Change and variation in morphonotactics

Wolfgang U. Dressler,[a] Katarzyna Dziubalska-Kołaczyk[b] and Lina Pestal[a]

[a]Department of Linguistics, Vienna University [b]School of English, Adam Mickiewicz University Poznań

In this contribution we discuss diachronic and variationist aspects of morphonotactics, a new research field that we have tried to establish over the last years (cf. Dressler & Dziubalska-Kołaczyk 2006). Morphonotactics is the area of interaction between morphotactics and phonotactics and represents a subfield of morphonology, which in turn is the area of interaction between morphology and phonology (cf. Dressler 1985, 1996). We claim that in this interaction morphotactics typically creates phonotactically marked structures which occur never or only exceptionally in monomorphemic words. In our contribution we deal with typical diachronic changes. Our claim about the markedness of morphonotactic sequences is tested mainly against data from Polish, Lithuanian and other Balto-Slavic languages. Our theoretical basis draws on models of Natural Phonology (cf. Hurch & Rhodes 1996, Dziubalska-Kołaczyk & Weckwerth 2002) and Natural Morphology (cf. Dressler et al. 1987, Kilani-Schoch & Dressler 2005), and especially on the subtheories of universal markedness (or universal preferences) and of typological adequacy.

Keywords: Balto-Slavic languages, phonotactics, morphonotactics, consonant clusters, markedness

1. Introduction

Phonotactic preferences hold for the basic forms of monomorphemic words; the less they are respected, the more marked phonotactic sequences arise. Morphonotactic sequences, on the other hand, which are due to morphological operations of inflection or word-formation, are much more likely to be marked. This contribution focusses on such morphonotactic consonant clusters.

Folia Linguistica 31 (2010), 51–67. DOI 10.1515/FLIH.2010.003
ISSN 0165-4004, E-ISSN 1614-7308

We define a morphonotactic consonant cluster as a cluster which comes into being only through morphotactic operations. The most obvious example of such an operation is probably concatenation, as in English *seem+ed*. Here, the word-final cluster /md/ can occur only in morphologically complex words, but never in monomorphemic ones. However, also other morphotactic operations may be responsible for morphonotactic consonant clusters. For example, the Polish word-initial cluster /lv/ in *lw-i* is due to adjective formation from *lew* 'lion' with morphological vowel deletion, and the cluster /lv/ does not occur initially in morphologically underived words. Since they do not occur in monomorphemic words, morphonotactic clusters indicate that the words exhibiting them are morphologically derived. Thereby, purely morphonotactic clusters facilitate morphological processing in perception, as we have shown for German in still unpublished experiments.

In addition to purely morphonotactic clusters, there are also clusters which are morphonotactic only by default. That is to say, they do occur in monomorphemic words, but only exceptionally. Also such default clusters facilitate morphological processing, although they signal the morphological derivedness of the words in which they occur not unambiguously, but only with a high probability. A case in point is the English word-final cluster /ps/, which nearly always contains a morpheme boundary (as in *lap+s*, *ape+s*), except in very rare examples of monomorphemic words such as *lapse*, *apse*, or *glimpse*.

In Beats-and-Binding phonotactics (see Dziubalska-Kołaczyk 2002, Dziubalska-Kołaczyk & Krynicki 2007, Dziubalska-Kołaczyk 2009, Dziubalska-Kołaczyk & Zielińska 2010, in press) the stability and preservation of consonant clusters is derived from universally required relationships among the perceptual contrasts between the involved consonants on the one hand, and those between the consonants and neighbouring vowels on the other hand. Since clusters are generally dispreferred, they can only survive, if they are sustained by some force counteracting the overwhelming tendency to reduce towards CVs. We take perceptual contrast to be such a force and model it in terms of Net Auditory Distance (NAD). On our definition, NAD is the sum of the distances between segments with respect to manner of articulation (MOA) and place of articulation (POA). On the basis of NAD, various preferences can be specified, which allow for the preservation of clusters, if respected. For instance, a word initial cluster

Table 1. Net Auditory Distance for Polish

Polish	Number of clusters that apply	Number of clusters that meet the preferences	Perc.
NAD (C1,C2) ≥ NAD (C2,V)	5,000	2,453	49.06
Morphologically complex	162	41	**25.31**
Morphologically simplex	4,838	2,412	**49.86**

of two consonants is predicted to be (relatively) preferred and therefore stable, if the net auditory distance between the first and the second consonant exceeds that between the second consonant and the following vowel.

For monomorphemic Polish and English words we have found that the phonotactic preferences which NAD specifies for clusters are respected to a moderately high degree (approximately 70%, cf. Dziubalska-Kołaczyk & Krynicki 2007). When one compares simplex and complex Polish words with respect to the phonotactic preference for word-initial biphonemic consonant clusters which the Net Auditory Distance principle specifies and which was introduced above, the following picture emerges (see Table 1).[1] As one can see, morphologically simplex Polish words observe the preference to a much higher degree than complex ones. So much for phonotactics.

