Renate Raffelsiefen, Anja Geumann

Phonological Analysis at the Word Level: The Role of Corpora

Abstract Notions such as "corpus-driven" versus "theory-driven" bring into focus the specific role of corpora in linguistic research. As for phonology with its intrinsic focus on abstract categorical representation, there is a question of how a strictly corpus-driven approach can yield insight into relevant structures. Here we argue for a more theory-driven approach to phonology based on the concept of a phonological grammar in terms of interacting constraints. Empirical validation of such grammars comes from the potential convergence of the evidence from various sources including typological data, neutralization patterns, and in particular patterns observed in the creative use of language such as acronym formation, loanword adaptation, poetry, and speech errors. Further empirical validation concerns specific predictions regarding phonetic differences among opposition members, paradigm uniformity effects, and phonetic implementation in given segmental and prosodic contexts. Corpora in the narrowest sense (i.e. "raw" data consisting of spontaneous speech produced in natural settings) are useful for testing these predictions, but even here, special purpose-built corpora are often necessary.

Keywords Speech corpora, German vowels, phonological grammar, abstractness, Optimality Theory

1 Introduction

Phonology is concerned with capturing the contrastive potential of a language, aiming at a comprehensive account of the ways in which differences in meaning can be conveyed through sound differences. Traditionally, a phonological description includes an inventory of phonemes, organized in terms of oppositions or distinctive features, along with rules for the combination and prosodic organization of the phonemes. Such a description then determines the lexical phonemic representations of words, which form the input to phonetic implementation.

The key intuition guiding phonemic analyses concerns a basic classification of linguistic material in terms of sameness versus distinctness, focusing on conditions for determining whether or not

– phonetically distinct sounds represent the same phoneme
– phoneme pairs represent the same (i.e. “proportional”) opposition

The answer to the first question again crucially refers to the notion of sameness since proof of phonemic distinctness presupposes the occurrence of distinct sounds in identical contexts. Applying this condition to German typically results in an inventory of fifteen or more vowel phonemes, which are then investigated and associated with IPA-symbols. Two descriptions with vowels arranged in accordance with IPA-conventions, one proposed by Kohler (1999: 87), see (1a), the other by Eckert & Barry (2005: 111), see (1b), are shown below.

(1) a. b.

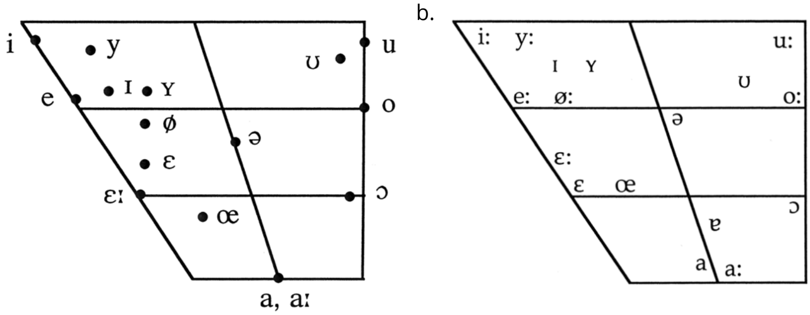

Figure 1: a. The German vowel system according to Kohler (1999). b. The German vowel system according to Eckert and Barry (2005).

While Eckert & Barry posit a vowel /ɐ/ to represent the unstressed syllable in words like *Vater* ‘father’, Kohler apparently considers that sound the same as other independently established phonemes. There is agreement that two vowel pairs differ in quantity only (i.e. /a/-/aː/ as in *prallen* ‘to bump’ - *prahlen* ‘to boast’, /ɛ/-/ɛː/ as in *stellen* ‘to put’ - *stählen* ‘to steel’), in contrast to all other oppositions, which are deemed to involve no phonemic quantity contrast (cf. 1a) or one linked to quality contrasts (cf. 1b). More radically different assessments of vowel sameness are seen in the works of others, including the view that there are no more than eight distinct vowels in German (Vennemann 1991, Becker 1998).

How can corpora help decide among such phonemic analyses or help evaluate the merits of abstract representation in general? Is there hope that ever larger corpora of spontaneous speech, subjected to ever more precise measurements and ever more sophisticated statistical modeling, could further our understanding of phonemic structure? How can quantitative methods capture the notion of phonemic sameness, which is rooted in the intuition that physical differences are abstracted away from and items are classified the same as long as those differences can be attributed to context? How can such methods capture abstractions in the minds of speakers which clearly are not amenable to direct measurement?

The approach to pinpointing phonemic structure to be illustrated below is rooted in the idea of a phonological grammar as a language-specific ranking of universal constraints (Prince & Smolensky 1993). While phonemic structure and the concept of abstractness are rarely addressed in such frameworks, we will argue that the interaction of constraints and their inherent properties yield insight into such structure. On this approach the focus shifts to data resources that shed light on constraints and their effects on phonological structure. Empirical support comes from the convergence of various types of independent evidence.

The paper is organized as follows. Section 2 presents some basic claims of constraint-based grammars, illustrating these with the role of roundedness in the vowel system of German. Section 3 focuses on the relevance of constraints in distinguishing between phonemic and subphonemic structure, to be illustrated with length versus quality differences in German vowels. Section 4 discusses some of the currently existing resources.

2 Constraint-based grammar: some basic ideas

Optimality Theory envisions phonological grammar as language-specific resolutions of conflicts among universal constraints (Prince & Smolensky 1993). The core conflict concerns the desirability to maximize contrast, by allowing all types of structure to distinguish morphemes, versus the desirability to minimize phonological markedness, to enhance ease of production and perception. Additional constraints concern correspondence of structure among words, requiring sameness of structure both at the syntagmatic level, to enhance cohesion (e.g. rhymes, alliteration), and at the paradigmatic level, to minimize allomorphy and enhance recognition of paradigmatic relatedness.

To illustrate a language-specific resolution of the core conflict between the maximization of potential contrast and satisfaction of markedness constraints, consider the roundedness contrast in German in (2). The stressed vowels are represented without duration marks, as duration will be argued to be a subphonemic property in German (cf. section 3).

(2) a. /ˈʃpilən/ <spielen> ‘to play’
/ˈkɪsən/ <Kissen> ‘pillow’
/ˈlezən/ <lesen> ‘to read’
/ˈkɛnən/ <kennen> ‘to know’

b. /ˈʃpylən/ <spülen> ‘to rinse’
/ˈkʏsən/ <küssen> ‘to kiss’
/ˈløzən/ <lösen> ‘to solve’
/ˈkœnən/ <können> ‘to be able to’

Contrastiveness as in (2) motivates the assumption of an active faithfulness constraint FAITH([±round]). Formally, such a constraint concerns the relation between an input and the corresponding output, requiring the "faithful" preservation of the input structure. To maximize potential contrast, it would be ideal if roundedness were contrastive for all vowels, including low and back vowels. The restriction of this contrast to the vowel pairs illustrated in (2) indicates a specific interaction among FAITH([±round]) and phonological markedness constraints prohibiting the cooccurrence of the feature [±round] with other features (e.g. *V{[+back][-round]} (Back unrounded vowels are prohibited), *V{[-back][+round]} (Front rounded vowels are prohibited)). The ranking in (3) says that in German for back and low vowels, it is more important to satisfy the relevant markedness constraints than to exploit the contrastive potential of lip roundedness. Only for non-low front vowels is the potential for contrast valued more than the satisfaction of the relevant markedness constraint (i.e. *V{[-back][+round]}). (Constraint domination is marked by the symbol ">>".)

