languages

MDPI

Article

Correspondence of Consonant Clustering with Particular Vowels in German Dialects

Samantha Link

Deutscher Sprachatlas, Philipps-Universität Marburg, 35037 Marburg, Germany; linksa@staff.uni-marburg.de

Abstract: Recent work found a correspondence between consonant clustering probability in monosyllabic lexemes and the three vowel types, short and long monophthong and diphthong, in German dialects. Furthermore, that correspondence was found to be bound to a North–South divide. This paper explores the preferences in consonant clustering of particular vowels by analyzing the PhonD2-Corpus, a large database of phonotactic and morphological information. The clustering probability of the diphthongs is positively correlated with frequency while the other vowels showed particular preferences that are not positively correlated with frequency. However, all of them are determined by a threefold pattern: short monophthongs prefer coda clusters, diphthongs onset clusters and long monophthong are balanced. Furthermore, it was found that this threefold pattern seems to have evolved from an originally twofold pattern (short monophthong prefers coda clusters and long monophthong and diphthong prefer onset clusters) in Middle High and Low German. This result is then further considered under the aspect of the compensation of the syllable weight and moraicity. Furthermore, some interesting parallels with the syllable vs. word-language typology framework are noted.

Keywords: phonotactics; German dialects; Middle High German; vowels; consonant clustering

check for updates

Citation: Link, Samantha. 2024. Correspondence of Consonant Clustering with Particular Vowels in German Dialects. *Languages* 9: 255. https://doi.org/10.3390/languages9070255

Academic Editors: Adrian P. Simpson and Jeanine Treffers-Daller

Received: 8 March 2024
Revised: 8 July 2024
Accepted: 17 July 2024
Published: 22 July 2024

1. Introduction

Consonant clusters are linear sequences of consonantal units occurring before or after the nucleus of a syllable and can be defined on the level of graphemes, phones or phonemes (Gregová 2010, p. 79). The composition of such a sequence is firstly determined by the Sonority Sequencing Principle (Selkirk 1984, p. 116), which is considered as a linguistic universal. However, there are also language-specific phonotactic rules: such rules might forbid certain consonant combinations, but they can also allow violations of the Sonority Hierarchy (Alber and Meneguzzo 2016, p. 27).

That particular phonotactic rules and preferences do not only apply to languages but also to dialects is shown by Lameli (2022, pp. 262–66) in describing a North–South divide of consonant clustering probability in monosyllabic words in the dialects spoken in the Federal Republic of Germany. From the Low German to the Upper German dialects, the consonant clustering probability was found to increase. One example of how this increase manifests in the data is the way the lemma *gesagt*, 'said', is realized in the examples (1)–(6) from different areas.

(1) zɛçt (Hohwacht, North Saxon, LG);
(2) əzɛçt (Astfeld, Eastphalian, LG);
(3) jəzaːt (Siebenbach, Mozelle Franconian, WCG);
(4) gəzaːxt (Linz, Thuringian, ECG);
(5) gsaIt (Bempflingen, Swabian, WUG);
(6) gsɔkt (Peterskirchen, Central Bavarian, EUG).

As illustrated by these examples[1], there are not only differences in syllabicity (mono- vs. bisyllabic) and consonant clustering (no clusters, onset and/or coda clusters) but also in whether they are realized with a short or long monophthong or a diphthong.

Thus, the question was raised in Lameli and Link (forthcoming) how the spatial North–South increase interacts with these three vowel types (V, Vː, VV). The results from Lameli and Link (forthcoming) given in Figure 1 reveal that although the North–South increase is present in all vowel types, they differ in terms of the preference of the cluster position: the short monophthongs prefer coda clusters, which contrasts with the onset preference of the diphthongs, and the long monophthongs have a more balanced pattern with a slight onset preference, except in East Upper German (EUG), which prefers coda clusters.

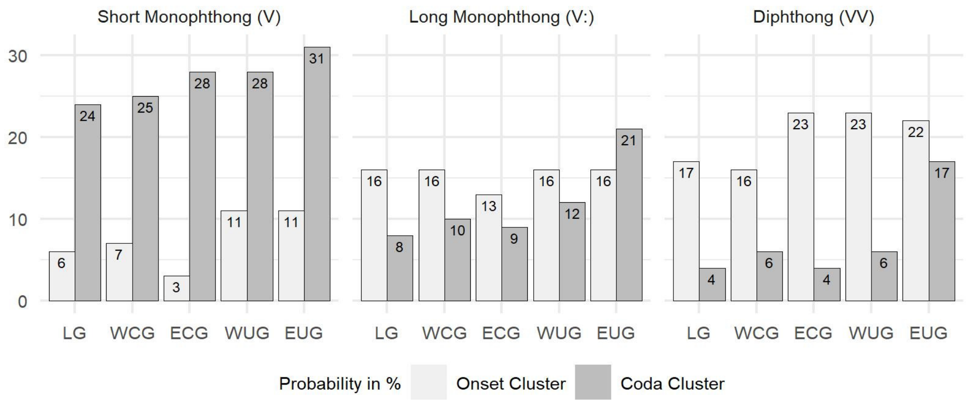

Figure 1. Probability of onset (lighter gray) and coda (darker gray) consonant clustering for different vowel types (V, Vː, VV) in dialects of Low German (LG), West Central German (WCG), East Central German (ECG), West Upper German (WUG) and East Upper German (EUG).

