Phonotactic vs. morphonotactic obstruent clusters in Slovak and German

Wolfgang U. Dressler, Miroslava Hliničanová, Matej Ďurčo, Karlheinz Mörth, Katharina Korecky-Kröll

Slovak has more consonants than German and also more consonant clusters (plus accented syllabic sonorants). In addition, their asymmetric distribution within the word is different: Slovak has more clusters in word-initial position, German many more in word-final position. This difference is reinforced by morphology, insofar as only Slovak has monoconsonantal prefixes, and only German has monoconsonantal suffixes. The main word-internal difference is that in compounding German increases clusters, whereas Slovak decreases them. Asymmetries in terms of type frequency are radicalised in token frequency, i.e. in the profitability of clusters. In addition, clusters arise in Slovak also via vowel deletion, which is extremely rare in German (except in casual speech and in dialects).

1. Introduction

In continuation of previous theoretical and contrastive work (Dressler & Dziubalska-Kołaczyk 2006, Dressler *et al.* 2010, Korecky-Kröll *et al.* 2014) we contrast Slovak and German patterns of consonantal morphonotactics vs. phonotactics from a phonological, morphological and corpus-linguistic perspective. So far we have done only some corpus-linguistic work on German, but the ongoing PhD thesis of the second author allowed a vast corpus-linguistic study of Slovak as well (which has never been done before). This enables us to discuss it also in the framework of typological differences between a Slavic and a Germanic language.

The theoretical background is Natural Phonology and Morphology (cf. Dziubalska-Kołaczyk & Weckwerth 2002, Kilani-Schoch & Dressler 2005), and this approach does not only strive towards descriptive und explanatory adequacy but also towards guaranteeing, at least partially, the psychological reality of the linguistic constructs. This demands a psycholinguistic perspective. In contrast to Korecky-Kröll *et al.* (2014), the psycholinguistic perspective of this contribution is devoted to acquisition and not to processing. In usage-based linguistic and psycholinguistic approaches (Bybee 2001, Bauer 2001, Tomasello 2003) it is often claimed that token frequency is important only for the question of storage (which is not an issue

Italian Journal of Linguistics, 27.1 (2015), p. 45-60 *(received September 2014)*

here), whereas only type frequency and discrepancy between high type frequency and low token frequency is relevant for productivity and profitability of patterns (cf. Du & Zhang 2010, Berg 2014). We compared type and token frequency, in order to evaluate these claims with fresh data.

We investigate prototypical and non-prototypical cases of morphonotactics, i.e. the prototypical case of merely concatenative shapes of morpheme combinations, particularly when they differ from the phonotactics of lexical roots and morphemes and thus signal morpheme boundaries, as in English *seem-ed* /si:m-d/ (i.e., there is no lexical final [-md] cluster in English). All German cases that we investigate are of this type. Slovak has, in addition, also the non-prototypical case of morphological combinations resulting in vowel deletion, which is extremely rare in German, for example in *Risiko* 'risk', adj. *risk-ant* 'risky'. Slovak has many more, e.g. *pes* 'dog', pl. *ps-y,* adj. *ps-í*.

In those cases in which the pattern investigated occurs both in lexical phonotactics and morphonotactically, such as word-final [pst] in German *Papst* 'pope' vs. *du gab-st* 'you gave', we will look also at the quantitative distribution of both patterns in large electronic corpora. In order to focus on the most radical contrasts between German and Slovak, we limit our study to obstruent clusters.

Consonant clusters are studied within the framework of Dziubalska-Kołaczyk's (2002) Beats-and-Binding model (cf. Dziubalska-Kołaczyk & Zielińska 2010), which starts with syllabic nuclei as beats and consonants bound to them, but does not assume syllabic boundaries. For the sake of better understanding we will however provide information about traditional syllabification for each group of examples. Words are defined as morphological words and thus include compounds (which also have just one main accent).

The corpus-linguistic research is done with the help of the huge electronic corpora of the Slovenský národný korpus (Slovak National Corpus), which contains 829 million word tokens, the Austria Media Corpus, which contains all the content of all Austrian journals and periodicals of a period comprised between 15 and 5 years ago (7 billion word tokens), and the Austrian Academy Corpus, which focuses on written texts of all genres of the nineteenth and earlier twentieth centuries (half a billion word tokens).

Examples are written in the national orthographies. In the German consonantal system, *ch* is a voiceless palatal or velar fricative, *sch* (and word-initial *s* before a stop) a voiceless shibilant. Its Slovak counterpart is written *š*, its voiced correspondent *ž*, *č* stands for a voiceless palatal affricate, *dž* for its voiced counterpart, *z* is the

voiced counterpart of *s*; *c* (as in German) and *dz* the corresponding affricates, *ch* a voiceless prevelar fricative, *h* its voiced counterpart, ť and ň are the palatalized counterparts of *t* and *n;* an accent above a vowel grapheme means length; *r, l* between obstruents are syllabic. In Slovak there is a biunique relation between phonemes and graphemes, with the exception of word-final obstruent devoicing and forward devoicing of voiced to following voiceless obstruents (both processes also occur in German, as well as the digraph *ch*).