(3) *V{[+back][-round]}, *V{[+low][+round]} >> FAITH(V[±round]) >> *V{[-back][+round]}

Phonological markedness constraints are presumably ultimately grounded in phonetics, expressing relative difficulties in articulating or perceiving certain structures compared to others (e.g. specific coordinations between tongue positions and lip roundedness). They are reflected in asymmetries in the distribution of sounds in the languages of the world documented in databases such as UPSID[1], which is based on 317 languages. Links between lip roundedness and tongue advancement are shown by the fact that 94 % of front vowels are unrounded whereas 93.5 % of back vowels are rounded (Maddieson 1984: 124). Among the vowels classified as low central monophthongs in the languages in question, 392 unrounded compare to a single rounded vowel (Maddieson 1984: 124).

A representation of phonological grammar in terms of rankings among universal constraints as in (3) is superior to a mere listing of phonemes in that it relates the actual to the potential. Such a model predicts that more marked structure (e.g. rounded front vowels in German) implies the existence of the corresponding less marked structure (e.g. unrounded front vowels in German). This is because there is no ranking of independently motivated markedness constraints which would describe a language where marked structures exist to the exclusion of the corresponding less marked structures.

1 This acronym stands for *UCLA Phonological Segment Inventory Database* (Maddieson & Precoda 1990). For more discussion see section 4.

To verify the existence of the respective less marked phonemes it is necessary to establish the relevant relations and to demonstrate the presence of consistent phonetic correlates. The relations in question are supported by correspondence patterns, including regular sound alternations in paradigms[2] and also so-called impure rhymes, which are characterized by specific relaxations of a general requirement for sameness. Consider the German word pairs in (4)[3], which function as rhymes despite the difference in vowel roundedness. These rhymes then support the specific phoneme correspondences illustrated by the minimal pairs in (2).

(4) /y/ : /i/ *grüßen* 'to greet' – *fließen* 'to flow'
/ʏ/ : /ɪ/ *Sünder* 'sinner' – *Kinder* 'children'
/ø/ : /e/ *schön* 'beautiful' – *stehn* 'to stand'
/œ/ : /ɛ/ *Töchter* 'daughters' – *Wächter* 'guard'

Reference to the feature [±round] in the grammar stated in (3) to capture the vowel opposition illustrated in (2) is motivated by the relevance of the respective markedness constraints. A consistent phonetic difference is confirmed by observing the degree of lip roundedness during the articulation of the vowels in each pair in (4). However, not all phonetic reflexes are easily assessed on an introspective basis and in general there are many advantages to conducting phonetic studies based on acoustic measurements. Such studies concern the resonances, known as formants, which change according to the size and the shape of the vocal tract thereby reflecting on articulatory properties (Peterson & Barney 1952). For example, the first formant frequency (F1) increases as the tongue lowers. F1 decreases, while the second formant frequency (F2) increases, as the tongue body advances. All formant frequencies, especially F2 and F3, decrease with increased lip roundedness as a result of the concomitant elongation of the vocal tract (Hixon et al. 2008).

Regarding the pairs in (4), there is accordingly a prediction that for each unrounded vowel, the values for F2 and F3 should be higher than those for the corresponding rounded vowels. This prediction is borne out by the measurements of the relevant vowel formants based on recordings of 26 female speakers in the Kiel Corpus of Read Speech (cf. section 4).[4]

2 Correspondence involving paradigmatic relations can be illustrated with plural-singular pairs (e.g. /ʃtylə/ <Stühle> 'chairs' – /ʃtul/ <Stuhl> 'chair', /flʏsə/ <Flüsse> 'rivers' – /flʊs/ <Fluss> 'river'), which confirm the existence of a less marked rounded back vowel corresponding to each more marked front rounded vowel.

3 These rhymes are adopted from the poem "Romanzen vom Rosenkranze" by Clemens Brentano.

4 Formant values were extracted automatically with PRAAT (Boersma & Weenink (2016)) at 50 % of the vowel duration. The numbers of tokens for individual vowels (stressed and unstressed) are as follows: /i/ 1215, /y/ 289, /ɪ/ 2,536, /ʏ/ 264, /e/ 978, /ø/ 149, /ɛ/ 1,070, /œ/ 245.

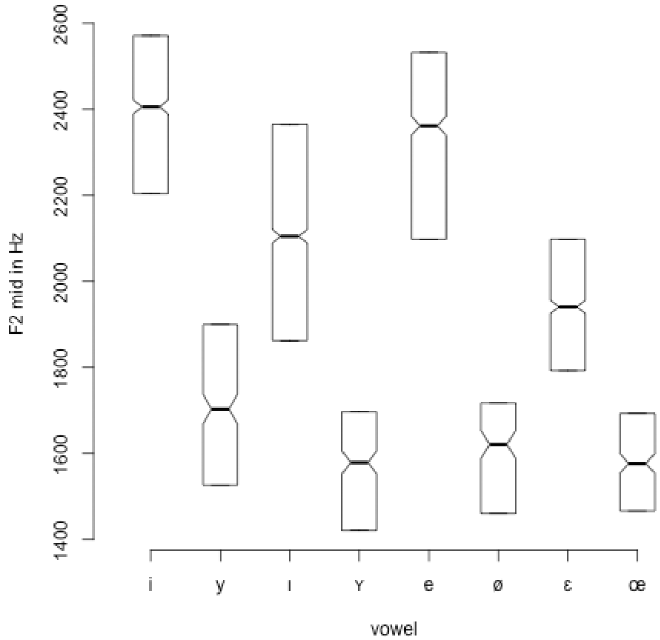

Figure 2a: *Kiel Corpus* 26 f speakers, Formant F2 values in Hz.

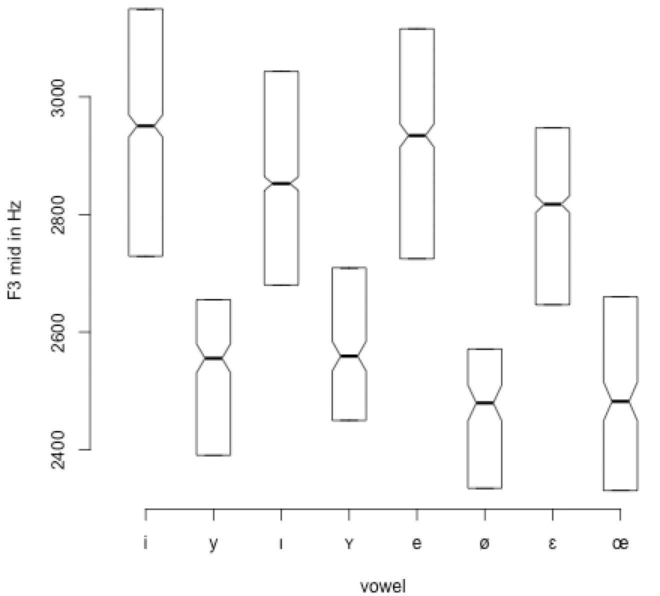

Figure 2b: *Kiel Corpus* 26 f speakers, Formant F3 values in Hz.

The boxplots in Figure 2 show non-overlapping indentations for all relevant pairs (e.g. /i/–/y/), which means that the median values differ significantly (Chambers et al 1983).[5]

A comparison of the four vowel pairs in Figure 2 shows that the respective differences among the formant values differ considerably. For instance, the pair /e/–/ø/ exhibits a larger difference among the values for both F2 and F3 than the pair /ɛ/–/œ/. Such disparities are consistent with the assumption of a single phonological opposition as long as they can be attributed to independent differences (e.g. larger difference among F2 values in pairs of peripheral vowels (/e/–/ø/, /i/–/y/) compared to the corresponding pairs of centralized vowels (/ɛ/–/œ/, /ɪ/–/ʏ/)). The analysis of all of the relevant pairs as instances of a single phonological roundedness opposition is expressed in terms of positing a single faithfulness constraint FAITH(V[±round]) and its interaction with other constraints as in (3).