As far as morphology is concerned, an initial basic premise of our argument is a universal preference for concatenative morphology. It represents a secondary preference derived from the universal preference parameters of iconicity and morphotactic transparency. From such a preference for concatenative morphology we derive that languages with a low degree of morphotactic complexity will exhibit only such morphonotactic phenomena as result from pure morphological concatenation. This is indeed true for languages such as English (e.g. preterites *scream-ed*, *liv-ed*, *robb-ed*)

[1] The preference was tested on 5,000 CCV clusters classified manually into 162 containing morphological boundary and 4,838 not containing any morphological boundaries. The procedure of extracting the list of 5,000 consisted in phonetically transcribing 120 000 entries of Great PWN dictionary which were subjected to semi-automatic heuristics (removing words with derivational morphemes and potential compounds which resulted in 13,691 words), cf. Dziubalska-Kołaczyk & Krynicki (2007).

or German (e.g. second person singular forms *lach-st*, *stopf-st*, *qualm-st*) or Italian (e.g. derivational prefixations *s-radicare*, *s-gridare*, *s-frenare* < Latin prefix *ex-*). In terms of diachronic phonology, the complexity of these consonant clusters is due to vowel loss. Typologically speaking, these Indo-European languages are all weakly inflecting languages.

In contrast, more strongly inflecting languages may exhibit also morphonotactic sequences resulting from morphotactic modifications such as synchronic morphological vowel deletion – in addition to clusters brought about through pure concatenation. An example is Latin, where the word-initial sequence /spr/ occurs only in the perfect and past passive participle *sprē-v-i*, *sprē-tus* from present *sper-n-o* 'I despise'.

Thus, the diachronic genesis of morphonotactic consonant clusters presupposes, on the one hand, a sufficient degree of morphological complexity. On the other hand, we expect morphonotactic clusters to arise only in languages with sufficiently complex phonotactics and one typical diachronic source is phonological vowel deletion. This paper demonstrates the plausibility of our expectations by giving an historical overview of the development of (mor)phonotactics from reconstructed Proto-Indo-European up to modern Balto-Slavic languages.

2. Polish

Let us start with a case of Polish prefixation and provide examples of morphonotactic clusters which arise through prefixation of sibilant-initial words with the prefix *w-*, phonetically [v] or, with assimilatory devoicing, [f].

Of the word-initial *ws-* [fs-] cluster, only morphologically conditioned examples exist, but no monomorphemic ones. *wsz-* [fʃ] occurs in the fossilized but frequent prefixoids *wsze*, *wszech*, *wszem* 'all, everybody' (with diachronic loss of the morpheme boundaries in what were originally case forms), in archaic *wszędy* 'everywhere', in frequent *wszystko* 'everything', and in archaic *wszak* 'after all'. *wsi-* [fɕ] appears in the Russian loan *wsio* 'everything' and in the colloquial pronunciation of the abbreviation *WSJO* from the recent term *Wyższa Szkoła Języków Obcych* 'college of modern languages'.[2]

[2] Examples from Dubisz's (2006) dictionary.

All other instances of the three initial clusters are morphonotactic. The first group consists of words with the prefix *w-* 'in', as in the verbs *w+sypać* 'to pour', *w+szyć* 'to sew in' (plus three other items), *w+siać* 'to sow in'. When comparing monomorphemic and bimorphemic clusters we can see that *wsi-* is a morphonotactic cluster by default, whereas *wsz-* is not.

The three double-consononat clusters discussed above are also part of triple-consonant clusters. The highly marked clusters *wsp-*, *wst-*, *wści-* appear in the following monomorphemic words: *wspaniale* 'splendid', *wspak* 'backward'; *wstążka* 'ribbon' (plus three other words); *wściekać* 'to get furious' (with 22 derivatives; all the words had morpheme boundaries formerly).

Comparable morphonotactic clusters span either one or two morpheme boundaries, for example, two in *w+s+kazać* 'to point' (plus 13 other items), one in *w+skoczyć* 'to jump in' (plus 29 other items) and *ws+pomagać* 'to help' (plus five other items). Another two-morpheme-boundary cluster [fsx-] occurs in *w+s+chodzić* 'to rise' and *w+s+chód* 'east, sunrise', and another one-morpheme-boundary cluster in *w+szczepić* 'to implant' (plus seven other items). Thus, the morphonotactic character of those triple clusters is only a weak default.

All comparable word-initial quadruple clusters are morphonotactic, with one morpheme boundary in *ws+tręt* [fstr-] 'disgust' (plus two derivatives) and *w+strzelać* 'to shoot in' (plus five other items), and two morpheme boundaries in *w+s+trząsać* 'to shake' (plus eight other items).

Diachronically speaking, these clusters have resulted from vowel deletion, because the Polish prefixes *w-* and *s-* and their cognates in the other Slavic languages go back to Proto-Slavic **wu-*, **su-* via the reduction of these vowels and their subsequent loss in rhythmically weak positions (cf. Walczak 1999: 45–46, Respond 2003: 120).

