

RECONSTRUCTING THE SONORITY HIERARCHY

Paula Orzechowska^{a,b} & Richard Wiese^b

^aAdam Mickiewicz University Poznań, Poland; ^bPhilipps-Universität Marburg, Germany
paulao@wp.pl; wiese@uni-marburg.de

ABSTRACT

Combinations of segments in a language are subject to co-occurrence restrictions. This paper focuses on phonotactic constraints that govern the formation of word-initial consonant clusters in German. A description of the inventory of clusters results in scales of phonotactic preferability and a novel approach to a ranking of onset clusters. A set of structural preferences for clusters can be established empirically and expressed by individual scales, which can be seen as constituent parts of sonority. Since scales are generalizations for categories of size, place and manner of articulation as well as voicing, we argue that they provide a more insightful description of clusters than the frequently used measure of sonority.

Keywords: German, phonotactics, onset clusters, preferability scales, sonority

1. APPROACHES TO SONORITY

The work by Sievers [7] stimulated further proposals for a sonority hierarchy, which all use the “inherent loudness” of segments as an explanatory principle for the sequencing of segments, as in [1] and [4]. Proposals for sonority do not agree on the details of the hierarchy, have an unclear empirical basis, and can only give a sonority value to segments and to distances between adjacent segments. Therefore, the concept of sonority can be questioned, [6]. However, there still seems to be some explanatory value in the sonority hierarchy.

The present study starts with the hypothesis that articulatory categories discussed below are all relevant for an adequate account of clustering. Much as we regard sonority to be a useful tool in the description of clusters' preferability, we posit an alternative approach. It not only consists in a more detailed account of phonotactic preferences, which sonority merges into a single criterion, but also allows for establishing a ranking of clusters.

In section 2, we present a listing of all existing word-initial clusters, and their phonetic description in terms of the aforementioned categories. We advocate preferability scales for every dimension derived from the observed structural generaliz-

ations in section 3, and arrive at a ranking of all clusters in terms of their adherence to empirical preferences in section 4.

2. CLUSTER INVENTORY

2.1. Previous research on phonotactics

German is phonotactically complex and rich in consonant clusters. Studies on these clusters [2, 3, 8] have led to a number of insightful observations, but have been able neither to provide an exhaustive list of cluster inventory (accounting for rare and borrowed clusters) nor to accommodate the most significant generalizations on phonotactics as such. This paper contributes to this discussion by providing some insights into German word-onset clusters, the selection of which is motivated by greater salience of initial, as opposed to final, position.

2.2. Dictionary-based list of clusters

Onset clusters were extracted from the extant literature and then checked against corpus data.¹ The corpus-based survey of clusters guaranteed an exhaustive list of CC and CCC. The complete word list was extracted from the corpus [9], which is based on newspaper texts of present-day German. The data contains lemmata, inflected word forms, proper names, borrowings and abbreviations. Words were transcribed by the *Festival* software. Word types with low frequencies usually consisted of spelling variants, errors and misparsings, either in the corpus or as made by the transcriber. Thus, when checking the clusters against the corpus data, only entries with a frequency ≥ 100 were used.

Table 1: Onset clusters of German.

CC	CCC
bl, bʁ, dʁ, fl, fʁ, gl, gm, gn, gʁ, kl, km, kn, kʁ, ks, kv, pfl, pflʁ, pl, pn, pʁ, ps, sf, sk, sl, sm, sn, sp, sʁ, st, sts, sv, ʃk, ʃl, ʃm, ʃn, ʃp, ʃr, ʃt, ʃv, tm, tʁ, tsv, tv, vl, vʁ	skl, kʁ, skv, spl, stʁ, ʃpl, ʃpʁ, ʃtʁ

An attempt was made to include all, even rare (e.g. /ʃk/: *Schkopau*), clusters in the data with the exception of unassimilated borrowings. Sequences of a consonant followed by /j/ (e.g. /bj/: *Björn*) were not considered due to their recognizably

foreign status. /pf, ts, tʃ, dʒ/ are commonly [5, 8] treated as affricates, and, therefore, were not classified as CC. All the remaining consonantal sequences were included and are given in Table 1.

The 53 clusters given were assigned the articulatory features of POA, MOA and voice according to the IPA specification as presented in Table 2. The description is not intended as a specific commitment to the correctness of this approach, but serves as a theory-neutral classification.