The claim that the constraint ranking in (3) captures the role of roundedness in German phonology is supported by independent evidence concerning historical change.

Here again, we find an asymmetry to the effect that an increase of markedness (the emergence of rounded front vowels) comes about through context-sensitive change whereas context-free change consistently leads to a decrease of markedness (unrounding of front vowels). This generalization can be illustrated with the development of the English verb *kiss* in (5), where an increase in markedness (/ʊ/ => /ʏ/) results from assimilation (fronting of /ʊ/ to agree with the following front vowel /i/). The subsequent loss of rounding in front vowels (/ʏ/ => /ɪ/) is context-free and reduces segmental markedness:

(5) Old Saxon *kussian* > Old English *cyssan* > Modern English *k*/ɪ/*ss* <kiss>

Additional sources of front rounded vowels in German are illustrated in (6).[6] The sporadic changes from less marked to more marked vowels invariably involve segmental contexts consisting of labial fricatives [v], [f] or [ʃ][7], all of which favor the perception of a rounded vowel:

(6) MHG *wirde* > NHG *W*/ʏ/*rde* <Würde> ‘dignity’
MHG *vinf* > NHG *f*/ʏ/*nf* <fünf> ‘five’
MHG *zwelf* > NHG *zw*/œ/*lf* <zwölf> ‘twelve’
MHG *lewe* > NHG *L*/ø/*we* <Löwe> ‘lion’
MHG *leschen* > NHG *l*/œ/*schen* <löschen> ‘to extinguish’

5 Outliers are not presented in the boxplots but are included in the calculations.

6 The changes are sporadic as unrounded vowels are often preserved in the contexts in question (e.g. NHG *W*/ɪ/*rbel* <Wirbel> ‘whirl’, NHG *W*/ɛ/*lle* <Welle> ‘wave’).

7 [ʃ] is pronounced with strongly protruded lips in German (cf. Wängler 1964).

In other contexts, changes involving roundedness consistently favor unmarked unrounded front vowels. The following changes concern vowels spelled with the grapheme <y>, which is historically linked to rounded /y/ or /ʏ/, but, unlike the grapheme <ü>, also associates with unrounded vowels in German (s. Dudenband 6: 913).

(7) G/ʏ/mnásium > G/ɪ/mnásium <Gymnasium> 'secondary school'
s/ʏ/mpátisch > s[ɪ]mpátisch <sympathisch> 'likable'
S/ʏ/stém > S[ɪ]stém <System> 'system'

The asymmetry in historical change illustrated above is predicted by the grammar in (3) if one were to assume inputs consisting of actual word forms encountered by hearers. Faithfulness constraints would then make their force felt only if a given sound property has been perceived. Otherwise markedness prevails and the unmarked segments will emerge. This approach also makes sense of the fact that reanalysis to unmarked vowels as in (7) is more common in unstressed positions because stressed syllables favor the perception of contrasts (cf. the stability of roundedness in words like *'P*/y/*thon* <Python> 'python', *'G*/y/*ros* <Gyros> 'gyros'). The connection in question can be expressed by way of linking faithfulness constraints to prominent positions (e.g. $\text{FAITH}_{\text{stress}}$) and by imposing a universally fixed ranking to the effect that $\text{FAITH}_{\text{POS}}$ (POS = "prominent position") dominates the corresponding general faithfulness constraint. This phenomenon, known as "positional faithfulness" (Beckman 1998), is also relevant to the analysis of speech errors illustrated in (8)[8], which appear to favor the alignment of marked structures with prominent positions. The correct and presumably intended forms are given in parenthesis.

(8) M[ɪ]s't[øː]rium (M[ʏ]s't[eː]rium <Mysterium> 'mystery')
S[ɪ]n't[øː]se (S[ʏ]n't[eː]se <Synthese> 'synthesis')
S[ɪ]l'v[œ]ster (S[ʏ]l'v[ɛ]ster <Sylvester> 'New Year's Eve')
Di[e]'z[øː]se (Di[ø]'z[eː]se <Diözese> 'diocese')
Z[i]'l[ʏ]nder (Z[y]'l[ɪ]nder <Zylinder> 'cylinder')

A phonological grammar in terms of ranked constraints as in (3) accounts for both the distribution of phonemes, thus capturing potential contrast, and the stability of phonological structure. Significantly, such a grammar provides clear guidance for research based on annotated speech corpora, singling out specific

8 The examples in (8) stem from personal communication, published speech error collections (Leuninger 1996), or common misspellings in internet data (e.g. *Zilynder*, *Silvöster*).

sound structures for comparison and focusing the investigation on the question of how certain abstract structures are implemented in various segmental and prosodic contexts. For instance, the juxtaposition of the measurements shown in Figure 2 indicates closer F2 and F3 values for roundedness contrasts involving centralized vowels, which may account for the higher rate of phonemic reanalysis for such vowels (cf. MND *flistern* > ˈ*fl*[ʏ]*stern* <flüstern> 'to whisper', but MND *vlise* > ˈ*Fl*[iː]*se* <Fliese> 'tile').

The data reviewed so far illustrate types of evidence to support grammatical descriptions in terms of interacting constraints as well as the use of speech corpora to verify the presence of consistent phonetic correlates. The following section illustrates ways in which evidence from constraint interactions can resolve questions concerning phonemic abstractness along with additional ways in which acoustic studies could verify such analyses.

3 Identifying phonemic oppositions

While there is a consensus that the minimal pairs listed in (2) illustrate a single rounding opposition, other cases raise substantial controversy. Recall the lack of consensus regarding the role of quantity versus quality in the analysis of German vowels addressed above. A complete list of relevant opposition members, represented phonetically in square brackets and referred to as "A-vowels" versus "B-vowels" for now, is illustrated in (9). The cases which have been claimed to involve a pure quantity opposition are listed in (9b), where the symbols /aː/ and /ɛː/ presented in the charts in (1) are replaced by symbols indicating quality differences (i.e. [ɑː] and [e*ː])

(9)	a.	A-vowels	B-vowels
		/m[iː]nə/ <Mine> 'mine'	/m[ɪ]nə/ <Minne> 'love'
		/d[yː]nə/ <Düne> 'dune'	/d[ʏ]nə/ <Dünne> 'thinness'
		/b[uː]lə/ <Buhle> 'paramour'	/b[ʊ]lə/ <Bulle> 'bull'
		/d[oː]lə/ <Dohle> 'jackdaw'	/d[ɔ]lə/ <Dolle> 'rowlock'
		/h[øː]lə/ <Höhle> 'cave'	/h[œ]lə/ <Hölle> 'hell'
		/ʃt[eː]lən/ <stehlen> 'to steal'	/ʃt[ɛ]lən/ <stellen> 'to put'
	b.	/ʃt[e*ː]lən/ <stählen> 'to steel'	?/ʃt[ɛ]lən/ <stellen> 'to put'
		/pʀ[ɑː]lən/ <prahlen> 'to boast'	/pʀ[a]lən/ <prallen> 'to bump'