As a strongly inflecting language, Polish has also morphonotactic clusters with a non-concatenative source, i.e. vowel deletion. Examples with deletion of the first-syllable root vowel in adjective formation are: *wieś* 'village', adj. *wsiowy*, *len* 'linen'–*lniany*, *lew* 'lion'–*lwi*, *mech* 'moss'–*mchowy*, *wesz* 'louse'–*wszawy*, or the comparative of the adverb *lekko* 'light' *–lżej*. The same deletion also applies in inflection: masculine *len*, Gen.Sg. *ln-u*; *mech*, *mch-u*; feminine *wieś*, *ws-i*; *wesz*, *wsz-y*. This synchronically morphological vowel deletion goes back to the phonological vowel deletion of rhythmically weak ultrashort, high front or back vowels in the eleventh century (the so-called 'jer' and 'jor', cf. Walczak 1999: 45–46).

3. Slovak and Russian

Let us turn now to another West Slavic language, namely Slovak (cf. Dvonč et al. 1966). Looking again at the prefix *v-*, we see that its prefixation sometimes creates morphonotactic clusters, as in: *v+biť* 'to hit on', *v+bok* 'on the side', *v+padnúť* 'to fall in', *v+pokon* 'later', *v+posled* 'earlier', *v+pred* 'forwards', *v+pravo* 'on the right side', *v+plyv* 'influence', *v+stať* 'to get up', *v+stup* 'entrance'.

Clusters that are morphonotactic by default occur in: *v+dýchnuť* 'to inhale', etc. (but cf. monomorphemic *vdova* 'widow'); *v+členiť* 'to include', *v+čas* 'in time', etc. (but cf. monomorphemic *včera* 'yesterday', *včela* 'bee'); *v+klad* 'inlet', *v+kliniť* 'to wedge in', etc. (but cf. monomorphemic *vkus* 'taste'). In contrast, the word-initial clusters /vl, vn, vr/ correspond to normal, well-formed phonotactic combinations, and the same is true of /ft, fʃ, vz/.

In comparison with Polish, morphological vowel deletion has often been analogically leveled in Slovak, e.g. *mach* 'moss', Gen. *mach+u*, adj. *mach-ový*, *ľan* 'linen', Gen. *ľan+u*, adj. *ľan-ový*, *lev* 'lion', Gen. *lev+i*, Adj. *lev-í*, *ľahký* 'light' comparative *ľahš-í*.

Also in Russian (cf. Townsend 1975) there are similar word-initial morphonotactic consonant clusters, as in: *v+begať* 'to run into', *v+biť* 'to hit into', *v+birať* 'to suck in', *v+blizi* 'nearby', etc.; *v+gibať* 'to bend into', *v+glub'* 'into the depth', *v+gljadetsja* 'to regard closely'; *vz+kriknuť*; *vz+ryť* 'to dig in', *vz+ryxiť* 'to loosen'; *vz+letať* 'to fly up', *vz+lom* 'burglary', etc.; *vz+max* 'swing', *vz+more* 'coast', etc.; *vz+nos* 'payment', *vz+nuždať* 'to bridle'.

Also in Russian, *v+d* is morphonotactic only by default, because of *vdova* 'widow'.

As far as word-initial geminates are concerned, all of them are morphonotactic, e.g. *v+verx* 'up', *v+vesti* 'to bring in', *v+vod* 'setting to work', *v+vedenie* 'introduction'.

As to clusters originating in root vowel deletion, Russian has several of them, but each of them occurs only in a single root. Since vowel deletion creates the same clusters nowhere else, they are purely morphonotactic but at the same time isolated and not involved in any further generalizations that might facilitate morphological processing. Thus, on the one hand, the isolated word-initial clusters /mx/ (Gen. *mx-a*) and /mʃ/ (adj. *mš-istnyj*) are derived from *mox* 'moss'; the cluster /lʲv/, as in Gen. *ľv-a* and in the derivations adj. *ľv-inyj*, *ľv-ica* 'lioness', *ľv-ënok* 'young lion' is derived from

lev 'lion'; the cluster /lʲd/, as in Gen. *ľd-a*, compounds *ľd-o-X* is derived from *lëd* 'ice'; the cluster /lb/, as in Gen. *lb-a* is derived from *lob* 'forehead'; the cluster /lʒ/, as in Gen. *lž-i*, in the derivations *lž-ec* 'liar', *lž-ivnyj* 'mendacious', and in compounds *lž-e-X* 'pseudo-X' is derived from *lož'* 'lie'. On the other hand, there are several word-initial monomorphemic /rv/ and /rʒ/ clusters vs. Gen. *rv-a*, etc. from Nom. *rov* 'ditch' (e.g. *rvať* 'to tear'), Gen. *rž-i*, adj. *rž-anoj* from Nom. *rož'* 'rye' (e.g. *rža* 'rust').

4. Lithuanian inflection

Let us now move to the closely related Baltic languages, and focus on Lithuanian. In Lithuanian, many word-final consonant clusters are of a purely morphonotactic nature. Our first examples are second-person singular imperatives in /k/, as in: *dirb+k* 'work!', *temp+k* 'bend!', *megz+k* 'knit!'; *skris+k* 'fly!'; *lauž+k* 'break!', *blokš+k* 'give a blow!'; *im+k* 'take!', *nurim+k* 'calm down!'; *gel+k* 'sting!', *kel+k* 'lift!'; *ar+k* 'plough!', *bar+k* 'scold!'. The parallel word-final cluster /nk/, as in *aiškin+k* 'explain!', *sodin+k* 'plant!', is morphonotactic only by default, because there exists monomorphemic *link* 'towards' (plus its derivations).