Table 2: Features used for the description of clusters.

	plos	affr	fric	nas	liq
bilab	p b			m	
lab-dent		pf	f v		
alv	t d	ts	s z	n	l
post-alv			ʃ		
vel	k g				
uvul					ʁ

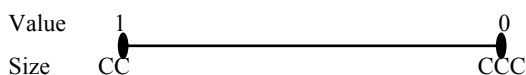
3. DERIVING SCALES

On the basis of all the CC and CCC collected in Table 1, it is argued that the distribution of clusters found in German is not arbitrary, but follows a set of identifiable preferences. Besides a size preference, the three broad dimensions of POA, MOA and voice are regarded as contributing to the patterning of onset clusters. For the 4 dimensions, we identify a range of parameters each of which is presented on a scale with values ranging from 1 to 0. The extremes are defined in such a way that preferred cluster types receive a numerical value higher than dispreferred ones. The analysis of preferred and dispreferred clusters led to the positing of 9 scales, whose description is pursued next.

3.1. Size

Inspection of the 53 clusters in Table 1 showed that 45 of them contain two consonants and only 8 contain three consonants. This fact provides evidence in favour of shorter over longer clusters. Therefore, scale expresses this preference by assigning the value of 1 to CC clusters, and the value of 0 to CCC clusters.

(1) Scale Cluster Size



3.2. Place of articulation

The POA dimension is characterised by the six distinctive values given in Table 2. Their patterning in clusters demonstrates that only a few of the

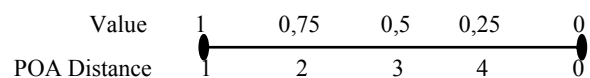
possible combinations of POA exist in onsets. Out of 36 logical combinations for CC, only 18 are attested. The proportion is even smaller for CCC: out of 216 potential combinations, only 8 are found. Inspection of the POA patterns found in all the CC and CCC led to positing three defining parameters of POA.

POA distances: Articulatory distance between adjacent consonants can be shown to be one of the defining parameters of the POA dimension. The distance of 1 is given to features in adjacent rows in Table 2. The bilabial and labiodental features are merged into a single category on grounds of complementary distribution, and thus distance 1 is given to alveolar segments combining with both labial ones. POA combinations such as alv+alv (e.g. /st/), post-alv+bilab (e.g. /ʃm/), lab-dent (e.g. /kv/) and bilab+uvul (e.g. /bʁ/) is assigned distances of 0, 2, 3, and 4, respectively. For CCCs, the mean value is computed for C1-C2 and C2-C3, as in /spl/=(1+1):2=1 and /ʃtr/=(1+3):2=2. In cases of rational numbers, the value obtained is rounded down to the closest whole number, as in /skv/=(2+3):2=2.5=2 and /ʃpl/=(2+1):2=1.5=1.

We observe a strong tendency for smaller POA distances to be preferred over larger ones, with the highest number of clusters of distance 1. In general, as the distance increases above 1, the number of clusters with a given pattern decreases. The cluster inventory contains 24 types with distance 1, 12 with 2, 7 with 3 and 5 with 4. 0 distance is represented by only 4 clusters.

Of all 6 possible identical POA CC sequences, only 1 is found: alv+alv. Of all possible identical POA CCC sequences (6), none is found. This result suggests that much as smaller POA distances are favoured, clusters with larger distance are strongly avoided. Thus, in scale, the value of 0 is assigned to clusters with zero distance. In contrast, clusters with the distance 1 score 1 point.

(2) Scale POA Distance

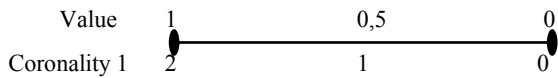


Presence of a coronal: As can be seen in Table 1, alveolar and post-alveolar consonants, subsumed under the class of coronals, occur once in 31 clusters, and twice in 12 clusters. In other words, coronal segments emerge 55 times, whereas labials and dorsals 29 and 30 times. Furthermore, out of 53 clusters, only 10 CCs involve no (post-)alveolar articulation. Out of these 10 clusters, 7 contain the uvular /ʁ/. The tendency for coronals to emerge in clusters is strengthened with the increase in the

number of segments in a string. All CCC clusters contain at least one (post-)alveolar consonant.

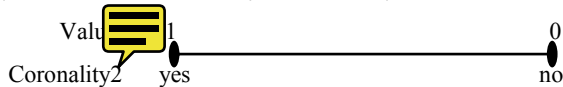
Presumably the strong preference for coronals corresponds to the preference for a smaller articulatory distance as their positioning in the oral cavity is central. (Post-)alveolar POA mediates between labial and dorsal, suggesting that it is optimal in cluster formation. From this generalization, scale is derived. The highest value of 1 is given to clusters containing 2 coronal consonants, followed by 0.5 and 0 points assigned to clusters with 1 or 0 coronal.

