Published in: Kleinhenz, Ursula (ed.): Interfaces in phonology. - Berlin: Akademie Verlag, 1996.
pp. 194-209. (studia grammatica 41)

Gaps in Word Formation

Renate Raffelsiefen (Freie Universität, Berlin)

0. Introduction[1]

Focusing on verbs derived by *-ize*-suffixation, I will argue that gaps in word formation result from the interaction of phonological wellformedness conditions and conditions on phonological transparency between derived forms and their bases. The relevant transparency conditions require identity of surface structure, including non-distinctive phonological features such as stress. While the data lend themselves to a description in terms of constraints, specifically the type of identity constraints introduced in McCarthy and Prince (1995), they cannot be described by models of the morphology-phonology interface in which morphological rules have access only to distinctive phonological information.

In section 1.1 I discuss some evidence that gaps in *ize*-formations reflect an irresolvable conflict between an identity constraint and a constraint against stress clash. In section 1.2 I present a constraint-based analysis of "truncation" in *ize*-formations arguing that, counter to Goldsmith's (1990) analysis, the patterns in question also reflect sensitivity to stress clash. In section 2 I investigate the interaction of identity constraints with the constraint against stress clash in English word formation involving suffixes other than *-ize*.

1. Gaps in *-ize*-formations

1.1 Cases which do not involve allomorphy

Table (1) lists a few of the several hundred attested verbs formed by suffixing *-ize* in English. Although the rule is generally productive, it almost never applies to monosyllabic or iambic words (cf. (1b)) ('X -> Y' means "X can be formed based on Y"):

(1)a.	rándomìze -> rándom	sálmonìze -> sálmon
	fóreignìze -> fóreign	sísterìze -> síster
	shépherdìze -> shépherd	rhýth[ə]mìze -> rhy'th[ə]m
b.	*Xize -> corrúpt	*Xize -> apt
	*Xize -> obscéne	*Xize -> firm
	*Xize -> políte	*Xize -> tense
	*Xize -> secúre	*Xize -> calm

[1] I thank an anonymous referee for comments and Bruce Straub for proof-reading.

*Xize -> arcáne *Xize -> bold

In Raffelsiefen 1992 a,b I argued that the gaps in (1b) reflect a dilemma: potential *ize*-formations would be either phonologically ill-formed due to stress clash (i.e. **corrúptize*) or the stem (i.e. the derived form minus the affix) would differ from the base in the position of the main stress (i.e. *córruptize *-> corrúpt*). That is, for a suffix with initial stress like *-ize-*, attachment to a base with final stress cannot yield any form which satisfies both constraints in (2):[2]

(2) *CLASH
Two adjacent stressed syllables are prohibited

IDENT
The stem of the derived word must be identical to the base.

Consider next the question of how to ensure that no output is preferred to a candidate which violates either *CLASH or IDENT. The approach to gaps proposed by Prince and Smolensky 1993 is to include the input (i.e. a structure in which affixes are unattached) in the candidate set, to posit a constraint "M-PARSE" which prohibits unattached affixes, and to rank that constraint below the constraints which cause the gap. The constraint M-PARSE is stated in (3):[3]

(3) M-PARSE
Morphemes are parsed into morphological constituents.

The gap illustrated in (1) is described by the constraint-ranking in tableau (4). The input consists of the affix *-ize* and a word which satisfies its syntactic subcategorization requirements (i.e. non-verbs).[4] The base is represented in its surface phonological form, which is crucial for proper evaluation w.r.t. the constraint IDENT:[5]

---

2 The undesirability of stress clashes was already pointed out by Liberman and Prince 1977.

3 Given my assumptions about word-basedness and concrete representations, I will follow the spirit but not the letter of the definition of M-PARSE and treat all and only unaffixed structures, specifically candidates which equal the input, as violations of that constraint. A better formulation of the constraint in question would be ATTACH defined as "Affixes must attach to their base".

4 Words with consonant-initial suffixes should also be excluded from the morphosyntactic domain of the suffix *-ize* since attested coinages like *cheerfulize* are very rare.

5 In her analysis of truncated forms like English [lær] 'Lar' from [læ.ri] 'Larry', Benua 1995 also emphasizes the need to refer to surface representations. Generally the vowels [æ] and [ɑ] are distributed such that [æ] appears before heterosyllabic [r] (e.g. [kæ.ri] 'carry') and [ɑ] appears before tautosyllabic [r] (e.g. [kɑr] 'car'). In order to account for the fact that the truncated form mimics the vowel in the base form (i.e. [lær] rather than [lɑr]) it is necessary to invoke an identity constraint which evaluates strings with respect to subphonemic features.
Non-identity of the syllable-final [r] in [lær] and the syllable-initial [r] in the base [læ.ri] is also found in pairs like *winte*[r]*ize* - *winte*[r] and raises the question of possible identity effects.