Another set of word-final morphonotactic clusters is formed by third-person futures in *-s*, as in: *kep+s* 'he/she/they'll bake', *dirb-s* ''ll work'; *kel+s* ''ll lift', *gel+s* ''ll sting'; *gin+s* ''ll defend', *sodin+s* ''ll plant'.

Clusters that are morphonotactic by default occur in: *im+s* ''ll take', *nurim+s* ''ll calm down' (because of the dative plural suffix *-ms*); *ar+s* ''ll plough', *bar+s*,'ll scold' (because of *nors* 'though'); *šok+s* ''ll jump', *aug+s* ''ll grow' (because of *koks* 'which' with its derivatives).[3]

More complicated is the case of verb roots ending in *-š*, *-ž*. Due to concatenative morphology, root-final consonants never occur word-finally. The two sibilants are the only exceptions, and occur word finally only when they are followed by the future suffix *-s*. The sequence /ʒ+s/ is assimilated to intermediate /ʃs/, which is phonotactically prohibited and simplified to /ʃ/. Therefore, the third person futures of *lauž+ti* 'to break', *blokš+ti* 'to give a blow' are *lauš*, *blokš*. The resulting cluster /rʃ/ in *terš-ti* 'to make dirty', Fut. *terš* is morphonotactic only by default because of the preposition *virš* 'over'.

Since Lithuanian is a strongly inflecting language with one of the most

[3] We don't include interjections such as *tvoks*, because interjections are extragrammatical formations which may easily differ from the phonological system of a language.

complex morphologies of all living Indo-European languages, we expect to find morphonotactic clusters that originate in morphological modifications involving non-concatenative morphology. As far as word-final clusters are concerned, we find only cases in which non-concatenative and concatenative morphological operations are combined. This occurs when concatenation involving a consonant results in metathesis. This is the case in the root-final cluster /zg/: *mezg-a* '(s)he/they is/are knitting', infinitive /mezg+ti/ → *megz+ti* [meksti], and analogically future /mezg+s/ → [megz-s → meks].

A more frequent constellation occures when the genitive singular case suffix *-s* is attached to the results of unproductive stem alternation. The first, isolated case is the Gen.Sg. *obel+s* of the *i*-stem *obeli+s* 'apple tree'. Next there are three genitives of derived *r*-stems: *sesuo* 'sister', Gen.Sg. *seser+s*, *dukte* 'daughter', Gen.Sg. *dukter+s*, *moteri-s* (earlier: *mote*) 'woman', Gen.Sg. *moter-s*. Finally there are two dozen derived *n*-stems, e.g. *akmuo* 'stone', Gen. Sg. *akmen+s*, *piemuo* 'shepherd' – *piemen+s*, *vanduo* 'water' – *vanden+s*, as well as isolated *šuo* 'dog', Gen.Sg. *šun+s*. There is a new trend in colloquial Lithuanian to replace these genitive forms with productive variants without a consonant cluster, e.g. *obel+s* > *obel+ies*, *moter+s* > *moter+ies*, *šun+s* > *šun+io*, *piemen+s* > *piemen+io* (Ambrazas et al. 1996: 79, 80).

5. Latvian

Latvian word-final phonotactics differs very strongly from Lithuanian, because the loss of unstressed last-syllable vowels and morphological analogies have given rise to many word-final consonant clusters in citation forms, e.g. Nom.Sg. *cilvek-s* 'man', *akmen-s* 'stone', *sun-s* 'dog', *augstum-s* 'height', *rag-s* 'horn', *pil-s* 'castle', *vaciet-s* 'German', *av-s* 'sheep'. These sound changes appear to be more important for explaining the apparent lack of purely morphonotactic word-final consonant clusters than the more weakly inflecting character and the less complex morphology of Latvian in comparison with Lithuanian.

6. Lithuanian compounding

In compounds, Lithuanian presents a marked contrast to Slavic languages. Slavic languages (like many other Indo-European languages) typically

insert a vocalic interfix (a linking vowel) between the two members of a compound (e.g. Pol. *teatr-o-logia*, *tor-o-mistrz*, *towar-o-znawstwo*). Thus, they avoid the creation of new consonant clusters. Lithuanian, however, deletes the thematic vowel of the first constituent of a compound in at least half of all cases.[4] Thus Lithuanian compounding creates new word-internal, morphonotactic consonant clusters, e.g. *juod+varnis* 'black raven' ← *juodas* 'black' and *varnas* 'raven'. Since the only monomorphemic words in which the cluster /dv/ occurs word-internally are Latinate loan-words such as *adventas* and *advokatas*, it is a morphonotactic cluster only in a weak sense, i.e only among native words or as a default. We will nevertheless regard such compounds as containing morphonotactic clusters, also if the same clusters arise by prefixation and suffixation.[5]

Of course, as in any other highly inflecting-fusional Indo-European language system or subsystem (e.g. the Latinate part of English morphology), interfixation exists also in Lithuanian, but only as a minor, subregular operation. Still, in some cases it produces more token-frequent variants of vowel-deletion cases. For example:

didž-ia+galv-i-s 'animal with a big head' ← *did-i-s* 'big' + *galv-a* 'head'
plač-ia+burn-i-s 'broad-mouth(y)' ← *plat-i* 'broad' + *burn-a* 'mouth'
kryži-a+snap-i-s 'crossbill' ← *kryži-u-s* 'cross' + *snap-a-s* 'beacon'

are the more frequent variants of the respective variants with vowel deletion below.