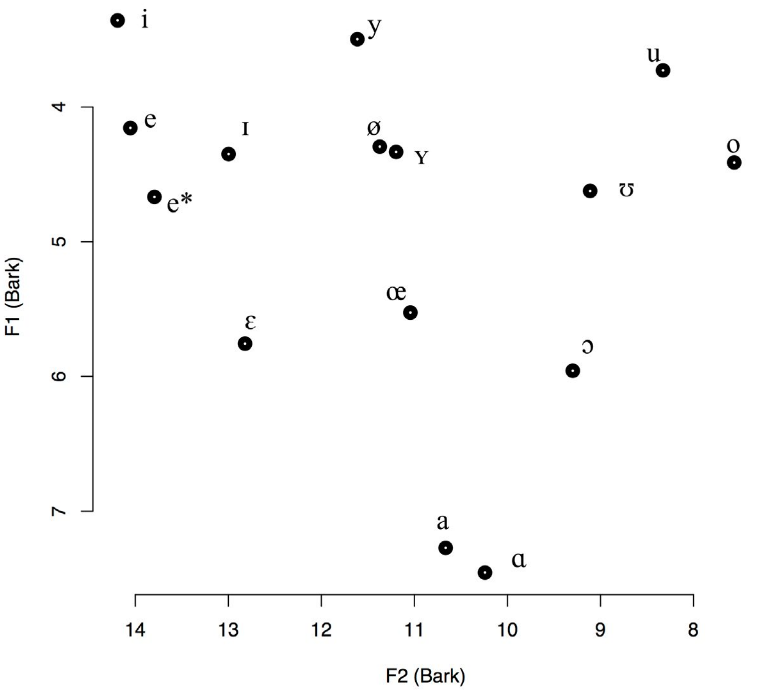

Figure 3a: Vowel chart F1/F2 plane in Bark for stressed German vowels.
Kiel Corpus of Read Speech, 26 female speakers.

Measurements of F1 and F2 for these vowels, again based on the female speakers of the Kiel Corpus, are given in Figure 3a. The respective values for duration are listed in Figure 3b.[9]

The values in Figure 3 are largely consistent with both phonemic analyses indicated in (1). It is doubtful that additional measurements, based on larger corpora, could answer the question of whether the length contrasts are phonemic for all, some, or no pairs. Indeed, none of the phonetic studies considered so far seem to offer a clear basis for deciding which vowels form opposition members in the first place. Proximity of positions within the formant charts alone is hardly decisive as for instance the vowels in /d[ʏ]nə/ <Dünne> 'thinness' versus /h[øː]lə/ <Höhle> 'cave' are represented with distinct symbols in all descriptions known to us, despite exhibiting greater similarity than any of those in (9b).

9 For our calculations we used the Burg algorithm, searching for 5 formants in the range from 0-5500 Hz for females. The number of tokens, all of them stressed, are as follows: /a/ 1,157, /ɑ/ 575, /ɛ/ 698, /e/ 619, /e*/ 36, /ɪ/ 645, /i/ 419, /ɔ/ 279, /o/ 231, /œ/ 81, /ø/ 138, /ʊ/ 324, /u/ 365, /ʏ/ 209, /y/ 205.

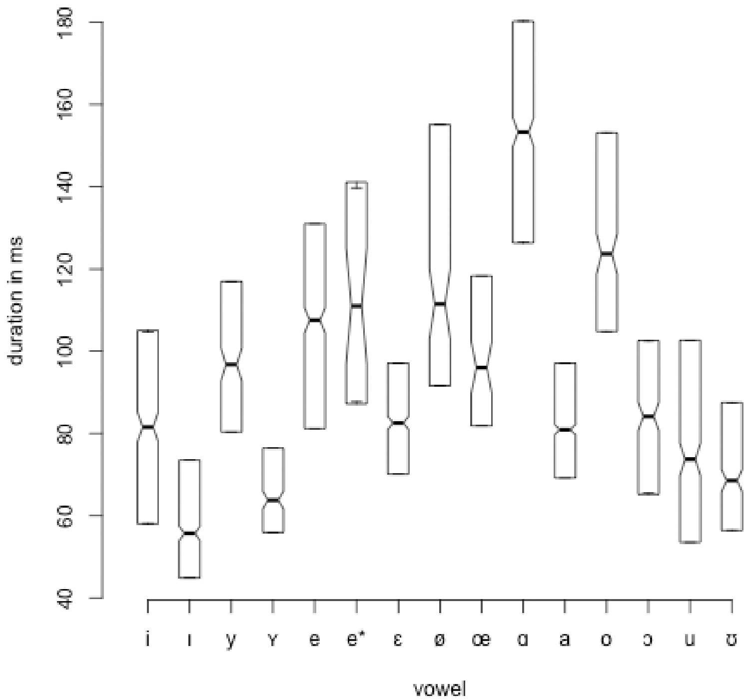

Figure 3b: Vowel durations in ms for stressed German vowels.
Kiel Corpus of Read Speech, 26 female speakers.

Below we will briefly indicate how a constraint-based approach may resolve these questions, focusing on the sort of data resources needed for establishing constraint interactions. The question of whether or not the pairs in (9) form a single opposition is addressed in section 3.1, while arguments for separating phonemic from subphonemic structure to identify that opposition are addressed in section 3.2. Arguments for identifying respective opposition members are reviewed in 3.3. Section 3.4 discusses corpus-based acoustic studies relevant to verifying the results.

3.1 Establishment of a single opposition

The analyzability of all vowel pairs in (9) as a single opposition depends on whether there are parallel restrictions indicative of single constraint interactions. The investigation focuses then on neutralization patterns, to establish the existence of contexts where all A-vowels can appear, to the exclusion of all B-vowels, and vice versa. One such context is given in (10a), as all A-vowels, but no B-vowels, occur before another syllabic vowel.

(10)	a.	/'n[ɑː]ə/ <nahe> ‘near’	b.	<< OHG nāh
		/'[eː]ə/ <Ehe> ‘marriage’		<< OHG ēwa
		/'ʀ[uː]ə/ <Ruhe> ‘quiet’		<< OHG ruowa
		/'m[yː]ə/ <Mühe> ‘effort’		<< OHG muohi
		/'dʀ[oː]ən/ <drohen> ‘to threaten’		<< OHG drouwen
		/'ʀ[iː]o/ <Rio> place name		Spanish [rrío]

The demonstration of systematic restrictions on phonological form is inherently problematic as it may seem to require an exhaustive examination of all relevant data. In addition, there is a possibility that the absence of specific patterns is synchronically accidental, caused by the imitation of the given and ultimately resulting from historical circumstances. Such conditions might fully account for the restrictions on the prevocalic vowels observed in (10a) as they go back to long vowels or diphthongs in Old High German (OHG) shown in (10b). Also in loan words the relevant structure could exist independently in the source language, adapted “faithfully” by the borrowers, without necessarily being represented in their phonological grammar.

There is a question then of which types of data are best suited to reveal genuine phonological restrictions caused by active phonological markedness constraints. All data involving potential modification of observable input structures are ideal as such modifications necessarily indicate the dominance of markedness constraints over faithfulness. Apart from cases of historical change and speech errors discussed above, the most significant sources include acronyms and the adaptation of loan words. The latter type is illustrated in (11), where apparent B-vowels in prevocalic position in the French source words are systematically replaced by A-vowels in German.[10]

(11) French /kl[ɔ]'ak/ <cloaque> ‘sewer’ => German /kl[o]'ɑkə/ <Kloake> ‘sewer’
French /n[ɔ]'ɛl/ <noël> ‘Christmas’ => German /n[o]'ɛl/ <Noël> ‘French Christmas carol’
French /p[ɔ]e'zi/ <poésie> ‘poetry’ => German /p[o]e'zi / <Poesie> ‘poetry’

10 The data in (11) raise a question concerning the status of the respective input and output forms. French acoustic forms could be mapped to forms perceived by German learners. Alternatively, French structures perceived by German speakers could be mapped to outputs they produce in speech. Either view involves modifications which presuppose an active markedness constraint.

Reference to loan words as a source for revealing active markedness constraints is potentially complicated by instances of so-called unassimilated loans illustrated in (12), which are characterized by special efforts on the parts of speakers to closely imitate the pronunciations of words in the respective donor languages.