(4)	kərÁpt-áyz		IDENT	*CLASH	M-PARSE
		kɔ́rəptàyz[6]	*!		
		kərÁptàyz		*!	
	√	kərÁpt-áyz			*

For words with final stress such as *corrúpt*, the non-affixed candidate is optimal, because it is the only candidate which satisfies both IDENT and *CLASH.[7] As a result there is gap. For other words there is always a candidate which satisfies both IDENT and *CLASH, which means that a verb can be coined:

(5)	rǽndəm-ayz		IDENT	*CLASH	M-PARSE
	√	rǽndəmàyz[8]			
		rǽndəm-àyz			*

Note that "gap-causing" dilemmas cannot arise due to phonological constraints or identity constraints alone, but require that both types of constraints dominate M-PARSE. The phonological constraint causing such a gap will be one which is violated as a result of the attachment of the affix. The identity constraint requires identity in surface forms.

The constraint-ranking in (5) describes the *potential* for forming new verbs in *-ize*. Of course, it is possible to have stress clashes in actual words as shown in (6):

(6) báptìze dónàte
cápsìze rótàte
fránchìse gýràte

The claim that suffixing *-ize* to the adjective *apt* would be ungrammatical due to a stress clash (i.e. **áptìze*) is by no means inconsistent with the existence of the verb *báptìze*. The word *baptize* is not derived by suffixation and therefore is not subject to M-PARSE.[9] In general, there is no reason to expect the sound patterns characteristic of the potential words derived by a certain affix to conform to the prevailing sound patterns since they each are determined by independent constraint-interactions. As it happens, stress clashes are often unstable in English. This instability manifests itself in a tendency to reduce stress (cf. 7a) or in stress shift if a syllable precedes the clash within the same prosodic word (cf. (7b):[10]

---

6 The quality of the first vowel is based on the spelling of the word. If the written representation is not known to a speaker, the indeterminacy of that vowel might suffice to rule out the formation in question.

7 For the sake of completeness, M-PARSE must also be dominated by a constraint which requires phonological identity of the suffixes. Such a constraint prohibits a stressless vowel in the suffix.

8 The relative prominence patterns in the verb are determined by independent phonological constraints. Final stress in verbs of three or more syllables which end in at most one consonant is rarely primary (cf. *réconcìle, círcumcìse, cónstitùte* versus *cècomménd, cìrcumvént, ìnterrúpt*)

9 In fact, *baptize* is not even historically derived by suffixation in English but was borrowed from French (cf. French *baptiser*).

10 Stress shifts like in *abdómen* > *ábdòmen, anchóvy* > *ánchòvy* show that there are also developments

(7)a.	móbìle > móbile	b.	advértìse > ádvertìse
	dándrùff > dándruff		recógnìze > récognìze
	ábdòmen > ábdomen		amórtìze > ámortìze
	pérfèct > pérfect		staláctìte > stálactìte
	récòrd > récord		adúmbràte > ádumbràte

The replacement of quantity-sensitive stress by alternating stress is nearly complete with the result that the type of stress clash illustrated in (7b) is rare in Modern English. It is worth noting that the description in (4) does not link the gap illustrated in (1b) to the instability of stress clashes illustrated in (7b), in this respect converging with Goldsmith's 1990 analysis of that gap in terms of the filter in (8):

(8) Prohibited: * * foot
* * syllable
σ]# σ]

Goldsmith links the illicitness of a stress clash directly to the word-boundary, exempting both underived words and words derived by affixes associated with stem allomorphy from his filter. As will be shown in the next section, this claim is incorrect in that potential *-ize*-formations always depend on stress, independent of stem allomorphy.

Additional evidence showing that the sound patterns of productive *-ize* formations do not conform to the stress patterns of other verbs is presented in (9):

(9)a.	recíprocàte	b.	féderalìze
	certíficàte		hóspitalìze
	intércalàte		rádicalìze
	exásperàte		pérsonalìze
	elíminàte		partícularìze
	exácerbàte		repúblicanìze
	extrápolàte		Américanìze
	equílibràte		pópularìze

The examples in (9a) show the regular, strictly alternating stress patterns in English verbs. The occurrence of stress lapses (i.e. sequences of stressless syllables) in (9b) are best analysed as an identity effect as shown in (10):

---

which result in stress clash. This is because the constraint against stress clash is overriden by a higher-ranking constraint requiring antepenultimate stress in nouns and adjectives. Forms such as *ánchòvy* are, however, unstable due to the constraint against stress clash and we expect it to be replaced by *ánchovy* eventually (Cf. *balcóny* > *bálcòny* > *bálcony*).

(10)	fédərəl-áyz		IDENT	*CLASH	M-PARSE	*LAPSE
	√	fɛ́dərəlàyz				*
		fədɛ́rəlàyz	*!			
		fɛ́dərəl-áyz			*!	