One question the findings from Lameli and Link (forthcoming) and Lameli (2022, pp. 262–66) bring up is which syllable types are actually responsible for shaping these particular clustering probabilities. This will be addressed in an analysis of the CV structure in Section 3.1.

1 Which syllable structures (CV) shape the pattern found by Lameli and Link (forthcoming)?

A second question, which is the main focus of this study, is which actual vowels are represented by the three groups V, V: and VV? Standard German is classified as a language with a large vowel system (Maddieson 2013). From its 15 vowels, there are 6 tense ones, which can be short [i y e ø u o] or long [iː yː eː øː uː oː], and 9 short lax ones [ɪ ɛ ʏ œ ʊ ɔ a ə ɐ] from which 2 can be lengthened [ɛː aː] (Hall 2000, p. 34). Furthermore, there are the three diphthongs [aɪ aʊ ɔɪ]. The Standard German vowel inventory is also shared by the dialects. However, they introduce a couple of particularities: [ɑ] or [ɒ] are found in many areas including Hessian (Durrell and Davies 1990, p. 225), Thuringian (Spangenberg 1990, pp. 269, 274, 277), High Alemannic (Russ 1990a, p. 369) and Southern Bavarian (Wiesinger 1990, p. 485). Furthermore, there are nasal vowels in Swabian (Russ 1990b, p. 347) and tone accents in Central Franconian (Schmidt 1986). A detailed overview of the vowel inventories of the German dialects is given in Wiesinger (1983b) and further descriptions can be found in the two edited volumes by Herrgen and Schmidt (2019) and Russ (1990c). But so far, there are no studies providing any information on German dialectal vowel frequencies. In contrast, there are a couple of studies discussing them for Standard German: a very early approach is Trubetskoy's (1967, p. 233), who calculated the vowel percentages of 200 words from a scientific text (K. Bühler) and from a fairytale (A. Dirr) which are given in Table 1. Due to these rather similar percentages, he concludes that German phoneme frequency is not influenced by stylistics.

Menzerath (1954, p. 72f) provides frequency counts for monosyllabic words that constitute the following order (from highest to lowest count). [ʌ] does not belong to the Standard German vowel inventory and might originate from a loanword in his data. The diphthong /ui/ is restricted to cases such as interjections (*Pfui!*) or names (*Luise*).

Short Vowels: a (385), i (252), u (214), ä (197), o (191), ü (15), ö (15), æ (3), ʌ (1);

Long Vowels: a: (164), o: (108), i: (107), e: (93), u: (92), ä: (31), ü: (19), ö: (26);

Diphthongs: ei (163), au (110), eu (31), ui (2).

Delattre (1964, p. 89) gives combined percentages for mono- and polysyllabic words, but does not further distinguish between long and short vowels:

ə (23.88%), ɪ (11.52%), a (10.72%), aɪ (7.88%), ɛ (7.08%), e (6.73%), i (5.88%),

ʊ (5.28%), ɑ (3.89%), u (3.84%), ɔ (3.74%), aʊ (3.14%), o (2.29%), y (1.30%),

ʏ (0.98%), ø (0.95%), ɔʏ (0.55%), œ (0.35%).

Pätzold and Simpson (1997, p. 223) present relative vowel frequency counts from the Kiel Corpus resulting in the order given below (the original data are given as a bar plot; thus, the percentages had to be extracted from the y-axis):

ə (15,6%), ɪ (11.7%), a (9%), i: (8%), aɪ (7.8%), ɛ (7%), ɐ (7%), e: (6.7%), a: (5.3%),

ʊ (5%), u: (3.3)%, ɔ (3%), aʊ (2.9%), o: (2.8%), y: (1.3%), ɔʏ (1.1%), ʏ (1%), ø: (1%),

œ (0.5%).

Since these four works rely on different data and analyze them in varying granularity, it is not surprising that their results for short and long vowels do not agree in every aspect. Pätzold and Simpson (1997, p. 222) further point out that vowel frequencies tend to be a speaker-dependent phenomenon. Nevertheless, all four works reveal a general tendency that can be summarized as the following: [ə] is the most frequent vowel in German, followed by the group of unrounded non-back vowels [a, ɪ, i, ɛ, e] with [ɪ] as the most frequent one, followed by the back vowels [u, ʊ, o, ɔ] and the rounded front vowels [y, ʏ, ø, œ] as the rarest group. However, this pattern contradicts the hypothesis of Menzerath (1954, p. 72f) that the most frequent vowels are those found in the extreme points of the vowel space. In contrast to the monophthongs, the results for the diphthongs are consistent: all four sources found them to be ordered as [aɪ, aʊ, ɔʏ].