2. Word-initial position

Word-initially, Slovak is much richer in obstruent clusters than German, i.e. due to morphonotactic concatenation through the prefixes *z-, s-, v-,* e.g. in *s+chrad+nú-ť* 'to get older', *z-drav-i-ť* 'to greet', *v-kvap-k-a-ť* 'to drip in' and because of greater phonotactic complexity, e.g. *pstruh* 'trout', *škvár-a* 'slag', *vdov-a* 'widow'. Since Slovak is very rich in word-initial obstruent clusters (which is not the case in German), Slovak has many obstruent clusters which are either purely phonotactic or purely morphonotactic or are ambiguous, i.e. both phonotactic and morphonotactic.

Purely phonotactic obstruent clusters are formed by the obstruents *džb* in *džbán* (2130 tokens) 'jug', *pch* in *pchať* (1850) 'mend', *more c*omplex *pstr* in *pstruh* (3552) 'trout', *pšt in pštros* (920) 'ostrich'. Purely phonotactic clusters are also combinations with the affricate č: *čk* in *čkať* (209) 'gulp', *čp* in *čpavok* (661) 'ammonia' and *čv* in *čvachtať* (124) 'clap', more complex *škv* in *škvarka* (313) 'bacon greave'. Of course, one has to classify the syllabic sonorants /r, l/ as non-consonantal, as in *štrngnúť* (520) 'impinge', *žbrnda* (220) 'sword'.

Purely morphonotactic obstruent clusters are *s-č* in *s-černieť* (351) 'to become black' or *s-ť* in *s-ťahovať* (14463) 'drag along' or *v-b* in *v-biehať* (244) 'come in', *v-bodnúť* (93) 'prick', *v-biť* (48) 'strike in'.

Slovak ambiguous obstruent clusters, which are both phonotactic and morphonotactic, are restricted to combinations with word-initial obstruents *s, z, v* (only exception: *ps*, see below), e.g. *sc* in *scenár* (31189) 'scenario' vs. *s-c* in *s-cvrknúť* (1475) 'shrink', *sch* [sx] in *schizma* (494) 'schisma' vs. *s-ch* in *s-chladiť* (2229) 'chill', *sf* in *sféra* (30612) 'sphere' vs. *s-f* in *s-formovať* (3628) 'form', *sk* in *skala* (25427) 'rock' vs. *s-kombinovať* (1576) 'combine', *vč* [fč] in *včera* (201703) 'yesterday' and *včela* (8429) 'bee' vs. *v-č* in *v-čas* (14111) 'in time', *v-člen-iť* (909) 'intersperse', *vzl* in *vzlyk* (1250) 'sob' vs. *v-zl* in *vzlet* (868) 'ascent', *zb* in *zbojník* (5799) 'robber' vs. *z-b* in *z-baliť* (4266) 'pack up', *zbr* in *zbraň* (78918) 'weapon' vs. *z-br* in

z-brúsiť 'smooth', *zv* in *zviera* (63895) 'animal' vs. *z-voliť* (43706) 'choose'. The exceptional cluster *ps-* is phonotactic in *psí* 'slag', morphonotactic as result of morphology-induced vowel deletion of the root vowel of *pes* 'dog' in Pl. *ps-y* and in the derived adjective *ps-í* (7916), *ps-in-ec* (98) 'dog den', compound *ps-o-vod* (2443) 'dog guide'.

Table 1 additionally provides information about the proportion of clusters that can be both phonotactic and morphonotactic in the lexicon. The as phonotactic clusters include the "ambiguous" clusters as well; the same holds for the morphonotactic counterparts.

Table 1. Slovak word-initial obstruent clusters

CLUSTERS	TOKENS	LEMMAS
morphonotactic (M)	42 058	58
phonotactic (P)	12 073 799	1 884
ambiguous / thereof morphonotactic	2 360 612	930
ambigues / thereof phonotactic	23 177 076	2 754
Total	37 653 545	5 626

Clearly word-initial phonotactic clusters (including those occurring in inflected word-forms) largely outweigh morphonotactic types and this contrast is radicalised in tokens. We interpret this as showing that the relative normalcy (cf. Wurzel & Schentke 1989) of phonotactic sequences within a language is not only a question of type frequency (i.e. number of lemmas or inflectional word forms in which each sequence occurs), but also a question of their profitability in term of tokens. This is especially important for consonant clusters, because very frequent words are known for their tendency to reduce consonant clusters.

The analysis shows that most clusters are of the ambiguous group and that the marked group of purely morphonotactic clusters is extremely limited. In other words, morphonotactic clusters are nearly always built on the patterns of existing phonotactic clusters.