The constraint *LAPSE is stated in (11):[11]

(11) *LAPSE
Two adjacent stressless syllables are prohibited

The fact that potential stress clashes, but not stress lapses, lead to gaps is accounted for by ranking M-PARSE between *CLASH and *LAPSE. The reason for ranking IDENT higher than *CLASH pertains to systematic preferences among the suboptimal candidates. While not accepting either form, native speakers consistently prefer *CLASH-violators like *corrúptìze* to IDENT-violators like *córruptìze* as derivatives based on words with final stress. These judgements are supported by the forms of the attested verbs in (12), which represent a few of the rare cases where coiners were apparently not deterred by creating a stress clash:

(12)		base:		
	bənálàyz	bənál	(cf.*bǽnəlàyz)	'banalize'
	rutíynàyz	rutíyn	(cf.*rúwtənàyz)	'routinize'
	bəbúwnàyz	bəbúwn	(cf.*bǽbənàyz)	'baboonize'
	kokéynàyz	kokéyn	(cf.*kówkənàyz)	'cocainize'

Should any of the formations in (12) gain currency in spite of their phonological ill-formedness, they are likely to undergo the stress shift illustrated in (7b) thereby adjusting to the regular sound patterns of the language. The question of whether or not such a stress shift occurs is, however, irrelevant to the ranking in (10) which describes the potential for creating new coinages in *-ize*. Similarly, the mechanisms for adopting loan words into English are irrelevant to that ranking, even if the etymological base of the loan word happens to exist as well. Some examples are given in (13):

(13)		
	kǽnəlàyz < French canaliser	'canalize'
	dívənàyz < French diviniser	'divinize'
	ímyənàyz < ?French immuniser[12]	'immunize'
	stáy(ə)làyz[13] < German stilisieren	'stylize'

[11] A similar constraint is also proposed by Kager (1994).

[12] The glide in English *immunize* suggests that the verb is perhaps not an adaptation of French *immuniser* but rather was coined in English. In the latter case we have to assume though, that the verb was originally pronounced *immúnìze* with subsequent stress shift.

[13] Interestingly, many speakers pronounce this verb with a syllabic *l* thereby avoiding the stress-clash.

The verbs in (13) are actual words in English but they are not potential words.[14] That is, a verb like *cánalìze* may well be borrowed into English but it could not possibly arise as a coinage based on *canál*.[15] The distinction between actual and potential word is a crucial prerequisite for any adequate description of the conditions characterizing the productive word formation patterns and is summed up in table (14):

(14)		Actual word?	Potential word?
	rúralìze (-> rúral)	Yes	Yes
	crúelìze (-> crúel)	No	Yes
	ímmunìze (-> immúne)	Yes	No
	sécurìze (-> secúre)	No	No

Only those *-ize*-formations which involve neither a stress clash nor non-identical stress patterns in the relation to their base are potential words.

1.2 Cases which involve allomorphy

The claim that new coinages in *-ize* must necessarily satisfy the constraint IDENT must be modified in view of the verbs in (15) all of which were coined in English:

(15)	émphasìze -> émphasis	sýnthesìze -> sýnthesis
	epénthesìze -> épenthesis	hypóthesìze -> hypóthesis
	parénthesìze -> parénthesis	metáthesìze -> metáthesis
	antíthesìze -> antíthesis	metástasìze -> metástasis

The formations in (15) follow the rules established this far in that they neither involve a stress clash nor do the derived forms differ from their respective bases with respect to the position of main stress. As was proposed by McCarthy and Prince (1995), identity constraints do not need to be monolithic but can target certain features of the pairs of strings to be evaluated. The illicitness of **córruptìze* based on *corrúpt* as opposed to the acceptability of *émphasìze* based on *émphasis* could be accounted for by modifying IDENT as follows ("S" = stress):

(16) IDENT(S)
Each stressed syllable in the base must correspond to a stressed syllable in the derived word

Consider next the question of why forms like *émphasìze* are preferred to forms which satisfy IDENT (i.e. **émphasisìze*). Coinages like *féderalìze*, *hóspitalìze*, etc. show that

---

[14] The point that the actual words of a language are not a subset of the potential words has been emphasized repeatedly by Aronoff (cf. Aronoff 1983). Needless to say, relations between actual words must also be described by the grammar to the extent that they are systematically recognized by the native hearer.

[15] This is because borrowed words are adopted as whole objects including affixes and therefore are not subject to the the constraint-rankings which describe the set of potential words.

violation of *LAPSE is not a crucial factor. Instead, the constraint pertains to syllabic structure and can be stated as follows ("O" = onset; "R" = rhyme):

(17) *O_iRO_i
Adajacent syllables must not have identical onsets.