The interfixation of a vowel in Slavic and other Indo-European languages is a means for avoiding the creation of marked consonant clusters. Indicatively, Lithuanian vowel deletion tends to lead to the creation of unmarked clusters.

First, it needs to be pointed out that, with regard to Dziubalska-Kołaczyk's (2009) Net Auditory Distance principle, and in contrast to what seems to be the case in word-initial and word-final position, clusters in medial position appear to be preferred when the NAD between them is small.[6] Therefore, one would expect word-medial morphonotactic clusters

[4] Whereas this option is only a minor pattern in German, Swedish, etc.
[5] We do not consider onomatopoetic words, because they may contain clusters which violate normal phonotactics.
[6] More precisely, the preference is for smaller NAD between the consonants than between each of them and the neighbouring vowels on both of their sides. Of course, the NAD between the two consonants should still be greater than zero.

to involve clusters with a relatively great NAD between their constituents, so that they can be easily distinguished as clusters that do not occur morpheme-internally. In Lithuanian compound formation, however, as well as in its derivational morphology, also purely morphonotactic clusters seem to respect the preference for a small NAD. If we look at clusters with miminal NAD, then we find that the majority of them are morphonotactic clusters, while language-specifically well-formed phonotactic clusters represent only a minority. Thus, among occlusive clusters only the three clusters /bd, gd, kt/ are phonotactically well-formed ones in Lithuanian, whereas all of the following nine clusters are purely morphonotactic:[7]

/bg/: *varp*(*a*)+*gal-i-s* 'end of a ear' ← *varp-a* 'ear' + *gal-a-s* 'end'
/pk/: *duob*(*ia*)+*kas-y-s* 'grave-digger' ← *duob-ė* 'pit' + *kas-ti* 'to dig'
/db/: *balt*(*a*)+*barzd-i-s* 'white-beard' ← *balt-a-s* 'white' + *barzd-a* 'beard'
/tp/: *rud*(*a*)+*plauk-i-s* 'brown-haired' ← *rud-a-s* 'brown' + *plauk-ai* 'hair' (Pl.)
/dg/: *did*(*žia*)+*galv-i-s* 'animal with a big head' ← *did-i-s* 'big' + *galv-a* 'head'
/tk/: *led*(*a*)+*kaln-i-s* 'iceberg' ← *led-a-s* 'ice' + *kaln-a-s* 'mountain'
/gb/: *pilk*+*balt-i-s* 'grey-white' ← *pilk-a-s* 'grey' + *balt-a-s* 'white'
/kp/: *rug*+*pjūt-i-s* 'August' ← *rug-y-s* ,rye' + *pjūt-ė* 'mowing'
/pt/: *kep*+*tauk-i-ai* 'cooking fat' ← *kep-ti* 'to roast' + *tauk-ai* 'fat' (Pl.)

Among occlusive–affricate sequences, only the cluster /tk/ is phonotactically well-formed, whereas four clusters are only morphonotactic:

/ʃt/: *didž*(*ia*)+*turt-i-s* 'wealthy man' ← *did-i-s* 'big' + *turt-a-s* 'wealth'
/dʒb/: *plač*(*ia*)+*burn-i-s* 'broad-mouth(y)' ← *plat-i* 'broad' + *burn-a* 'mouth'
/dʒd/: *treč*(*ia*)+*dal-i-s* 'one third' ← *treči-a* 'third' + *dal-i-s* 'part'
/dʒg/: *plač*/*plad*+*gal-y-s* 'oar blade' ← *plat-u-s* 'broad' + *gal-a-s* 'end'

Among fricative clusters, the two clusters /sv, šv/ are phonotactically well-formed, whereas the following five clusters are only morphonotactic:

/fs/: *gyv*+*sidabr-i-s* 'mercury' ← *gyv-a-s* 'lively' + *sidabr-a-s* 'silver'
/fʃ/: *diev*+*šauk*+*i*+*s* 'praying person' ← *diev-a-s* 'God' + *šauk-ti* 'to call'
/ʃs/: *kryž*(*ia*)+*snap-i-s* 'crossbill' ← *kryži-u-s* 'cross' + *snap-a-s* 'beacon'
/ʒv/: *maž*+*vaik-i-s* 'childish person' ← *maž-a-s* 'small' + *vaik-a-s* 'child'
/vʒ/: *purv*+*žem-ė* 'muddy earth' ← *purv-a-s* 'mud' + *žem-ė* 'earth'

[7] /pk/ has one monomorphemic exception, thus this cluster represents a strong morphonotactic default. In the following examples, in parentheses, variants are presented of the same compound with either no vowel deletion or vowel deletion with morphonological change.