The German standard has no monoconsonantal prefixes, in contrast to Bavarian-Austrian dialects, as in *g'storben* 'died', *b'soffen* 'drunk', *z'ruck* 'back(wards)', etc., corresponding to Standard German *ge-storb-en, be-soff-en, zu(-)rück*. And the in the highest level of phonotactic complexity is reached in the sequence of two obstruents plus a sonorant, as in *Strafe* 'fine', *Sprache* 'language', *Splitter* 'splinter', *Strontium, Skrupel* 'scruple', *Sklave* 'slave'.

The other German word-initial obstruent clusters are: [*kv*] as in

Quelle 'source', [ks] as in *Xenophobie,* [ps] as in *Psychologie* 'psychology', *zw* [tsv] as in *Zwang* 'coercion', *sc*, *sk* [sk] in loanwords such as *scannen, Skalp*, *st* in loanwords such as *Stil* 'style' (compare *st* [ʃt] in *Stock* 'stick'), *sp* in loanwords such as *Spatium* 'space' (vs. *sp* [ʃp] in *Spaß*), [sf] in loanwords such as *Sphäre* 'sphere', [sv] in loanwords such as *Sweater*, [sç] in loan words such as *Schizophrenie* 'schizophrenia', *schw* [ʃv] as in *Schwanz* 'tail'; limited to learned loanwords, one should also mention *pt, ft*, such as *Pteridin, Phthisis, cht* [xt], such as *chthonisch, sz* [sts] as in *Szene* 'scene'.

3. Word-internal position

Word-medially both languages have comparable phonotactic consonant clusters. In German, the syllable boundary is in compounds immediately before the second compound member and in prefixations always after the prefix, otherwise word-internal syllable onsets always follow the pattern of word-initial onsets. In Slovak, the syllable boundary is always after the first consonant of clusters (Dvonč *et al.* 2013).

Morphonotactic obstruent clusters arise in word formation in different distributions according to the typological difference between Germanic compounding languages and the rather derivational Slavic languages. Thus German compounds such as *Dienst+grad* 'military rank' have no correspondence in Slovak, not only because common Slovak is much poorer in compounds than German, but also because for Slovak the compounding interfix *-o-* is typical, as in *štrk-o-piesky* 'gravel beach' (from *štrk* 'gravel', *piesok* 'sand', Pl. *piesk-y*), *mozg-ov-o-ciev-n-y* 'cerebrovascular' (*mozog* 'brain', Adj. *mozg-ov-ý, ciev-a* 'vessel', Adj. *ciev-n-y*). In contrast, German has the obstruent interfix *-s-*, which increases consonantal complexity, as in *König-s-krone* 'royal crown', *Fried-hof-s-mauer* 'cemetery wall'.

Slovak word formation creates, via many consonant-final prefixes and consonant-initial suffixes, many complex morphonotactic obstruent clusters, such as *roz-drob-en-ý* 'broken up', *gróf-stvo* 'county'. Furthermore, in contrast to German, Slovak declension creates via vowel deletion new morphonotactic obstruent clusters, as in *mozog* 'brain', Gen.Sg. *mozg-u* vs. *vosk* 'wax', *laket'* 'ellbow', Gen.Sg. *lakt'a* vs. *akt* 'act', *otec* 'father', Gen.Sg. *otc-a, list-ok* 'leaflet', Gen.Sg. *list-k-a / u, o-pas-ok* 'belt', Gen.Sg. *o-pask-a* vs. *po-tok* 'creek', Gen.Sg. *po-tok-a / u,* prefixed from *tok* 'river, stream'.

Since Slovak does not allow word-final obstruent groups *-zg, -tk, -stk*, one could be induced to assume that we do not deal here with

vowel deletion, but with automatic phonological vowel insertion in the Nom.Sg. /mozg, list-k/, etc. As a consequence, these obstruent clusters would not be morphonotactic, but phonotactic. However with some obstruent clusters, these occur also word-finally, i.e. one would have to assume a non-automatic, lexically restricted morphonological vowel insertion, for example in Nom.Sg. *o-pas-ok* 'belt', Gen. *o-pas-k-u / a* (a morphosemantically opaque diminutive derivation from *pás* 'strap, belt') vs. *vosk* 'wax', *stisk* 'pressure', *vý-prask* 'clout', *kiosk* 'kiosk'. Note that the assumed underlying /mozg/ would be subject to obligatory word-final obstruent devoicing to [mosk]: this word-final obstruent cluster would be perfectly pronounceable.