The candidate **émphasisìze* violates the constraint in (17) because the vowel-initial suffix would cause the stem-final consonant to be in onset position. That constraint also plays a role in the cases of haplology listed in (18):

(18) *mRm

máximìze -> máximum	Cf. rádiumìze -> rádium
óptimìze -> óptimum	médiumìze -> médium
mínimìze -> mínimum	vácuumìze -> vácuum

*nRn

féminìze -> féminine	Cf. másculinìze -> másculine
phenómenìze -> phenómenon	skéletonìze -> skéleton

*tRt

áppetìze -> áppetìte	Cf. párasitìze -> párasìte

The constraint *O_iRO_i may seem dubious since it is often violated in English words. Some examples are given in (19):

(19)

*sRs	assassin, necessary, sausage, sesame, sissy
*mRm	mammoth, memory, camomile, mimosa, mimic, mammal
*nRn	anonymous, synonym, nanny, unanimous, banana
*tRt	tattoo, ratatouille, tatter, tentative, tetanus, tutu

As was discussed earlier, due to the interaction of phonological constraints and M-PARSE some sound patterns may not be allowable in potential words and yet be common in the vocabulary. Additional productivity gaps reflecting the dominance of *O_iRO_i over M-Parse are shown in (20):[16]

(20)

*Xt - ətɪ	*acutity, *completity, *obsoletity, *remotity
*Xf - əfay	*deafify, *toughify, *stiffify, *beefify, *roughify

---

16 The gaps in (i) reflect a related constraint prohibiting identical segments in the onset and the coda of the same syllable:

(i)

*Xl - əl	*appealal, *annulal, *exhalal, *assailal, *revealal, *instalal
*Xs - əs	*biasous, *atlasous, *tennisous, *mennaceous
*Xn - ən	*cleanen, *meanen, *plainen, *greenen, *keenen

The gaps in (20) reflect the ranking in (21):[17]

(21) IDENT, *O_iRO_i > M-PARSE

Provided that *O_iRO_i plays a role in the word formation patterns in (15) and (18), why is "truncation" not confined to the final consonant? The preference of *emphasize* over *emphasiize* can be accounted for by invoking the constraint in (22):

(22) *VV
Adjacent vowels are prohibited.

Additional evidence in support of the constraint *VV can be gleaned from the "truncation"-patterns in the native coinages in (23):[18]

(23)

mémorìze -> mémory	apóstrophìze -> apósthrophe
jéopardìze -> jéopardy	priórìtìze -> priórity
epítomìze -> epítomy	súbsidìze -> súbsidy
súmmarìze -> súmmary	análogìze -> análogy
fántasìze -> fántasy	párodìze -> párody
scrútinìze -> scrútiny	ecónomìze -> ecónomy
mahóganìze -> mahógany	énergìze -> énergy

The analysis of "truncation" is shown in tableau (24). I quote "truncation" because the word-formation patterns in question do not reflect a partial deletion of the base but rather the preference of one candidate over another. [19]

---

[17] A more comprehensive analysis of the suffixes would show that M-PARSE is in each case dominated n IDENT, but by more restricted identity constraints. Note also that the constraint ranking in (i) rules out n the actual nouns ending in *-ity* listed in (ia) nor the potential formations in (ib). The latter are not rule because the relevant onsets are non-identical.

(i)a. entity, identity, quantity, sanctity

b. chastity, vastity, augustity

[18] The fact that "truncation" in *-ize*-formations is always restricted to the final VC_0 sequence may indicate a constraint which requires all onsets in the base to correspond to identical onsets in the derived form.

[19] Clearly, more constraints are needed to account for the preference of *emphasize* over all competing candidates. For example, there is an undominated constraint ruling out any type of epenthesis in *-ize*-formations. The only potential counter-examples involve the ending *-ma* as shown below:

(i) cinematize -> cinema aromatize -> aroma
stigmatize -> stigma schematize -> schema

Since it is unlikely that the nasal in *-ma* conditions *t*-epenthesis phonologically, the data above are probably best described in terms of "correlative pairs" introduced in Marchand (1969) (i.e. X*matize* -> X*ma*).

(24)	émfəsɪs-áyz		IDENT(S)	*OiROi	*VV	*CLASH	M-PARSE
		émfəsɪsàyz		*!			
		émfəsɪàyz			*!		
	√	émfəsàyz					
		émfəàyz			*!		
		émfàyz				*!	
		émfəsɪs-áyz					*!