Clusters which are a little less optimal medially but still observe the preference are predominantly phonotactically well-formed in the language, whereas purely morphonotactic ones are a small minority. Thus among occlusive-fricative and fricative-occlusive clusters, 16 clusters are phonotactically well-formed whereas only four are purely morphonotactic.

Among sonorant clusters, there are five phonotactically well-formed ones, namely /lm, ln, rl, rm, rn/, and six purely morphonotactic clusters /lr, ml, mn, mr, nm, nr/. Among occlusive–sonorant and sonorant–occlusive clusters (which are generally dispreferred medially), there are 17 phonotactically well-formed occlusive–sonorant clusters, but only six morphonotactic ones, while there are 18 phonotactically well-formed sonorant-occlusive clusters, but only four morphonotactic ones.

Summing up, in Lithuanian compounding, the deletion of the thematic vowel of the first component clearly does not result in the creation of new and marked clusters. Instead, among the clusters found in medial position, the "compound clusters" are the only ones that are truely phonotactically preferred. Thus, in Lithuanian compound formation, the rise of phonotactically marked consonant clusters appears to be disfavoured, and in this respect Lithuanian is similar to other languages which employ vowel interfixation for that purpose.

7. More on diachrony

Inflection creates marked phonological clusters in those conservative Indo-European languages which have inflectional infixation. For example the present stems of the Latin roots /tag/, /iug/, /rup/ show a nasal infix: *tang-o* 'I touch', *iung-o* 'I join', *rump-o* 'I break'. These clusters are phonotactically marked since they violate what seems to be a preferece for word medial clusters with a small NAD among their constituents. Nevertheless, all of them are also phonotactically well-formed in Latin, and therefore quite normal clusters. Thus the clusters created by Latin inflectional infixation are neither new nor marked. The same holds for Ancient Greek, Sanskrit and marginal Lithuanian infixation. This leads us to the conclusion that morphology tends to create marked phonological clusters only at the beginnings and the ends of words. Since inflectional affixes tend to be peripheral, i.e. more marginal than derivational affixes, it is mainly inflection which is signalled by marked consonant clusters.

In the light of this, let us look at the diachronic genesis of Lithuanian morphonotactic clusters. Word-final morphonotactic clusters created in the Lithuanian imperative are a relatively recent phenomenon, because the imperative suffix *-k* goes back to the particle *-ki*, which lost its final vowel only during the last centuries (Stang 1966: 219, Kazlauskas 1968: 382ff). Also, the Baltic third person future without a personal ending is a Baltic innovation (Stang 1966: 397ff, Kazlauskas 1968: 365ff). It has effects on morphonotactics, but is not caused by a general phonological change. Furthermore, the signalling of composition by deleting a stem vowel is also an innovation. Thus, the only Indo-European heritage is stem change plus *-s* genitive in the unproductive Lithuanian declension class *vanduo* 'water', Gen.Sg. *vanden+s* (cf. Hittite *watar*, Gen. *weten+aš*). But these genitive singular consonant clusters have become morphonotactic only because of the phonological or morphological simplification of word-final consonant clusters in other inflection patterns. This suggests the following conclusion: although Lithuanian is the most conservative living Indo-European language and has the most complex morphology and morphonology, the high degree to which it exploits morphonotactics is only slightly connected with Proto-Indo-European phonotactics and morphonotactics.

If we look at reconstructed Proto-Indo-European, we see that common reconstructions involve many morphonotactic consonant clusters that result from concatenative and non-concatenative morphological operations. We give just a few examples of the effects of vowel deletion by regular ablaut, i.e. the switch to zero grade (Schwundstufe) in inflection and derivation. We focus on the reconstructed laryngeals $*h_1$, $*h_2$, $*h_3$, which were presumably fricatives (Mayrhofer 1986: 121ff; 2004: 25ff; Schindler 1972: 2ff). In the transition to specific daughter languages these laryngeals were lost or transformed, most of the time lengthening an immediately preceding vowel, e.g. $*eh_1 > \bar{e}$, $*eh_2 > \bar{a}$, $*eh_3 > \bar{o}$. The canonic root structure of PIE verbs in the full grade is *CeC-, and less frequently *CCeC- or *CeCC-.

In the zero grade the /e/ is deleted and morphonotactic clusters may arise. For example, $*peth_2$- 'expand' (preserved in the Greek aorist *e-péta-sa*) has the zero grade in the infixed present stem $*pt\text{-}ne\text{-}h_2$- (reflected by Greek *pítnā-mi*). In the transition to the daughter languages such consonant clusters were repaired by the introduction of a weak vowel. The same happened where a laryngeal was involved in the morphonotactic consonant cluster, but where that laryngeal was colouring the weak vowel, e.g. in the past participles $*d^hh_1\text{-}tó\text{-}s$, $*sth_2\text{-}tó\text{-}s$, $*dh_3\text{-}tó\text{-}s$ (Greek $t^hetós$, *statós*,

dotós) of the full-grade forms **d^{h}eh$_1$-*, **steh$_2$-*, **deh$_3$-* 'to lay, stand, give' (Greek presents *tí-t^{h}ē-mi*, *hístā-mi*, *dí-dō-mi*) or in the third person present indicative **h$_1$s-énti* (Hitt. *ašanzi*, Mycenean *e-e-si* = /ehensi/) of the root **h$_1$es-* 'to be' (Lat. *es-se*).