Of the other obstruent clusters we would like to cite the special case of disyllabic *chrbát* 'back', Gen. *chrbt-a / u* vs. the Latinate loanwords *recept* 'receit'*, adept, koncept, Egypt.* The insertion of a long [a:] in /xrbt/ → *chrbát* would not only be isolated, but also phonologically totally implausible. The assumed underlying Nom.Sg. /xrbt/ would be subject to phonological voicing assimilation /xrpt/. The isolatedness of the deletion of long *-á-* in all other inflection forms of Masc. *chrbát,* plus in the derivations Fem. *chrbt-ic-a* ‚back bone', Adj. *chrbt-ov-ý* has stimulated the formation of rare alternative forms without vowel deletion. The token frequencies in the Slovak National Corpus are, in the case of Gen.Sg., *chrbt-a* 6368, *chrbt-u* 172 vs. *chrbát-a* 68, *chrbát-u* 19. This token frequency distribution is further evidence for the basic status of /a:/ in *chrbát.*

Furthermore, vowel deletion (which is also assumed in Růžička *et al.* 1966, Sokolová *et al.* 1999) instead of vowel insertion is also supported by the two following arguments. First, the alternating vowel is restricted to the unmarked base form of the Nom.Sg., whereas vowel deletion signals in an anticipatory manner (cf. Dressler 1985) marked inflection, derivation or compounding via the new morphonotactic consonant cluster; see e.g. the above-mentioned example *mozg-ov-o-ciev-n-y* 'cerebrovascular' (from *mozog* 'brain'). Second, the alternating vowels originated by sound change law from the Protoslavic ultra-short vowels jer and jor,[1] which later were automatically deleted in unstressed positions (cf. Krajčovič 1971, Sabol 1989, e.g. in Old Church Slavonic *xrъbъtъ* > Slovak, Serbo-Croatian *chrbát*, Bulgarian *chrbét*, Slovene *hrbét*). In current theories of language change there is no plausible explanation, why in these alternations vowel deletion should have been reanalysed as vowel insertion.

German correspondences with obstruent-initial suffixes are morphonotactic *Kind-chen* 'child-DIM'*, Gast-lich-keit* 'hospitality' (/ç/ is the only obstruent followed by the allomorph *-keit* of *-heit* in German, thus obstruent clusters and especially plosive clusters are avoided),

lenk-bar 'guidable', *Schick-sal* 'fate', *furcht-sam* 'timid', *ab-wärts* 'downwards', unproductive *Jag-d* 'hunting', *Schlach-t* 'battle', *Schrif-t* 'writing', *Ein-kunf-t* 'income', *Ge-lüb-de* 'vow', *Klap-s* 'smack', *Knack-s* 'crack', *äch-zen* 'to groan', the isolated suffixoid in *Witz-bold* 'witty fellow'.

Verb prefixation too creates new word-internal obstruent clusters. For example, the separable prefix/particle *ab-* motivates the exclusively morphonotactic clusters /p-d, p-t, p-g, p-k, p-ʃ, p-ts, p-v/, as in *ab-drehen, ab-treten, ab-geben, ab-kommen*, *ab-schaffen*, *ab-ziehen, ab-wickeln* (with the addition of longer clusters, as in *ab-streiten*). Also some of the few non-separable verbal prefixes create new clusters, as with *ent-*, and the earlier but now only vestigial suffix *ant-*. All nominal and adjectival prefixes are non-separable. In addition, prefixes (and compounding) create geminate consonants which are disallowed morpheme-internally, and, phonotactically even worse, pseudogeminates are created by syllable- and morpheme-final obstruent devoicing, as in *ab-bauen* with /p-$b/.

4. Word-final position

In word-final position, German has morphonotactic *-Cst* clusters in 2nd singular verb forms, superlatives and unproductive nominalisations with the homophonous *-st* suffix.

As Dressler & Dziubalska-Kołaczyk (2006) have shown, exclusive morphological motivation exists for the clusters /-xst, -fst/, as in *lach-st* 'you laugh', *schläf-st* 'you sleep' as well as for the clusters /-pfst, -rkst -lkst, -nkst, -lpst, -mpst/.

A strong default is represented by the regional variant /-nkst/ (as in *häng-st* 'you hang' vs. monomorphemic *Angst* 'anxiousness' as well as by /-rpst. I.e. there are extremely few phonotactic cluster occurrences vs. an extensive, potentially unlimited number of morphonotactic cluster occurrences. The default is slightly weaker (i.e. there are significantly more phonotactic exceptions) in postvocalic /-pst/, as in *lieb-st* 'you love' as well as in postvocalic /-kst/ and in the affricate /ts/ followed by /-(s)t/, as in *reiz-(s)t* 'you/(s)he irritate(-s)'. And the morphonotactic default is even weaker with base-final sonorant as in /-rst, -nst/ clusters.

A further source of word-final morphonotactic obstruent groups is the nominal *-s* Gen.Sg., as in *des Lob-s* 'of the praise', *Prinzip-s* 'principle', *Kalb-s* 'calf', *Skalp-s* 'scalp', *Korb-s* 'basket', *Ge-zirp-s* 'chirping', *Schiff-s* 'ship', *Archiv-s* 'archive', *Schilf-s* 'reed', *Dorf-s*

'village', *Nerv-s* 'nerve', *Krieg-s* 'war', *Lack-s* 'varnish', *Werg-s* 'tow', *Quark-s* 'curd', *Talg-s* 'tallow', *Ulk-s* 'trick', *Dach-s* 'roof'.