Both *OiROi and *VV are ranked higher than M-PARSE in order to account for the gaps illustrated in (25):[20]

(25)a. *Xìze -> crísis
*Xìze -> scépsis
*Xìze -> thésis
*Xìze -> básis
*Xìze -> cathársis
*Xìze -> neurósis
*Xìze -> ellípsis
*Xìze -> psychósis

b. *Xìze -> sílly[21]
*Xìze -> énvy
*Xìze -> mónkey
*Xìze -> assémbly
*Xìze -> attórney
*Xìze -> balóney

The examples in (25) differ from those in (15), (18), and (23) in that "truncation" of the final VC string leads to a stress clash. As a result the unaffixed structures are optimal. The same conditions account for the gaps in (25a):[22]

(26)a. *Xìze -> Híttìte
*Xìze -> línen
*Xìze -> hórror
b. Sémitìze -> Sémìte
cóttonìze -> cótton
vígorìze -> vígor
c. áppetìze -> áppetìte
féminìze -> féminine
hórrify -> hórror[23]

---

20 The constraint *VV may have to be restricted in view of the examples in (i):

(i)a. ghéttoìze -> ghétto
zéroìze -> zéro
échoìze -> écho
b. státuìze -> státue
vírtuìze -> vírtue
Zúluìze -> Zúlu

Alternatively, it might be possible to analyse the stem-final vowels in (i) as diphthongs (e.g. *ghétt*[ow]*ìze*) in which case *VV would not be violated. Even if it were argued that words like *silly* also end in a diphthong (i.e. *síll*[iy]), the acceptability of the formations in (i), as opposed to the gap in (25b), could be accounted for on purely phonological grounds.

21 There are a few attested verbs like *dándyìze*, *Tóryìze*, but they typically correspond to nouns in *-ism* (*dándyism*, *Tóryism*). Other counter-examples to my analysis, which correspond to *-ism* forms, are *clássicìze - clássicìsm*, *Márxìze - Márxism*. Perhaps these systematic exceptions can also be accounted for by positing a correlative pattern *Xize -> Xism*, cf. footnote 19.

22 The existence of the cognates *páralỳse* and *parálysis* does not contradict the analysis of the gap in (24) but rather confirms the claim that it is necessary to distinguish between actual and potential words. The verb *paralyse* was not coined in English but is an adaptation of the French loanword *paralyser*. See also *canonize* from French *canoniser*, *terrorize* from French *terroriser*.

*Xìze -> Lénin Stálinìze -> Stálin
*Xìze -> cándid líquidìze -> líquid

The dilemma in question is illustrated in tableau (27):

(27)	kəθársɪs-áyz		IDENT(S)	*O_iRO_i	*VV	*CLASH	M-PARSE
		kəθársɪsàyz		*!			
		kəθársɪàyz			*!		
		kəθársàyz				*!	
		kǽθərsàyz	*!				
	√	kəθársɪs-áyz					*

In tableau (27) *CLASH is ranked lower than the other phonological constraints in order to account for the intuition that *cathársìze* is the least objectionable of the affixed candidates. The preference of *CLASH-violators to the other suboptimal candidates has been noted before and is also confirmed by the exceptional coinage *synópsìze* based on *synópsis*.[24]

The claim that certain gaps arise due to violations of phonological constraints such as *O_iRO_i may seem to be contradicted by coinages such as *ruralize*, *memorize*, etc. The elimination of **crisisize* based on the *O_iRO_i violation versus the acceptability of *memorize*, which also violates *O_iRO_i, is explained in tableau (28):

(28)	méməri-áyz		IDENT(S)	*O_iRO_i	*VV	*CLASH	M-PARSE
		méməriàyz		*	*!		
	√	méməràyz		*			
		méməàyz		*	*!		
		méməàyz		*		*!	
		méməri-áyz		*			*!

Crucially, all relevant candidates based on *mémory* violate *O_iRO_i which means that none is eliminated because of this violation.

Tableau (28) is still inadequate in that it fails to express the fact that truncated *ize*-formations are acceptable only when truncation is needed to satisfy specific phonological constraints. The default status of full identity in the relation between the stems of *ize*-suf-

---

23 Cf. also the pairs *terrify - terror* versus *honorify - honor*.

24 It should be noted, however, that the non-truncated candidate is preferred to the *CLASH-violator if the latter has a monosyllabic stem. (cf. **crísisize* versus ***crísize*). That observation calls for a constraint requiring that minimally the first two syllables in the base correspond to identical syllables in the derived form. The claim is that the major problem with the candidate ***crísize* concerns the relation to the base: in contrast to the relation between **cathársize* and *cathársis*, the relation between ***crísize* and *crísis* is unrecognizable due to the insufficient amount of identical phonological structure (cf. also the candidates ***emphize* versus **emphasisize*).

fixations and their bases is expressed by ranking the original constraint IDENT stated in (2) lower than M-PARSE but higher than *LAPSE as is shown in (29):

(29) IDENT(S) > *O_iRO_i,*VV > *CLASH > M-PARSE > IDENT > LAPSE

The ranking in (30) accounts for the fact that *féderalize* is a potential coinage based on *féderal*, but **féderize* is not.