(12) [nɔ'e] <Noé> French male given name
[nɔ'ɑ:j] <Noailles> French place name

Such cases are best set aside as long as they are limited to a few uncommon words (e.g. proper names) and/or are characterized by varying pronunciations. Conceivably, individual words can be imitated as wholes, with no impact on the phonological grammar. In general, the lack of modification of structure seen in unassimilated loan words does not contradict the assumption of active markedness constraints, but cases of systematic modification strongly support that assumption.

The second type of data mentioned above, acronyms, is also characterized by systematic restrictions on output forms which cannot be due to imitation of given forms. The data in (13) illustrate again the systematic exclusion of B-vowels in prevocalic position.

(13)	/'t^{s}[o:]at^{s}/	ZOAZ	Zentrales ([ɔ])/Organisations- und Abrechnungszentrum
	/'f[e:]ap/	VEAB	Volkseigener ([ɛ])/Erfassungs- und Aufkaufbetrieb für landwirtschaftliche Erzeugnisse
	/'f[i:]ak/	VIAG	Vereinigte ([ɪ])/Industrie-Unternehmen AG
	/'ʀ[i:]as/	RIAS	Rundfunk ([ɪ])/im amerikanischen Sektor

The restriction to the A-vowels marked in the acronyms in (13) can be linked neither to the vowels contained in the relevant source words (cf. the right-hand column in (13)), nor to conventions concerning grapheme-phoneme correspondence[11]. Although acronym formation or the adaptation of loan words may appear to be marginal phenomena, both are associated with highly regular modifications of sound structure which can be explained only by active phonological markedness constraints. A consistent convergence seen in the relevant output restrictions (e.g. the restriction to A-vowels in prevocalic position in both /kl[o]'ɑkə/ <Kloake> 'sewer' and /'t^{s}[o:]at^{s}/ <ZOAZ>) documented for all relevant opposition pairs may indeed suffice to establish systematic gaps in the distribution of phonemes, obliterating the need for an exhaustive examination of dictionaries.

11 For instance, the grapheme <E> associates with a lax vowel in /pɛk/ <PEG> (based on *Perkutane endoskopische Gastrostomie*) versus a tense vowel in /pekɪp/ <PEKIP> (based on *Prager Eltern-Kind Programm*).

As for the overall distribution of A- versus B-vowels in German, a thorough investigation of neutralization patterns indicates strictly parallel patterns within each class. For instance, the restriction to A-vowels, which are “tense” and phonetically long, in prevocalic position in (13) also extends to low vowels illustrated in (14):

(14) /'l[ɑː]ɔs/ Laos (possibly adopted from French [la'oːs] <Laos> ‘Laos’
/'t^{s}[ɑː]ɛt/ ZAED Zentralstelle für Atomenergie-Dokumentation

The exclusion of all phonetically short B-vowels in the stressed prevocalic position supports the presence of a single opposition. Many additional contexts can be found, where either only A-vowels occur, to the exclusion of all B-vowels or only B-vowels occur, to the exclusion of all A-vowels.[12] This parallelism strongly argues in favor of a single opposition distinguishing A- versus B-vowels, not a mixed system as suggested by Kohler’s depiction in (1a).

3.2 Identifying the nature of the opposition

As was noted above, the assumption of phonological markedness constraints rests on cross-linguistic asymmetries in the distribution of sounds. Their proper identification in individual languages is accordingly determined primarily by the overall neutralization patterns. As for the opposition of A- versus B-vowels in German, the observed restrictions suggest reference to syllable structure, invoking markedness constraints of the type “No B-vowels in open syllables”, “No A-vowels in closed syllables”. This particular context is consistent with a quality contrast, as is shown by the so-called *Loi de Position* in French, which bans vowels in word-final open versus closed syllables based strictly on their quality, regardless of length (e.g. “/o/ and /ø/, but no /ɔ/ or /œ/, in open syllables”, “/ɛ/, but no /e/ in closed syllables”). However, in general the syllable structure contexts in question may also be consistent with a quantity opposition, provided that rules known as “Open Syllable Lengthening” and “Closed Syllable Shortening” can in fact be shown to be neutralizing.

As for German, the syllable-based restrictions in question appear to target quality rather than quantity. This is because the relevant neutralization patterns are also observed in unstressed position, where all vowels are short (cf. the data in (11)). Moreover, there are additional neutralization patterns clearly betraying

12 The restriction to only B-vowels is for instance seen before sonorant-obstruent clusters which include a non-coronal segment (e.g. /'v[ɔ]lkə/ (*/'v[oː]lkə/) <Wolke> ‘cloud’, /'f[a]lkə/ (*/'f[ɑː]lkə/) <Falke> ‘falcon’).

reference to quality rather than quantity. The words in (15a) illustrate a restriction to B-vowels before a closed syllable containing a corresponding B-vowel, an apparent harmony effect as vowels before similar consonants in (15b) are not affected (cf. Raffelsiefen 2016).

(15) a. /'b[ɔ]tʀɔp/ <Bottrop> ‘place name’
/'n[ʊ]bʊk/ <Nubuk> ‘nubuck’
/'v[ɪ]tɪp/ <witib> ‘widow’
b. /'l[o:]tʀɪŋən/ <Lothringen> ‘Lorraine’
/'t[u:]bɑ/ <Tuba> ‘tuba’
/'v[i:]tɑ/ <Vita> ‘vita’

Evidence for an active harmony constraint in (15a) is highly significant as harmony is known to universally refer only to quality features, never to length. The data in (16a) illustrate the relevant harmony effect for low vowels, further demonstrating the parallel behavior of all A- versus B-vowels. The acronyms in (16b) show the synchronic productivity of the relevant restrictions.[13]

(16) a.	/'t[a]bak/ <Tabak> ‘tobacco’	/'ʀ[ɑ:]bə/ <Rabe> ‘raven’
	/'m[a]dras/ <Madras> ‘place name’	/'p[ɑ:]dʀə/ <Padre> ‘padre’
b.	/'h[a]pak/ <HAPAG>	/'[ɑ:]po/ <APO>
	/'t[a]kraf/ <TAKRAF>	/'n[ɑ:]gʀɑ/ <NAGRA>

Given the necessary reference to quality features to capture the neutralization patterns in (15) and (16), as opposed to the absence of cases where reference to quantity is needed to capture potential contrast patterns, the opposition referred to as A- versus B-vowels in (9) can be analyzed as a fundamental quality opposition. Vowel length constitutes then a subphonemic property. Additional research, including investigations of cross-linguistic patterns, is needed to properly identify the quality feature in question. Here, we tentatively choose the feature [± peripheral] (Lindau 1978). The restriction of the relevant contrast to stressed syllables indicates an active positional faithfulness constraint $\text{FAITH}(\pm\text{PER})_{\text{STRESS}}$ whose interaction with various markedness constraints captures the distribution of peripheral versus centralized vowels in German (Raffelsiefen 2016). Section 3.4 focuses on studies based on speech corpora suited to the empirical testing of the analysis.

13 HAPAG: Hamburg-Amerikanische Paketfahrt-Aktien-Gesellschaft, TAKRAF: Tagebauausrüstungen, Krane und Förderanlagen, APO: Außerparlamentarische Opposition, NAGRA: Nationale Genossenschaft für die Lagerung radioaktiver Abfälle.