Parallel phonotactic clusters occur in *Schnaps* 'brandy', *Gips* 'gypsum', *Raps* 'colza', *Kollaps* 'collapse', *Krebs* 'cancer', *Rülps* 'belch', *Knirps* 'dwarf', *Keks* 'cookie', *Koks* 'coke', *Klecks* 'blot', *Fuchs* 'fox'. Morphonotactic word-final *(r/l)fs, xs, lks, rks* have no phonotactic parallels.

A problem is represented by imperatives of the type *knicks!*, *schubs!* 'curtsy!', 'push!'. First, it is unclear whether the word-final -s is synchronically still a derivational suffix. Second, even if not, it is unclear whether such imperatives are to be classified as base forms (if yes, then phonotactic) or as morphologically derived from the infinitive as lexical entry.

In contrast to German, Slovak word-finally has no consonantal inflection suffixes, but creates morphonotactic obstruent clusters in suffixless genitive plurals via deletion of the stem vowel, as in *rod-i-sk-o* 'birth place', Gen.Pl. *rodísk*, *mzd-a* 'wage', Gen.Pl. *miezd*.

Examples of Slovak word-final obstruent clusters, which are both morphonotactic and phonotactic, (with their token frequency values within brackets) are e.g. Gen.Pl. (all neuters) *prac-ov-í-sk* (9573) 'of the working places', *voj-sk* 'troups' (8015) vs. Nom.Sg. *zisk* 'yield' (20788) or *blesk* 'lightning' (4852); Gen.Pl. *účilíšť* 'educational establishments' (935) vs. *plášť* (2120) 'coat'; *miest* (9982721) 'towns' vs. *rast* 'growth' 37958. Among the female Gen.Pl., one can mention [st] in *hviezd* (153253) 'of stars' vs. *zjazd* (3239) 'assembly'; from *naft-a* 'petroleum', the Gen.Pl. *náft* (6) vs. *kšeft* 'shop' (from Austrian German *G(e)schäft*') (551); from *pošt+a* 'post', Gen.Pl. *pôšt* (4813) or from *vražd-a* 'murder', Gen.Pl. *vrážd* [vra:št] (4723) vs. *mušt* (214) 'must'; from *lopt-a* 'ball', Gen. Pl. *lôpt* (2819) vs. *koncept* 'concept'(3025); Gen. Pl. *siekt* (905) 'seckts' *vs. projekt* 'project' (51215); Gen.Pl. *jácht* (336) 'yachts' or *šácht* (317) 'pits' vs. Nom.Sg. *necht* (206) 'nail'.

There are also purely morphonotactic word-final obstruent clusters, which originate only in Gen.Pl.: *ct* in *pôct* (325) 'of honours', *-pch* in *zápch* (90) 'obstructings' and *-rst* in *vierst* (211) 'wersts'.

In German, there are many word-final morphonotactic obstruent clusters due to verb inflection, especially 2nd and 3rd Sg. or 2nd. Pl., in combination of *ch* /x/ with *st* or *t*: e.g. only morphonotactic *ch-st* from *machen* 'make' in 2nd Sg. *mach-st*, whereas *ch-t* in 3rd Sg and 2nd Pl. *mach-t* has a phonotactic ccorrespondence in *Macht* 'power'. In contrast, Slovak has such morphonotactic word-final obstruent clusters only in Gen.Pl., as for example in Gen.Pl. *jácht* 'of yachts'. Thus, word-finally, German is richer in obstruent clusters than Slovak. Such obstruent clusters occur morphonotactically in the German verbs: e.g.

flitz-t 'flits', *herz-t* 'caresses', *lach-t* 'laughs', *schaff-t* 'creates', *park-t* 'parks', *klapp-t* 'flaps', *lieb-t* 'loves', *verschlamp-t* 'loses by neglect', *lass-t* 'lets', *schirr-st* 'harness' (2nd Sg.), *schiel-st* 'squint', *schien-st* 'put in splints', *salb-st* 'anoint', *kerb-st* 'notch', *hack-t* 'hacks', 2nd Sg. *hack-st, krank-t* 'is sick'.

German phonotactic correspondences of these word-final obstruent clusters occur in *Schaft* 'shaft', *Adept* 'adept', *First* 'ridge', *acht* '8', *Arzt* 'physician', *prompt* 'prompt', *Wulst* 'bulge', *Wanst* 'belly', *Markt* 'market', *Papst* 'pope', *Herbst* 'autumn', *Takt* 'tact', *Text* 'text', *Punkt* 'point'. Thus phonotactic word-final obstruent clusters occur in both languages, but they have a much higher type and token frequency in German than in Slovak, the opposite of what holds for word-initial obstruent clusters.

Only morphonotactic are the German word-final obstruent clusters in *salb-t* 'anoints', 2nd Sg. *salb-st, rupf-t* 'plucks', 2nd Sg. *rupf-st, welk-t* 'withers', 2nd Sg. *welk-st,* only 2nd Sg. *lach-st* 'laugh', *krank-st* '(you) are sick'.