The *-ize*-formations discussed in this section differ from those discussed in section 1.1. in that they involve some type of stem allomorphy. Goldsmith analyses this difference in terms of strata: all *-ize*-formations involving segmental alternations or truncation are claimed to be derived at stratum 1, and are therefore not subject to the stress clash filter in (8). This analysis fails to account for the fact that truncation does not apply if it would result in a stress clash. By contrast, on the analysis presented here all forms in (30) are ruled out because of *Clash-violations, thereby describing the gap in a uniform manner:

(30)a.	*áptìze (-> apt)	b.	*corrúptìze (-> corrúpt)
	*síllìze (-> sílly)		*assémblìze (-> assémbly)
	*crísìze (->crísis)		*cathársìze (-> cathársis)

2. Other suffixes

2.1 The suffix *-ation*

The cognates in (31a) might give the impression that for the suffix *-ation*, IDENT ranks lower than M-Parse:

(31)a.	èxplanátion - expláin	b.	*rèmanátion -> remáin
	pèrturbátion - pertúrb		*dìsturbátion -> distúrb
	ìnspirátion -> inspíre		*dèsirátion -> desíre
	prèservátion -> presérve		*dèservátion -> desérve
	èxpirátion -> expíre		*rètirátion -> retíre
	àdorátion -> adóre		*ìgnorátion -> ignóre
	òbscuration -> obscúre		*sècurátion -> secúre
	ìnvitátion -> invíte		*dèlightátion -> delíght
	àdaptátion -> adápt		*àdoptátion -> adópt
	cònsultátion -> consúlt		*ìnsultátion -> insúlt

However, the unacceptability of the starred formations in (31b), all of which are formed on the basis of verbs which are phonologically similar to those in (31a), shows that the word-formation pattern is non-productive in Modern English.[25] In fact, acording to the

[25] The ungrammaticality of the nouns in (31b) cannot be explained by blocking since several verbs lack derived nominals other than X*ing* (e.g. *ignore*, *remain*, *secure*, *deserve*), whereas others coexist with their alleged blockers (e.g. *adaptation*, *adaption*; *perturbation*, *perturbance*, *perturbancy*; *accusation*,

OED, all nouns in *-ation* which are (etymologically) related to iambic verbs are borrowings (e.g. Engl. *explanation* < Latin *explanation-em*, Engl. *perturbation* < Old French *perturbacion* < Latin *perturbation-em*, etc.).[26] In English, *-ation*-suffixation is typically confined to specific "productivity niches", i.e. verbs ending in *-ate* or in the suffix *-ize*.[27] Sporadically, the suffix also applies to verbs which end in a stressless syllable or are monosyllabic. Some examples are given in (32):

(32)a. bòtherátion -> bóther
elìcitátion -> elícit
bàckwardátion -> báckward
flìrtátion -> flirt
stàrvátion -> starve
crìspátion -> crisp

Tableau (33) accounts for the unacceptability of the formations in (31b) and for the well-formedness of the examples in (32a):

(33)		rəméyn-éysən	IDENT	*CLASH	M-PARSE
		rɛ̀mənéysən	*!		
		rəmèynéysən		*!	
	√	rəméyn-éysən			*

So far the analysis is identical to that of the suffix *-ize* in (4). The two suffixes differ, however, in that *-ation* applies to monosyllabic bases, thereby incurring a violation of *CLASH. Conceivably, there is a significant phonological difference between the type of stress clash in *flìrtátion* and the types in **áptize*, **corrúptize* and **remàinátion*, in which case a reformulation of *CLASH would be called for. It is also possible that the examples in (32b) are simply exceptions, since they exhaust the list of attested cases.

The suffixes *-ize* and *-ation* also differ regarding productivity. Whereas for *-ize*, a candidate which violates neither IDENT nor *CLASH is (subject to semantic and pragmatic restrictions) likely to be coined, this is not true for *-ation*-suffixations like *?vòmitátion*, *?bàrterátion*, *?ìnjurátion*, etc. Perhaps this difference is due to the fact that the suffix *-ize* subcategorizes for nouns and adjectives, which typically do not end in a stressed syllable, whereas *-ation* subcategorizes for verbs, which predominantly end in a stressed syllable. As a result the potential domain for the suffix *-ize* is much larger, allowing it to "gain momentum" , whereas the productivity of the suffix *-ation* is stifled.

2.2 Fused suffixes

Aronoff (1983) drew attention to the fact that, for words ending in the suffix *-able* the productivity of the suffix *-ity* clearly exceeds the productivity of its rival *-ness*, whereas

---

accusal). For more discussion, see Raffelsiefen 1992a,b.

26 The only exception is the noun *indentation* which, however, also has a corresponding form in French (i.e. French *indentation*).