Also nominal inflection and derivation is reconstructed as creating morphonotactic consonant clusters via zero grades, e.g. **h$_1$néh$_3$men-* 'name', oblique cases **h$_1$nh$_3$men-* (reflected by Doric *énuma* < **énoma*); **d^{h}eh$_1$s-* 'god' (reflected in Lat. *fēstus* 'festive') vs. **d^{h}h$_1$s-*, as in Greek *t^{h}és-p^{h}atos* 'said by God'; **d^{h}ég'^{h}ōm* 'earth' (Hitt. *tekan*), Locative **d^{h}g'^{h}ém* (cf. Greek *k^{h}am-aí*), Gen. **d^{h}g'^{h}m-és* (cf. Schindler 1977).

Thus Proto-Indo-European had presumably many more morphonotactic consonant clusters resulting from morphological vowel deletion than Polish or Russian. Both Slavic languages are much less strongly inflecting languages than reconstructed Proto-Indo-European. However, there is no direct diachronic link between Proto-Indo-European and contemporary Slavic morphonotactics, because the intermediate stage of Proto-Slavic was very different in that respect. That stage had practically no morphonotactic consonant clusters because of its open-syllable character connected with the loss of word-final and syllable-final consonants, the monophthongization of falling diphthongs, the metathesis of (*t*)*ort* groups, changes of the groups consonant + *j*, and the simplification of consonant triples.

Therefore, the similarity between PIE and Polish morphonotactics does not reflect diachronic continuity but results from typological characteristics of the inflecting-fusional type. We have seen a similar divergence between typological similarity and diachronic development in Lithuanian. But this is just the morphological perspective on morphonotactics. If we look at the phonological side, we find a typological change from highly consonantal Proto-Indo-European to much more vocalic Proto-Slavic and then back to much more consonantal Polish. This illustrates the nature of morphonotactics as an interface between morphology and phonology very well.

8. Synchronic variation

Finally, let us look briefly at the role of synchronic variation in morphonotactics. In addition to the examples of morphological variation already given, the following morphological strategies for preventing the rise of morphonotactic clusters deserve to be noticed.

Word-initially, Polish has an admittedly rather obsolete morphonological /e/-insertion in *w+spinać się* [fsp-] 'climb' (imperfective) ~ perfective 3.Sg. *wespnie się*. This prevents the word-initial cluster [fspn]. /e/-insertion occurs also in *w+ścielać* [fɕtɕ-] (impfv.) 'make the bed' ~ pfv. *wesłać*. While the second pair is generally rather obsolete, /e/-insertion in *wespnie się* is also avoided by using the periphrasis *będzie się wspinać* 'will be climbing' or by even using *wspnie się* [fspɲ-] without insertion. There is, however, also an example of /e/-insertion in current use, namely impfv. *w+spierać* [fspʲ-] 'support' ~ pfv. *wesprzeć* [vespʃ-].

Such morphological repair mechanisms may go back diachronically to earlier phonostylistic repair. Such phonostylistic repair mecahnisms are found synchronically in fast or casual speech reductions. This happens with many masculine singular preterits in *-ł* /w/ preceded by a root-final obstruent, as in *szed+ł* 'he walked', *rós+ł* 'he grew' where the word-final morphonotactic clusters /dw, sw/ quite normally lose the final preterit marker /w]. Similar reductions occur in root-final clusters in zero-genitive /plural forms, e.g. as in *przestępstw* [-mpstf → -mstf, -ms] from neuter nominative *przestępstwo* 'crime' or in *mężczyzna* 'man' ~ gen. pl. *mężczyzn* [-zn → s]. Simpler cases of either obligatory or phonostylistic repair occur also in other languages such as German (cf. Dressler & Dziubalska-Kołaczyk 2006).

9. Conclusion and outlook

In the diachronic developments discussed we have found that the main mechanisms by which morphonotactic consonant clusters were changed into phonotactic ones and vice versa have been regular sound changes. This includes prosodic changes in the typology of syllable structure, towards or away from an open syllable. A further phonological trigger has been irregular sound change within specific affixes, probably related to earlier synchronic phonostylistic variation. Whereas phonostylistic change may be the antecedent of later morphonotactic patterns, these may in turn be transformed in diachrony into isolated, fossilized lexical instances, obviously because of morphological analogic levelling. This corroborates the general view of both Neogrammarians and Natural Morphology that when phonology and morphology interact in diachrony, the former is typically active, the latter reactive.