3.3 Identifying corresponding opposition members

As was noted above, the identification of individual opposition members is supported by evidence pertaining to violations of strict correspondence constraints pertaining to both paradigmatic and syntagmatic relations. Paradigmatic correspondences are illustrated in (17), where a centralized vowel in an unstressed closed syllable alternates with a peripheral vowel in an unstressed open syllable. The vowel alternation is caused by a vowel-initial suffix carrying main stress, which conditions the syllabification of the preceding consonant as an onset, as opposed to the coda syllabification of the corresponding consonant in the base. Stresslessness is crucial as stressed vowels exhibit regular paradigm uniformity effects (see 3.4). The derived formations in (17b) are marked with question marks because they are not attested.[14]

(17) a. *Ják*[ɔ]*b* <Jakob> 'male name' *Jak*[o.]*bíner* <Jakobiner> 'Jacobin'
Tíb[ɛ]*t* <Tibet> 'Tibet' *Tib*[e.]*táner* <Tibetaner> 'Tibetan'
Lím[ɪ]*t* <Limit> 'limit' *lim*[i.]*tíeren* <limitieren> 'to limit'
Sább[a]*t* <Sabbat> 'Sabbat' *Sabb*[ɑ.]*tíst* <Sabbatist> 'sabbatist'
b. *Kál*[ʏ]*m* <Kalym> 'kalym' ?*kal*[y.]*míeren*
Báf[œ]*g* <Bafög> 'funding for students' ?*baf*[ø.]*gíeren*

The evidence from the paradigmatic alternations in (17) agrees with evidence pertaining to rhyme. The examples for assonance in (18) exhibit identical values for all contrastive vowel features other than [±peripheral].[15]

(18) ˈ[ʊ]*nter* <unter> 'under' – ˈ*Gr*[u:]*be* <Grube> 'pit'
ˈ*S*[ʏ]*nde* <Sünde> 'sin' – ˈ*w*[y:]*hlen* <wühlen> 'to rummage'
ˈ*f*[a]*ngen* <fangen> 'to catch' – ˈ*gr*[ɑ:]*ben* <graben> 'to dig'
ˈ*tr*[ɛ]*ffen* <treffen> 'to meet' – ˈ*L*[e:]*hrer* <Lehrer> 'teacher'
ˈ*M*[ɛ]*sser* <Messer> 'knife' – ˈ*Tr*[e*:]*nen* <Tränen> 'tears'

Assuming that the stressed vowels in the examples *Tränen* and *Lehrer* cited in (18) are indeed distinct, the assonance patterns support the correspondence relations indicated in (9), where both of the peripheral vowels in question correspond to centralized /ɛ/.

14 The relevant alternations ought to also be tested experimentally, ideally with illiterate speakers to exclude possible correspondence effects pertaining to graphemes.

15 The examples in (18) are also adopted from Brentano's poem "Romanzen vom Rosenkranze".

3.4 Verifying phonological analyses

The establishment of a single quality opposition for the vowel pairs illustrated in (9) predicts the presence of a consistent phonetic correlate. The measurements in Figure 4 are based on all 15 vowels in stressed position pronounced by female speakers in the *Kiel Corpus*[16] and demonstrate that each A-vowel is more peripheral than the corresponding B-vowel. In particular, it is shown that for a specific central position the peripheral vowel is always further away than the corresponding centralized vowel. The central position is calculated individually as the mean F1 and F2 value for all relevant vowels in a given (sub)corpus. The distance is then calculated as the Euclidian distance to the central position for each individual vowel in the F1 by F2 vowel space.

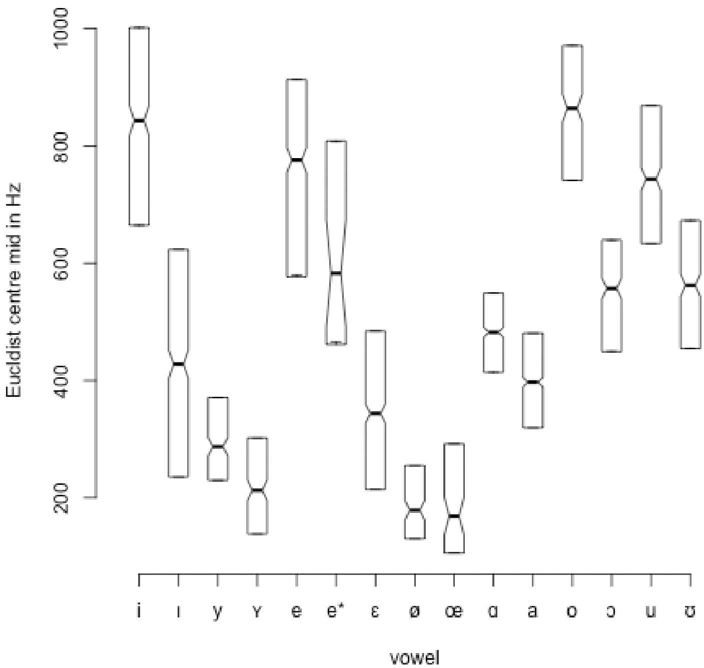

Figure 4: Boxplot of Euclidian distance in Hz for A- vs. B-vowel pairs for all stressed German vowels. *Kiel Corpus of Read Speech*, 26 female speakers.

16 See footnote 9.

The results in Figure 5 focus on the relations between the pairs /ɑ/–/a/, /e*/–/ɛ/, and /e/–/ɛ/, two of which have been claimed to exhibit a pure quantity contrast (cf. (1)). Our measurements show that all of these pairs exhibit the expected phonetic correlate, in accordance with their analysis as part of a single quality opposition on German.

The objection that at least some varieties of standard German might have a pure quantity contrast, at least for some oppositions, calls for a detailed study, focusing on the speech of maximally homogeneous groups or even individuals. This is because for phonetically similar sounds there is a danger that significant differences in the pronunciation of individuals become obscured by merging data. Even for a single speaker, systematic differences can be obscured by merging results pertaining to different segmental and prosodic contexts. The data in Figure 6 are based on the OLLO speech corpus, which contains minimally contrasting segment strings (cf. section 4). They demonstrate significant differences for the relative Euclidian distances for the /a/:/ɑ/ contrast compared in various segmental contexts, indicating for instance stronger contrasts in velar compared to labial contexts.

The ideal phonological corpora for establishing phonemic contrast are based on carefully controlled studies, where simplexes appear in identical carrier sentences and the speech of individuals can be examined separately.[17] In general, it holds that the demonstration of significant phonetic differences in a single context for a single speaker suffices to establish an active FAITH constraint in the phonological grammar of that individual.

Apart from demonstrating consistent phonetic correlates for phonological oppositions, there are various additional ways to test phonological analysis with speech corpora. The analysis predicts specific vowel qualities in the neutralization contexts, including only centralized vowels in unstressed closed syllables as in ˈ*Gyr*[ɔ]*s* ‘gyros’ or only peripheral vowels in unstressed open syllables as in *B*[i]*kín*[i] ‘bikini’, [ɑ]*lásk*[ɑ] ‘Alaska’. All subphonemic properties are predicted to conform to certain contextually determined restrictions such as only enhancement (rather than weakening) of gestures in strong prosodic positions (e.g. possible lengthening, never shortening, of vowels in stressed syllables). Subphonemic properties are further predicted to not exhibit paradigm uniformity effects (e.g. no difference in vowel length for the first vowel in *platónisch* ‘Platonic’ and *Platáne* ‘plane tree’, despite the presence of a long vowel in the base ˈ*Pl*[ɑː]*to* ‘Plato’). At the same time, it is predicted that phonemic structure, including quality contrasts concerning the feature [±peripheral], can show paradigm uniformity effects (e.g. a peripheral unstressed vowel in plural ˈ*Aut*[o]*s* ‘cars’, distinct