27 The ending *-ate* is always truncated in verbs (e.g. *séparàte - sèparátion*, *álternàte - àlternátion*, etc.). The phonological motIvation for truncation in these cases is perhaps to avoid identical nuclei in adjacent syllables.

the opposite holds for words ending in the suffix *-ive*. As for the lack of productivity of the suffix *-ness* w.r.t. *-able*-suffixations, Aronoff concludes that "a purely formal study would never have given the slightest suggestion that this could be the case, for there are no formal grounds for doubting the viability of this particular WFP" (Aronoff 1983:169f). However, the suffix-combinations *-ability* and *-ableness* differ in that the former is phonologically more fused than the latter. The claim that this type of fusion is crucial for the preference of *-ability* is described in the following scheme, where X is a word:[28]

(34) If the relation $Xsuf_isuf_j$ - X is phonologically transparent,
and the relation $Xsuf_isuf_j$ - $Xsuf_i$ is phonologically non-transparent,
then for each word $Ysuf_i$ - there is a word $Ysuf_isuf_j$.

While describing the preferences in (35a), the scheme in (34) allows also for the preferences in (35b), since **terr*, **horr* and **miser* are not words:

(35)a.	drinkability, *drinkableness	b.	terribleness, *terribility
	derivability, *derivableness		horribleness, *horribility
	developability, *developableness		miserableness, *miserability

The scheme in (34) also accounts for the fact that the suffix *-ation* is preferred to the generally more productive suffix *-ment* for words derived by *-ize*-suffixation:

(36) ràndomizátion - rándom is phonologically transparent
ràndomizátion - rándomìze is phonologically non-transparent
therefore for each word Y*ize*, there is a word Y*ization*.

The crucial factor in the description of productivity according to (34) is identity versus non-identity of phonological surface form in the relation between words. Aronoff's observation that the suffix *-ity* failed to develop productivity with respect to the suffix *-ive* reflects the relevance of another factor, namely stress clash. Crucially, the suffix-combinations in (37a) and (37b) differ in that the latter have initial stress:

(37)a.	abel+ity: abílity	b.	ive+ity: ívity
	ize+ation: izátion		ment+al: méntal
			ist+ic: ístic

Whereas stress clash never affects any words derived by the fused suffixes in (37), there are potential dilemmas for the suffixes in (b). The unacceptability of forms like *obsessivity* is described by the familiar constraint-ranking in (38):

---

[28] The scheme in (34) describes phonological conditions for morphological wellformedness. The question of whether words are actually coined depends as always on semantic and pragmatic factors.

(38)	ɔbsés-ívɪti		IDENT	*CLASH	M-PARSE
		ɔ̀bsəsívɪti	*!		
		ɔbsɛ̀sívɪti		*!	
	√	ɔbsés-ívɪti			*

The same ranking also accounts for the wellformedness patterns in (39). Since the suffix *-ive* only attaches to verbs with final stress, there are no data to compare to the *-ivity*-formations in (a). The data in (39c) are adopted from Strauss:[29]

(39)a. *Xívity -> invént/ive
*Xívity -> abúse/ive
*Xívity -> caréss/ive

b. *Xméntal -> emplóy/ment — devèlopméntal -> devélop/ment
*Xméntal -> discérn/ment — gòvernméntal -> góvern/ment
*Xméntal -> contáin/ment — àrguméntal -> árgue/ment

c. *Xístic -> cartóon/ist — fàtalístic -> fátal/ist
*Xístic -> escápe/ist — règalístic -> régal/ist
*Xístic -> deféat/ist — hùmorístic -> húmor/ist

The fact that suffix fusion is much more common for *-istic* than for *-ivity* or *-mental* can perhaps also be explained by the fact that the latter two fused suffixes subcategorize for verbs, which typically have final stress, whereas *-istic* subcategorizes for nouns and adjectives, which typically end in a stressless syllable. As was mentioned before, this factor affects the size of the overall domain of a suffix, possibly delimiting their potential to gain momentum.

2.3 The suffixes *-ee*, *-ese* and *-eer*

At this point it may appear that the constraint ranking in (38) describes the phonological properties of gaps for all English suffixes with initial stress. It would then be unnecessary to specify the relevant constraint-ranking for each suffix, although other morphophonological properties like "truncation" might still require individual specification. However, there is some evidence that stressed suffixes may differ with respect to the ranking of the constraints IDENT, *CLASH and M-PARSE. Consider the native coinages in (40):

(40) assìgnée -> assígn — appòintée -> appóint
divòrcée -> divórce — detàinée -> detáin
advìsée -> advíse — selèctée -> seléct

[29] A more detailed discussion of these examples is given in Raffelsiefen (1992a,b). For an alternative analysis of the gap in (39b) see Aronoff (1976) and for an alternative analysis of the gap in (39c) see Strauss. Goldsmith (1990) extends his stress clash filter analysis also to both the gaps in (39b) and (39c).