Identifying morphonotactics as a proper subfield of morphonology opens new perspectives for diachronic studies. Our contribution has attempted to indicate some major directions for such studies. In particular, and in addition to what was observed above, the issue of phonotactic markedness of morphonotactic clusters needs further investigation. The major prediction that morphonotactic consonantal clusters will be more marked (either universally or language-specifically) than intramorphemic phonotactic ones appears to hold true for inflection and, largely, derivation, but not for compounding. Can this be derived from the different diachronic sources of patterns of inflection, derivation and compounding?

References

Ambrazas, Vytautas ed. 1996. *Dabartinės lietuvių kalbos gramatika*. Vilnius: Mokslo ir enciklopedijų leidykla.

Dressler, Wolfgang U. 1985. *Morphonology*. Ann Arbor: Karoma Press.

Dressler, Wolfgang U. 1996. A functionalist semiotic model of morphonology. In R. Singh, ed. *Trubetzkoy's Orphan*. Amsterdam: Benjamins, 67–83, 102–105.

Dressler, Wolfgang U. & Katarzyna Dziubalska-Kołaczyk. 2006. Proposing morphonotactics. *Rivista di Linguistica*, 18: 249–266.

Dressler, Wolfgang U., Willi Mayerthaler, Wolfgang U. Wurzel & Oswald Panagl. 1987. *Leitmotifs in natural morphology*. Amsterdam: Benjamins.

Dubisz, Stanisław. 2006. *Uniwersalny słownik języka polskiego PWN*. Warsaw: Wydawnictwo Naukowe PWN.

Dvonč, Ladislav, Gejza Horák, František Miko, Jozef Mistrík, Ján Oravec, Jozef Ružička & Milan Urbančok. 1966. *Morfológia slovenského jazyka*. Bratislava: Vydavatel′stvo Slovenskej akadémie vied.

Dziubalska-Kołaczyk, Katarzyna. 2002. *Beats-and-Binding phonology*. Frankfurt am Main: Peter Lang.

Dziubalska-Kołaczyk, Katarzyna. 2009. NP extension: B&B phonotactics. *PSiCL* 45: 55–71.

Dziubalska-Kołaczyk, Katarzyna & Grzegorz Krynicki. 2007. Universal phonotactic calculator: B&B phonotactics and morphonotactics at work. Paper presented at the Poznań Linguistic Meeting 2007.

Dziubalska-Kołaczyk, Katarzyna & Jarosław Weckwerth, eds. 2002. *Future challenges for natural linguistics*. Munich: Lincom.

Dziubalska-Kołaczyk, Katarzyna & Daria Zielińska. 2010. Predicting phonotactic difficulty in second language acquisition. In Andréia S. Rauber, Michael A. Watkins, Rosane Silveira & Rosana D. Koerich, eds. *The acquisition of second lan-*

guage speech: Studies in honor of Professor Barbara O. Baptista. Florianópolis: Graduate Program in English/Federal University of Santa Catarina.
Dziubalska-Kołaczyk, Katarzyna & Daria Zielińska. In press. Universal phonotactic and morphonotactic preferences in second language acquisition. To appear in *New Sounds*.
Hurch, Bernhard & Richard Rhodes, eds. 1996. *Natural Phonology: The state of the art*. Berlin: Mouton de Gruyter.
Kazlauskas, Jonas. 1968. *Lietuvių kalbos istorinė gramatika*. Vilnius: Mintis.
Kilani-Schoch, Marianne & Wolfgang U. Dressler. 2005. *Morphologie naturelle et flexion du verbe français*. Tübingen: Narr.
Mayrhofer, Manfred. 1986. Lautlehre. In M. Mayrhofer, ed. *Indogermanische Grammatik I*. Heidelberg: Winter, 73–216.
Mayrhofer, Manfred. 2004. *Die Hauptprobleme der indogermanischen Lautlehre seit Bechtel*. Vienna: Verlag der Österreichischen Akademie der Wissenschaften.
Respond, Stanisław. 2003. *Gramatyka historyczna języka polskiego*. Warsaw: Wydawnictwo Naukowe PWN.
Schindler, Jochem. 1972. *Das Wurzelnomen im Arischen und Griechischen*. Universität Würzburg dissertation.
Schindler, Jochem. 1977. A thorny problem. *Sprache* 23: 25–35.
Stang, Christian. 1966. *Vergleichende Grammatik der Baltischen Sprachen*. Oslo: Universitetsforlaget.
Townsend, Charles E. 1975. *Russian word-formation*. Cambridge, Mass.: Slavica Publishers.
Walczak, Bogdan. 1999. *Zarys dziejów języka polskiego*. Wrocław: Wydawnictwo U. Wrocławskiego.

Authors' addresses
Wolfgang U. Dressler
Department of Linguistics and Communication Research
Austrian Academy of Sciences
Kegelgasse 27
A-1030 Wien
Austria
wolfgang.dressler@univie.ac.at

Katarzyna Dziubalska-Kołaczyk
School of English
Adam Mickiewicz University
al. Niepodległości 4
61-874 Poznań
Poland
dkasia@ifa.amu.edu.pl

Lina Pestal
Department of Linguistics
University of Vienna
Sensengasse 3a
A-1090 Wien
Austria
lina.pestal@univie.ac.at