Probably, segmentally identical phonotactic and morphonotactic clusters have different vowel durations (cf. Plag 2013), but it is yet unclear whether these differences lie above the threshold of perceptibility.[1] In any event, Plag is right in objecting to linguistic models which crucially contain a flow-chart from one submodule to another one in a way which presupposes bracket erasure (criticised also in Brown & Hippisley 2012: 273). Our model of morphonotactics (Dressler & Dziubalska-Kołaczyk 2006, Dressler *et al.* 2010, Korecky-Kröll *et al.* 2014) does not presuppose such bracket erasure. This also fits the above mentioned Slovak word-medial patterns: assuming that in a flow-chart inflectional morphology follows derivational morphology, the derivational boundary in *po-tok* must not be erased in order to prevent vowel deletion in Gen.Sg. *po-tok-a / u,* in contrast to the oblique cases of *líst-ok* and *otec*.

5. Phonotactic markedness

For German and other languages, there exist phonotactic analyses which assume any third consonant in a tautosyllabic consonant cluster as extrasyllabic or extrametrical (see Wiese 1988, 2000), which makes such a cluster a marked one. Additional consonantal morphemes increase markedness.

For our purpose, one of the most extensive and sophisticated markedness analyses can be found in Dziubalska-Kołaczyk's (2002,

2009, Dziubalska-Kołaczyk & Zielinska 2010) Beats-and-Binding model, which goes beyond numerical complexity and purely sonority based models. This phonotactic theory measures the degree of binding of the consonants to each other and to beats (vowels and syllabic sonorants) by means of Net Auditory Distance (NAD). Being calculated as the distance which occurs between two segments in terms of manner and place of articulation, NAD evaluates word-initial and word-final consonant clusters as the more preferred the greater the relative NAD is between the beat and the adjacent more peripheral consonant in comparison to the NAD obtained between the peripheral consonants themselves.

A basic hypothesis in this model assumes that morphology renders phonotactics more marked. Thus morphonotactic obstruent clusters should be more marked than comparable phonotactic clusters (in the narrow sense). And indeed, if we compare the preferred cluster-rich positions of Slovak and German, then in Slovak word-initial positions the five most frequent morphonotactic biconsonantic obstruent clusters *s-č-, z-v-, v-ž-, s-ch-, z-ž-* are more marked than the five most frequent purely phonotactic obstruent clusters *sv-, št-, dv-, tv-, šk-*.

As to German word-final consonant clusters, the purely morphonotactic triconsonantal clusters *-f-st, -pf-st, -x-st* are among the most marked word-final clusters (cf. Korecky-Kröll *et al.* 2014). However more analyses of other clusters are needed.

6. On the acquisition of German (mor)phonotactic obstruent clusters

This psycholinguistic section is devoted to the relationship between phonotactic and morphonotactic consonant clusters in first language acquisition, because phonotactic markedness has to do with how patterns are acquired by young children. Freiberger (2007) investigated morphonotactic and phonotactic clusters in the longitudinal spontaneous speech data of a Viennese boy (aged 1;6 to 3;0) and, in contrast to Zydorowicz (2009) on Polish and English and to Kamandulyte (2006) on Lithuanian, did not find a clear advantage for the acquisition of morphonotactic as opposed to phonotactic clusters, but also no disadvantage. This can be interpreted as showing that in morphology-rich languages, such as the Slavic languages, the presence of a grammatical morpheme within a word-final, word-medial or word-initial consonant cluster helps phonological acquisition (see the notion of bootstrapping, cf. Weissenborn & Höhle 2001), because children are more stimulated to focus on the acquisition of a richer than of a poorer morphology (Xanthos *et al.* 2011).

Unfortunately no acquisition data are so far available for Slovak.

Here, we briefly present the results of an analysis of spontaneous speech data from 16 typically developing Viennese children (aged 2;11 to 3;6) and their main caretakers. There were two 30-minute-recordings per child, thus 16 hours of speech in total. Our analysis is limited to *-Cst* clusters, which are the most frequent triconsonantal clusters in German.

In contrast to the Austrian German electronic corpora and to the Slovak National corpus of written adult speech, we find a much higher percentage of morphonotactic than of phonotactic consonant clusters in the children and caretakers speech: 85 % of the children's and 90 % of the parents' word tokens with word-final *-Cst* clusters contain morphotactic boundaries. A frequency ranking of word tokens containing *-Cst* clusters shows that the top three clusters in child speech (CS) and in child-directed speech (CDS) are all morphonotactic and the same in children and parents, namely final morphonotactic *-n-st, -l-st* and *-k-st*. These are only 2nd sg. verbal forms, such as *du kann-st* 'you can', *will-st* 'want' or *mag-st* 'like' which are extremely important and frequent words in parent-child interaction. Apart from the top three clusters, CDS and CS are slightly different. In CDS the final morphonotactic cluster *-x-st*, as in *du brauch-st* 'you need', is more frequent, whereas *-f-st*, as in *du darf-st* 'you may' is more frequent in CS. In the fifth place we find word-final morphonotactic *-m-st*, as in *du komm-st* 'you come'. In CDS and CS the top phonotactic cluster is *-nst*, as in *sonst* 'otherwise'. Phonotactic *nst* is also the most frequent medial cluster of our consonant clusters, but in comparison to the most frequent morphonotactic clusters it still is very rare.