The stress pattern of the *-ee*-formations does not occur in underived words and can only be interpreted as an identity effect. The tableau in (41) describes the patterns:[30]

(41)	əsáyn-íy		IDENT	M-PARSE	*CLASH
	√	əsàyníy			*
		æ̀səníy	*!		
		əsáyn-íy		*!	

The low ranking of *CLASH w.r.t. IDENT and M-PARSE is a unique property of the suffix *-ee*. The patterns are similar to the cases discussed earlier in that IDENT dominates *CLASH. However, as is shown by the native coinages in (42), there is at least one suffix for which *CLASH dominates IDENT:[31]

(42) Tàiwanése -> Taiwán Sùdanése -> Sudán
Nèpalése -> Nepál Viètnamése -> Vietnám

Tableau (43) shows the ranking describing *-ese*-formations:

(43)	taywán-íyz		*CLASH	M-PARSE	IDENT
	√	tàywəníyz			*
		taywə̀níyz	*!		
		taywán-íyz		*!	

Not all monosyllabic suffixes with an initial high, tense vowel differ from the suffixation patterns observed in the preceding sections. The suffix *-eer* shows the same restrictions as the stress-initial suffixes discussed above in that the suffix does not combine with words with final stress:

(44) mùffinéer -> múffin *Xéer -> bagúette
jàrgonéer -> járgon *Xéer -> strike
màrketéer -> márket *Xéer -> fair

Tableau (45) accounts for the gap in (44).

(45)	bəgét-íyr		IDENT	*CLASH	M-PARSE
		bæ̀gətíyr	*!		
		bəgə̀tíyr		*!	
	√	bəgét-íyr			*

[30] Relative prominence in the candidates in (41) is determined by an independent constraint which requires that tense high vowels in word-final syllables carry main stress.

[31] For *ese*-formations, the constraint *O_iRO_i also dominates IDENT as is shown by examples like *Lèbanése -> Lébanon* versus *Pèntagonése -> Péntagon*, *Àragonése -> Áragon*, etc.

Interestingly, tableau (45) also accounts for *-ese*-formations based on words other than names. That is, new coinages are generally possible only when there is a candidate which satisfies both IDENT and *CLASH[32].

3. Conclusion

In this paper I have presented evidence that English word-formation patterns are determined by the interaction of phonological constraints and constraints which require identity of surface phonological structure between a derived form and its base. Specifically, certain systematic productivity gaps appear to be manifestations of irresolvable conflicts between these two types of constraints. The analysis raises the question if not the entire morphology-phonology interface can be described in terms of conflicts between phonological constraints and identity constraints. Such an analysis would be preferable to previous descriptions in that it would refer to a single level of phonological representation, which is furthermore amenable to direct observation.

4. References

Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Linguistic Inquiry Monograph no. 1. Cambridge, Mass. M.I.T. Press.

Aronoff, Mark. 1983. Potential words, actual words, productivity and frequency. In S.Hattori & K. Inoue (eds.). *Proceedings of the XIII. International Congress of Linguists*. Tokyo: Permanent International Committee on Linguistics.

Benua, Laura. 1995. Identity effects in Morphological Truncation. In J. Beckman, S. Urbanczyck and L. Walsh, eds. *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*, Graduate Linguistics Student Association, Amherst, Massachusetts.

Goldsmith, John. 1990. *Autosegmental and Metrical Phonology*. Basil Blackwell, Oxford.

Kager, René. 1994. Ternary Rhythm in Alignment Theory. Ms. Utrecht.

Liberman, Mark and Alan Prince. 1977. *On Stress and Linguistic Rhythm*. Linguistic Inquiry 8, 249-336.

Marchand, Hans. 1969. *The Categories and Types of Present-day English Word-Formation*, 2nd edition, Beck, München.

McCarthy, John J. and Alan S. Prince. 1995. Faithfulness and Reduplicative Identity. In René Kager, Harry van der Hulst and Wim Zonneveld (eds.) *Proceedings of the OTS/HIL Workshop in Prosodic Morphology*.

Prince, Alan and Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in generative grammar*. Ms. Rutgers University and University of Colorado, Boulder.

Raffelsiefen, Renate. 1992a. A Nonconfigurational Approach to Morphology. In Mark Aronoff (ed.) *Morphology Now*. State University of New York Press.

Raffelsiefen, R. 1992b. *Relating Words: A New Approach to English Morphology*. University of Washington doctoral dissertaion.

Strauss, Steven L. 1983. Stress Assignment as morphological adjustment in English. *Linguistic Analysis* 11, 419-27.

---

[32] Cf. jàrgonése, jòurnalése, compùterése, nòvelése, càblése, translàtionése, etc.