(3) Scale *Coronality 1 (presence of a coronal)*



Coronal in C1: Coronality is relevant also in terms of the emergence of a (post-)alveolar segment in the cluster-initial position. 31 clusters contain a coronal in C1, 12 start with a labial and 10 with a dorsal. Out of 45 CCs, 23 begin with a coronal. Again, labial-initial (12) and dorsal-initial (10) clusters are found to equal degrees. All CCCs start with a (post-)alveolar segment. The observed strong tendency for coronals to be in C1 is expressed by binary POA scale. One extreme, marked 'yes', receives value 1 for the initial coronal. Others score 0.

(4) Scale *Coronality 2 (coronal in C1)*



3.3. Manner of articulation

The five MOA classes in Table 2 yield 25 logical combinations for CC, and 125 for CCC. However, only a small subset of the possible combinations is attested, namely 10 for CC and 2 for CCC clusters. Below we discuss three parameters defining the MOA dimension.

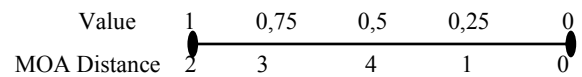
MOA Distances: Articulatory distances are relevant for the MOA dimension. The distance of 1 is assumed for consecutive manners given in Table 2. Segments belonging to the same MOA class have distance 0, as in fric+fric (e.g. /sv/). Adjacent consonants in a cluster can have distances of 2, 3 and 4, as in fric+liq (e.g. /fɹ/), plos+nas (e.g. /km/), and plos+liq (e.g. /pl/). CCC distances are computed by averaging the sum of distances for C1-C2 and C2-C3, as in /skv/=(2+2):2=2, and /spr/=(2+4):2=3.

Medial MOA distances are preferred in onset clusters, with distance 2 being represented by the

largest number. The data contains 19 clusters with distance 2, 15 with 3, and 6 with 1. Types with the largest distance 4 are found in 10 clusters of the plos+liq pattern. The other extreme is represented by clusters with 0 distance. Of all possible identical MOA CC sequences (5), only one pattern is attested, namely fric+fric in /sv/, /sf/, /ʃv/. 7 out of 8 CCCs follow a single fric+plos+liq pattern with distances of 2 for C1-C2 and 4 for C2-C3. The only exception constitutes the fric+plos+fric pattern in /s/ with a smaller distance for C2-C3 (2).

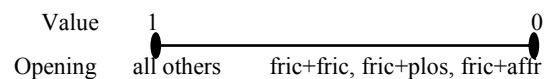
Scale reflects these preferences. A value of 0 is assigned to clusters with the distance of 0 (fric+fric) due to their scarcity. The opposite end is marked by the most abundant group of clusters with distance 2 (plos+fric, fric+plos, fric+liq, fric+plos+fric). The intermediate point values 0.75, 0.5 and 0.25 encompass clusters with distances 3 (affr+liq, plos+nas, fric+plos+liq), 4 (plos+liq), and 1 (affr+fric, fric+affr, fric+nas), respectively.

(5) Scale *MOA Distance*



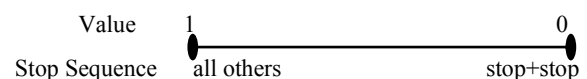
Increase of opening: For most of the clusters, the degree of opening increases throughout the sequence (with 18 counterexamples). Therefore, clusters in which C2 or C3 cause more airflow resistance than the initial segment are assigned the zero value, in contrast to all the remaining clusters, which score 1 in scale.

(6) Scale *Increase of Opening*



Stop sequence: As shown in Table 2, no cluster contains two adjacent oral or nasal stops. On this basis we derive another preference, according to which all stop+stop clusters are disfavoured. The dispreferred combination plos+liq (/gm, gn, km, kn, pn, tm/) receives 0 on scale, while all the other clusters are assigned value 1.