17 cf. the formant maps in Ramers 1988: 181ff

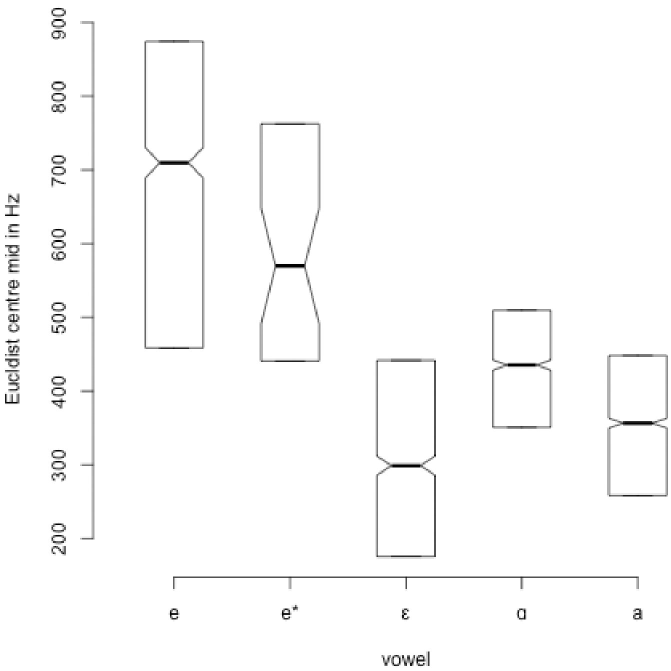

Figure 5: Boxplot of Euclidian distance in Hz for /e/, /e*/, /ɛ/, /ɑ/, /a/. *Kiel Corpus of Read Speech,* 26 female speakers.

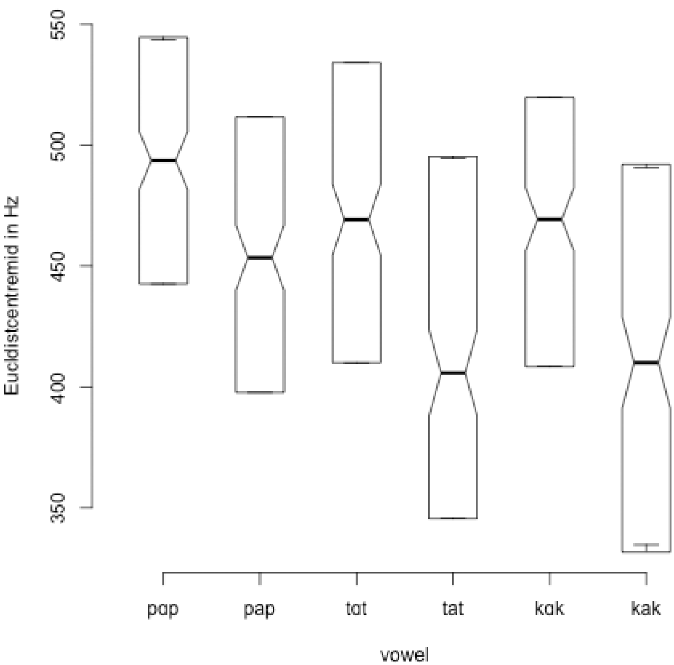

Figure 6: Boxplot of Euclidian distance in Hz for /a/ vs. /ɑ/ with different consonant contexts, each represented with 176–180 tokens. *OLLO corpus,* 5 Bavarian female speakers.

from the centralized vowel in '*Gyr*[ɔ]*s* ‘gyros’, to match the peripheral vowel in the singular '*Aut*[o] ‘car’). For some preliminary studies of these types to verify the [± peripheral] opposition of German vowels, see Raffelsiefen (2016).

4 Data resources

Below we will briefly discuss the data resources used in our research: speech corpora, typological databases, electronically searchable word lists, and various word collections.

The corpora mentioned above, the *Kiel Corpus* (Kohler 1994) and *OLLO* (Wesker et al. 2005), have the advantage that they are provided with complete corrected segmental annotations but differ greatly in scope. The Kiel Corpus contains recordings of read connected speech, including 31,000 word tokens from 53 native speakers of German. *OLLO* contains recordings of read nonce words of the type CVC and VCV, presented in conventional German orthography (e.g. <pahp>, <papp>). It is based on 40 speakers divided into four separate regions and contains 2,700 recorded tokens per speaker. While confined to a subset of German phonemes, and arguably not containing German language material proper, the highly controlled environments yield valuable information about subtle contrasts, contextual influences, and regional differences.

A third corpus for German we frequently use is *Deutsch Heute* (Brinckmann et al. 2008), which includes recordings of roughly 1,000 words, including many loanwords, by 670 speakers covering all German-speaking areas. This corpus is well-suited to studying regional variation. It is, however, not suited to studying subtle contrasts as there are almost no minimal pairs and the words are read without carrier sentences. Moreover, some of the material is currently provided only with automatic segmental annotation, which needs to be corrected manually. We resort to special purpose-built corpora when necessary to study subtle phonological contrasts or specific paradigm uniformity effects.

Generally speaking, annotations cannot be assumed to be adequate, even when manually corrected. For instance, annotations for the *Kiel Corpus* mark all word-final full vowels as long, regardless of stress. As a result, a word like *Alaska* ‘Alaska’ has identical representations for the first two vowels, distinct from the last, which is transcribed as long (e.g. /Qal'aska:/, where Q = glottal stop). As was noted above, a study of neutralization patterns in German indicates a restriction to peripheral vowels (or schwa) in open syllables, in contrast to centralized vowels in closed syllables (i.e. /ɑ.'las.kɑ/). It goes without saying that proper annotations are a crucial prerequisite for meaningful phonological studies. (The reference to the *Kiel Corpus* in our measurements of the low vowels is restricted to stressed syllables for this reason.)

To establish markedness constraints, we consult typological databases such as UPSID (cf. section 2). At close sight, the results of such studies often raise questions. Consider again the case of markedness involving lip roundedness, which in fact involves two parameters, vertical lip compression and lip protrusion (cf. Ladefoged and Maddieson 1996: 295). These are implemented jointly in most languages, but what is the claim for each individual parameter? Worse problems arise with respect to the sort of phonetic length and quality differences observed in German. Basing a typological study on the results presented by Kohler (1992) or Eckert & Barry (2005), compared to the results proposed here, will greatly affect the outcome of typological work.[18] If for a relatively well-studied language like German there is so little consensus of how to present the basic vowel system then how does this bode for hundreds of less studied languages? Again, the central issue here is abstractness: comparisons are valid only if all studies subsumed in typological surveys conform to specific methods for conducting phonological analyses.

To study neutralization patterns, we use electronically searchable word lists including the CELEX databases for German and English (cf. Baayen et al. 1995) and pronunciation dictionaries (e.g. Wells (2000) for English, Krech et al. (2009) and Dudenband 6 (2015) for German). The CELEX databases have the advantage that they are searchable with regular expressions, allowing for the extraction of word lists matching specific patterns. These databases are useful for finding examples or getting a first impression concerning certain patterns. Their disadvantage is that they are far too small (ca. 50,000 entries for German CELEX), include no information on variation, and tend to exclude precisely the most valuable "marginal" words discussed above.

The pronunciation dictionaries are much more comprehensive (for instance roughly 150,000 entries in Krech et al. (2009)) and also include some useful information regarding variation (especially Wells (2000) and Dudenband 6 (2015)). However, they, too, contain relatively little information on the "marginal" words, especially acronyms. For foreign proper nouns they often list only the entirely unassimilated pronunciation pertaining to the source language (e.g. [prɔˈvɑ̃ːs] 'Provence' in Dudenband 6).[19] Electronic searches are tedious, as only specific grapheme strings can be submitted in search queries.