The interaction between parents and their young children represents a very specific genre of oral speech, which abounds in second singular verb forms. This shows that the distribution of morphonotactic vs. phonotactic consonant clusters may be very genre-specific. But clearly for the acquisition and establishment of markedness relations oral CDS is the most important speech genre.

7. Conclusion and outlook

Phonotactic asymmetries between word-initial, word-final and word-medial positions are well-known. This starts with how the universal preference for CV structures (Dziubalska-Kołaczyk 2002, 2009) is realized in the three positions and depending on whether a word is monosyllabic, disyllabic or polysyllabic. What is interesting for the

contrastive analysis of Slovak and German obstruent clusters is their polar difference in the asymmetry between word-initial and word-final position: whereas Slovak is much richer in type and token frequencies of word-initial obstruent clusters, German is much richer in type and token frequencies of word-final obstruent clusters. Type frequency asymmetries proved to be radicalised in token frequency differences, which means that the dominant patterns are more profitable. These asymmetries hold both for phonotactic clusters in the narrow sense and for the morphonotactic ones, even for purely morphonotactic obstruent clusters. This basic difference between the two languages appear to go back to prehistoric or early historic major vowel deletions in German word-final positions as opposed to late Protoslavic deletions of ultrashort jer and jor in preceding word positions.

In word-medial positions both languages have a rich array of phonotactic obstruent clusters (not analysed in this contribution) and of morphonotactic clusters originating in derivational morphology. There is however a strong contrast in compounding which creates morphonotactic clusters in German, but avoids them in Slovak.

Slovak is a more consonantal language than German (27 vs. 21), and allows the phonemes /r, l/ as stressed syllabic nuclei. This general profile is presumably also responsible for the greater amount of consonant clusters, which include also more marked clusters than German. Morphology favours this consonantal abundance in Slovak due to frequent vowel deletion, but disfavours it word-internally in compounding via vocalic interfixes (in contrast to German consonantal interfixes). But this is less relevant, because German has much more productivity in compounding than Slovak.

What makes a difference between phonotactic and morphonotactic obstruent clusters is the greater markedness of morphonotactic than of phonotactic clusters. Nevertheless the interaction of morphology and phonology has been found to favour morphonotactic clusters in processing and in the acquisition by typically developing children (as opposed to children with Specific Language Impairment), at least in the morphology-rich languages Polish and Lithuanian. Evidence for German acquisition is still unclear, and there is not yet any study for the acquisition of Slovak.

Nevertheless also the evidence from our huge electronic corpora of adult speech must be considered with caution. It is not entirely clear whether these data are representative for the languages analysed. All three electronic corpora are limited to written data produced by adults for adult readers (adult-directed speech). Moreover, the Austrian Media Corpus is restricted to journalistic prose, whereas the much smaller

Austrian Academy Corpus includes also literary, political, and scientific texts (including scientific popularisation). In contrast, the Austrian acquisition data at our disposal are oral data, and for adult productions limited to child-directed speech. For all these reasons, we plan to extend our analyses to adult-directed adult oral speech.

Addresses of the Authors

Wolfgang U. Dressler, Universität Wien, Institut für Sprachwissenschaft, Sensengasse 3a. Stiege, 3. Stock, A-1090 Wien & Institut für Corpuslinguistik und Texttechnologie der Österreichischen Akademie der Wissenschaften, Sonnenfelsgasse 19, 1010 Wien, Austria <wolfgang.dressler@univie.ac.at>

Miroslava Hliničanová, Katedra germanistiky, Filozofická Fakulta UMB, Tajovského 40, 974 01 Banská Bystrica, Slovakia <miroslava.hlinicanova@umb.sk>

Matej Ďurčo, Austrian Centre for Digital Humanities, Österreichische Akademie der Wissenschaften, Sonnenfelsgasse 19, A-1010 Wien, Austria <matej.durco@assoc.oeaw.ac.at>

Karlheinz Mörth, Austrian Centre for Digital Humanities, Österreichische Akademie der Wissenschaften, Sonnenfelsgasse 19, A-1010 Wien, Austria <Karlheinz.Moerth@oeaw.ac.at>

Katharina Korecky-Kröll, Institut für Sprachwissenschaft Universität Wien, Sensengasse 3a, A-1090 Wien, Austria <katharina.korecky-kroell@univie.ac.at>

Notes

[1] In the tradition of generative Lexical Phonology, Rubach (1993) assumes also for the synchronic analysis of today's Slovak the underlying phonemes jer and jor, which never occur on the surface, an abstract analysis which we reject (as already implicit from Postal's 1968 naturalness principle).