(7) Scale *Stop Sequence*

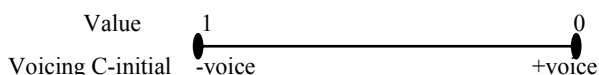


3.4. Glottis features

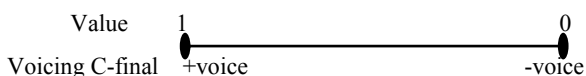
German uses a glottis-related distinction phonemically only for plosives and fricatives. The phonetic nature of this distinction has been a contentious issue but is not crucial for current purposes. Analysis of [voice] allows for two

generalizations: out of 53 clusters in Table 1, 44 begin with a voiceless segment, while 43 end with a voiced one. From the 8 logical possibilities for CCC, only one is attested: [-voice][-voice][-voice]. This observation translates into scales and, in which the preferred settings for [-voice] C-initially and [+voice] C-finally are assigned the value equal 1.

(8) Scale C-initial Voicing



(9) Scale C-final Voicing



4. RANKING OF CLUSTERS

The sonority-based approach is here extended to account for specific parameters of the following 4 dimensions relating to cluster description: (a) size, (b) place of articulation (POA), (c) manner of articulation (MOA), and (d) voice. 3 parameters are proposed for POA and MOA, 2 for voice and 1 for size. For each of these 9 parameters, a preferability scale is established, with values ranging from 1 to 0. Every cluster is assigned a value for each scale leading to a ranking of all clusters in terms of preferred options. Applying scales to all clusters by summing up their values results in an overall value for each cluster, as presented in Table 3. Clusters are grouped according to their cumulative values and assigned a rank.

Table 3: Ranking of clusters.

Rank	Value	Clusters
1	9	ʃl
2	8,5	tv
3	8,25	ʃn, ʃʁ
4	8	sl, sʁ
5	7,75	sm, tsv
6	7,5	fl, ʃm, tʁ
7	7,25	pfl, sn, tm
8	7	pl, ʃt
9	6,75	kl
10	6,5	dʁ, kʁ, kv, ps, sp, sv, ʃk, vl
11	6,25	fʁ, ks, pn, sk, spl, stʁ, ʃp, ʃpl, ʃv
12	6	bl, kn, pʃʁ, skl, skʁ, skv, st, ʃtʁ
13	5,75	gl, pʁ
14	5,5	gʁ, sf, ʃpʁ
15	5,25	km, sts, vʁ
16	5	gn
17	4,75	bʁ
18	4,25	gm

In contrast to the sonority hierarchy, the analysis yields a ranking providing a more fine-grained evaluation of all clusters. Some of the clusters in Table 3 (/tv/, /skv/) occur only in loan words. However, as discussed in section 7, German allows for clusters to be borrowed, provided that the constituent segments are always part of the phoneme system. Scales suggested are formulated on the basis of preferences pertaining to German; their language-specific or crosslinguistic nature needs to be evaluated. Extensions of the proposal would consist in adding weights to the scales, in the study of other typologically different languages, and possible correlations with type and token frequencies.

Perception gains from sufficient contrast. The present study demonstrates that contrast in onset clusters is needed, but is not maximised. Maintaining the distance in terms of POA and MOA is optimal for the interplay of perception and production. POA strives for minimal, but not zero distance, whereas MOA strives for medial distance. This division of labour serves verbal communication to be both speaker- and listener-friendly.

5. REFERENCES

- [1] Clements, G.N. 1990. The role of the sonority cycle in core syllabification. In Kingston, J., Beckman, M. (eds.), *Laboratory Phonology 1*. Cambridge: Cambridge University Press, 283-333.
- [2] Hall, T.A. 1992. *Syllable Structure and Syllable-related Processes in German*. Tübingen: Max Niemeyer Verlag.
- [3] Kohler, K.J. 1995. *Einführung in die Phonetik des Deutschen*. Berlin: Erich Schmidt Verlag.
- [4] Ladefoged, P. 1993. *A Course in Phonetics*. Fort Worth: Harcourt Brace Jovanovich.
- [5] Meinhold, G., Stock, E. 1980. *Phonologie der Deutschen Gegenwartssprache*. Leipzig: VEB Bibliographisches Institut.
- [6] Ohala, J.J. 1990. Alternatives to the sonority hierarchy for explaining segmental sequential constraints. *Annual Regional Meeting, Chicago Linguistic Society* 26, 319-338.
- [7] Sievers, E. 1881. *Grundzüge der Phonetik*. Leipzig: Breitkopf & Härtel.
- [8] Wiese, R. 1996. *The Phonology of German*. Oxford: Oxford University Press.
- [9] Wortschatz-Portal. <http://wortschatz.uni-leipzig.de/>

i We gratefully acknowledge the support of Barbara Samlowski and Bernd Möbius (Stuttgart, Bonn).