As was noted above, for the time being the perhaps most valuable data to establish constraint interaction consist of loan words and acronyms, even speech

18 Cf. Becker-Kristal (2010: 7ff) for an overview of different perspectives on vowel length in typological surveys.

19 These omissions are understandable given the main purpose of these dictionaries to provide information on the "correct" pronunciation, not least to meet their users demand to avoid possible social stigma.

errors, all of which involve a relation among an output and a given input form. Comparisons of these forms allow for systematic modifications of sound structure to be established, thereby providing a window on active constraints. The relevant adaptation patterns typically involve discrete decisions. Is French [bis'tʀo] 'bistro' borrowed into German by imitating peripheral [i] or by replacing it with a centralized [ɪ] (/bis'tʀo/ or /bɪs'tʀo/), by imitating the final stress or by shifting it to the initial syllable (/bɪs'tʀo/ or /'bɪstʀo/)? Does the pronunciation of the acronym GAL rhyme with /bal/ <Ball> 'ball' or with /vɑl/ <Wal> 'whale'? Discrete decisions of this type lend themselves to documentation in the form of transcriptions, as the choices of symbols can be assumed to be fairly reliable.[20] Unfortunately, the relevant data are nonetheless difficult to obtain, as even specialized dictionaries of abbreviations and acronyms (e.g. Steinhauer 2005), give no information regarding the pronunciation. As far as we know, there are currently no corpora for speech errors with phonological transcriptions or organized around phonological questions.[21]

The most valuable data to shed light on correspondence constraints also concern pairs of words, that is rhymes and paradigmatic alternations. Again, these data are often hard to find and, like loan word adaptation patterns and acronyms, ought to be backed up by experimental studies. The relevant collections will always pale in size compared to regular speech corpora but are likely to yield valuable insight into phonological systems. It is unclear how a strictly corpus-driven approach based on "raw speech" corpora alone could achieve this. In fact, the wider issue emerging from the above discussion of the problematic annotations in the *Kiel Corpus* is that proper annotation presupposes a thorough phonological analysis, to yield classifications of sounds which can be meaningfully compared.

References

Baayen, R. Harald, Richard Piepenbrock and Leon Gulikers. 1995. *CELEX2 LDC96L14*. Web Download. Philadelphia: Linguistic Data Consortium.

Becker, Thomas. 1998. *Das Vokalsystem der deutschen Standardsprache.* Frankfurt et al.: Peter Lang.

20 The outright rejection of transcriptions as "second-hand data", on the basis that it is impossible to verify their validity (Delais-Roussarie & Yoo 2014: 203), seems more justified when concerning subtle phonetic detail.

21 Here recordings would in fact be valuable to investigate the claim that only phonemic structure is affected by speech errors (Stampe 1973).

Becker-Kristal, Roy. 2010. *Acoustic typology of vowel inventories and Dispersion theory: Insights from a large cross-linguistic corpus*. Ph. D. Dissertation, Los Angeles: UCLA.

Beckman, Jill. 1998. *Positional Faithfulness*. Ph. D. dissertation. Amherst: UMass.

Boersma, Paul and David Weenink. 2016. *Praat: doing phonetics by computer* [Computer program]. Version 6.0.17 "http://www.praat.org/" (21 April 2016).

Brentano, Clemens. 1852. Romanzen vom Rosenkranz. Wikisource website. https://de.wikisource.org/wiki/Romanzen_vom_Rosenkranz (July 2017).

Brinckmann, Caren, Stefan Kleiner, Ralf Knöbl and Nina Berend. 2008. German Today: an areally extensive corpus of spoken Standard German. In *Proceedings 6th International Conference on Language Resources and Evaluation* (LREC 2008), Marrakesch, Marokko.

Chambers, John M., William S. Cleveland, Beat Kleiner and Paul A. Tukey. 1983. *Graphical Methods for Data Analysis.* Belmont, CA: Wadsworth & Brooks/Cole.

Delais-Roussarie, Elisabeth and Hiyon Yoo. 2014. Corpus and Research in Phonetics and Phonology. In Jacques Durand, Ulrike Gut and Gjert Kristoffersen. *The Oxford Handbook of Corpus Phonology*, 193–213. Oxford: Oxford University Press.

[Dudenband 6] = Kleiner, Stefan, Ralf Knöbl and Max Mangold. 2015. *Duden – Band 6: Das Aussprachewörterbuch.* Berlin: Dudenverlag & Mannheim: IDS.

Eckert, Hartwig and William Barry. 2005. *The Phonetics and Phonology of English Pronunciation.* Trier: Wissenschaftlicher Verlag Trier.

Hixon, Thomas J., Gary Weismer and Jeanette D. Hoit. 2008. *Preclinical speech science: Anatomy, physiology, acoustics, perception.* San Diego, CA: Plural Publishing Inc.

Kohler, Klaus J. (ed.). 1992. Phonetisch-Akustische Datenbasis des Hochdeutschen. Kieler Arbeiten zu den PHONDAT-Projekten 1989–1992. *Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel (AIPUK),* 26.

Kohler, Klaus J. 1994. Lexica of the Kiel PHONDAT Corpus Read Speech. Volume I. *Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel (AIPUK),* 27.

Kohler, Klaus J. 1999. German. In: *Handbook of the International Phonetic Association.* Cambridge: Cambridge University Press.

Krech, Eva-Maria, Eberhard Stock, Ursula Hirschfeld and Lutz Christian Anders. 2009. *Deutsches Aussprachewörterbuch.* New York: De Gruyter.

Ladefoged, Peter and Ian Maddieson. 1996. *The sounds of the World's languages.* Oxford: Blackwell.

Leuninger, Helen. 1996. *Reden ist Schweigen, Silber ist Gold.* München: Deutscher Taschenbuch Verlag.

Lindau, Mona. 1978. Vowel Features. *Language* 54: 541–563.

Maddieson, Ian. 1984. *Patterns of sounds.* Cambridge: Cambridge University Press.

Maddieson, Ian and Kristin Precoda. 1990. Updating UPSID. *UCLA-Working Papers in Phonetics* 74: 104–114.

Peterson, Gordon and Harold Barney. 1952. Control Methods used in a study of the vowels. *Journal of the Acoustical Society of America* 24: 174–185.

Prince, Alan and Paul Smolensky. 1993. Optimality Theory: Constraint Interaction in Generative Grammar. *Rutgers University Center for Cognitive Science Technical Report* 2, New Brunswick: Rutgers University Press.

Raffelsiefen, Renate. 2016. Allomorphy and the question of abstractness. Evidence from German. *Morphology* 26, Issue 3.

Ramers, Karl Heinz. 1988. *Vokalquantität und –qualität im Deutschen.* Tübingen: Niemeyer.

Stampe, David. 1973. *A Dissertation in Natural Phonology.* Ph.D. dissertation, University of Chicago, published 1979 by Garland Press, New York.

Steinhauer, Anja. 2005. *Duden. Das Wörterbuch der Abkürzungen.* 5th revised ed. Mannheim et al.: Dudenverlag.

Vennemann, Theo. 1991. Skizze der deutschen Wortprosodie. *Zeitschrift für Sprachwissenschaft* 10: 86–111.

Wängler, Hans-Heinrich. 1964. *Atlas deutscher Sprachlaute.* 3rd revised ed. Berlin: Akademie Verlag.

Wells, John. 2000. *Longman Pronunciation Dictionary.* 2nd edition, Harlow: Pearson Education Limited.

Wesker, Thorsten, Bernd Meyer, Kirsten Wagener, Jörn Anemüller, Alfred Mertins and Birger Kollmeier. 2005. *Oldenburg Logatome Speech Corpus (OLLO) for Speech Recognition Experiments with Humans and Machines*, Proc. of Interspeech, 4–8. Lisbon, Portugal.