[2] In an ongoing project, we are currently testing the differences of vowel duration before phonotactic vs. morphonotactic consonant clusters, trying to fill in this research lacuna.

Bibliographical References

Bauer Laurie 2001. *Morphological Productivity*. Cambridge: Cambridge University Press.

Berg Thomas 2014. On the Relationship between Type and Token Frequency. *Journal of Quantitative Linguistics* 21(3). 199-222.

Brown Dunstan & Andrew Hippisley 2012. *Network Morphology*. Cambridge: Cambridge University Press.

Bybee Joan 2001. *Phonology and Language Use*. Cambridge: Cambridge University Press.
Dressler Wolfgang U. 1985. *Morphonology*. Ann Arbor: Karoma Publishers.
Dressler Wolfgang U. & Katarzyna Dziubalska-Kołaczyk 2006. Proposing morphonotactics. *Italian Journal of Linguistics* 18. 249-266.
Dressler Wolfgang U., Dziubalska-Kołaczyk Katarzyna & Lina Pestal 2010. Change and Variation in Morphonotactics. *Folia Linguistica Historica* 31. 51-67.
Du Lifang & Xuezhong Zhang 2010. A Survey of the Measurements of Morphological Productivity. *English Language Teaching* 3. 60-63.
Dvonč Ladislav and collective 2013. *Pravidlá slovenského pravopisu*. Bratislava: Veda.
Dziubalska-Kołaczyk Katarzyna 2002. *Beats-and-Binding Phonology*. Frankfurt am Main: Lang.
Dziubalska-Kołaczyk Katarzyna 2009. NP extension: B&B phonotactics. *PSiCL* 45. 55-71.
Dziubalska-Kołaczyk Katarzyna & Jarosław Weckwerth 2002. *Future Challenges for Natural Linguistics*. Munich: Lincom.
Dziubalska-Kołaczyk Katarzyna & Daria Zielińska 2010. Predicting phonotactic difficulty in second language acquisition. In Rauber Andreia S., Watkins Michael A., Silveira Rosane & Rosana D. Koerich (eds.), *The acquisition of second language speech: Studies in honor of Professor Barbara O. Baptista*. Florianópolis: Editora Insular. 281-304.
Freiberger Eva Maria. 2007. Morphonotaktik im Erstspracherwerb des Deutschen. *Wiener Linguistische Gazette* 74. 1-23.
Kilani-Schoch Marianne & Wolfgang U. Dressler 2005. *Morphologie naturelle et flexion du verbe francais*. Tübingen: Narr.
Korecky-Kröll Katharina, Dressler Wolfgang U., Freiberger Eva Maria, Reinisch Eva, Mörth Karlheinz & Gary Libben 2014. Morphonotactic and phonotactic processing in German-speaking adults. *Language Sciences* 46. 48-58.
Krajčovič Rudolf 1971. *Náčrt dejín slovenského jazyka*. Bratislava: Slovenské pedagogické nakladateľstvo.
Plag Ingo 2013. Against homophony: the acoustics of English {s} morphemes. Paper presented at the 3rd NetWordS workshop, Dubrovnik, September 2013.
Postal Paul 1968. *Aspects of Phonological Theory*. New York: Harper and Row.
Rubach Jerzy 1993. *The Lexical Phonology of Slovak*. Oxford: Clarendon Press.
Růžička Jiří and collective 1966. *Morfológia slovenského jazyka*. Bratislava: Vydavateľstvo Slovenskej akadémie vied.
Sabol Ján 1989. *Syntetická fonologická teória*. Bratislava: Veda.
Sokolová Miloslava, Moško Gustáv, Šimon František & Vladimír Benko 1999. *Morfematický slovník slovenčiny*. Prešov: Náuka.
Tomasello Michael 2003. *Constructing a Language: A Usage-based Theory of Language Acquisition*. Cambridge: Harvard University Press.
Weissenborn Jürgen & Barbara Höhle 2011. *Approaches to Bootstrapping:*

Phonological, lexical, syntactic and neurophysiological aspects of early language acquisition. Amsterdam: Benjamins.
Wiese Richard 1988. *Silbische und Lexikalische Phonologie*. Tübingen: Niemeyer.
Wiese Richard 2000. *The Phonology of German*. Oxford: Oxford University Press.
Wurzel Wolfgang U. & Manfred Schentke 1989. *Inflectional Morphology and Naturalness*. Dordrecht: Kluwer.
Xanthos Aris, Laaha Sabine, Gillis Steven, Stephany Ursula, Aksu-Koç Ayhan, Christofidou Anastasia, Gagarina Natalia, Hrzica Gordana, Ketrez F. Nihan, Kilani-Schoch Marianne, Korecky-Kröll Katharina, Kovačević Melita, Laalo Klaus, Palmović Marijan, Pfeiler Barbara, Voeikova Maria D. & Wolfgang U. Dressler 2011. On the role of morphological richness in the early development of noun and verb inflection. *First Language* 31. 461-479.