Table 1. Vowel percentages from Trubetskoy (1967, p. 233).

	a	ə	u	o	au	i	e, ä	ei	ü	ö	öü
Bühler	15%	22%	7%	10%	3%	17%	18%	4%	3%	0.5%	0.5%
Dirr	18%	22%	9%	10%	4%	16.5%	11%	7.5%	1%	1%	0%

In Section 3.2, this work presents a comparative analysis of vowel frequency in the five German dialect areas. Coming back to the initial question of what stands behind the three groups V, V: and VV, Section 3.3 features an analysis of consonant clustering similar to Lameli and Link (forthcoming) but separately for all vowels and diphthongs. Based on the results from Sections 3.2 and 3.3, two questions are considered:

2 Are some vowels more inclined to have consonant clusters than others?
3 Do more frequent vowels also have a higher probability for clustering?

Another aspect from Lameli and Link (forthcoming) is the important role of sound change processes from both Middle High German (MHG) and Middle Low German (MLG) to the modern German dialects. MHG is the historic stage of German, which was spoken in a couple of varieties in the Central and Upper German area around 1050 to 1350 (Paul 2007, p. 10). MHG is thus a diachronic reference point for the vowel systems of the modern dialects of this area (Wiesinger 1983b, p. 1044). The Low German dialects correspond to MLG (Wiesinger 1983b, p. 1045), which is dated from the start of the 13th century to the 16th century (Lasch 1974, pp. 3, 5). Like modern German, MHG and MLG have short and long monophthongs and diphthongs, but due to a couple of sound change processes, there are some major differences: as an example, there is the New High German (NHG) Diphthongization, which caused MHD <î>, <iu> and <û> to become [aɪ], [ɔʏ] and [aʊ].

In order to approach the question below, Section 3.4 presents another analysis of consonant clustering probability based on splitting the data according to the underlying MHG/MLG vowels.

4 Can differences in clustering be traced back to MHG/MLG vowels?

Furthermore, Section 3.4 also considers the role of the (CV) syllable structure and takes a look at the actual MHG/MLG words in terms of an analysis of hierarchical clustering.

2. Materials and Methods

The PhonD2-Corpus is an open access online database on phonotactic and morphological structure of the dialects in the Federal Republic of Germany (see Lameli et al. 2023). The corpus features translations of Wenker sentences[2] into dialect by 172 subjects from 172 sites all across Germany (approx. 80,000 words and approx. 300,000 sounds).

The audio data used for the corpus have been taken from the project *Phonetischer Atlas der Bundesrepublik Deutschland* (PAD) (Göschel 1992, 2000). The PhonD2-Corpus contains broad phonetic IPA and SAMPA transcriptions, which had the aim of documenting features that are phonologically contrastive on the dialectal level[3]. The transcribed words are syllabified based on the Sonority Sequencing Principle (Selkirk 1984, p. 116) and the Maximum Onset Principle (Hall 2000, p. 217). Since the syllabification also considers syllabic consonants (liquids and nasals; see Hall 2000, p. 216), consonant sequences caused by the loss of schwa such as [bn] in [ha:bn] 'to have' are treated as syllables (with [b] as onset and [n] as nucleus). The corpus provides further phonotactic information such as the CV structure, syllable scheme, sonority measures, syllable count, syllable weight and syllable type and also gives some morphological data. It is available under https://dsa.info/PhonD2/ (last accessed on 17 July 2024), where the data for each lemma can be downloaded as a CSV or EXCEL file.

All data analysis was performed with R (v4.1.3; R Core Team 2022) and all the bar plots were made with the package *ggplot2* (Wickham 2016).

In recent phonotactic research, it is common to classify consonant clusters into two groups (see Dressler and Dziubalska-Kolaczyk 2006): There are morphonotactic clusters, as in va:ʃt (warst), 'you were', which result from the affixation of consonantal suffixes *-s*, *-t*, *-st* or the *g*-prefix in the Upper German area due to the schwa syncope of the past participle marker *ge-* like in *ghapt* (gehabt), 'have', (see Rowley 1997, p. 76). Clusters within a word consisting of a single morpheme such as vuɐʃt (Wurst), 'sausage', are referred to as phonotactic. In Lameli and Link (forthcoming), it was revealed that morphologically complex monosyllabic lexemes have a higher clustering probability than those consisting of a single morpheme, which is in line with Dressler and Kononenko (2021, p. 43). Since this was the case in all five dialect areas and for all three vowel types alike, the morphonotactic vs. phonotactic distinction was not further considered for the spatial analyses of Lameli and Link (forthcoming; this result is given in Appendix A in Figure A1).

This study is based on the same data as Lameli and Link (forthcoming): monosyllabic lexemes from the Wenker sentences of the PhonD2-Corpus. But to ensure consistency, its focus is set on purely phonotactic clusters and thus only monosyllabic lexemes without morpheme boundaries (85.29% of the original data) are considered for analysis. Nevertheless, all analyses were also made using the original data and these results are given in Appendix A for comparison.

As in Lameli and Link (forthcoming), the data were split into the five large dialect areas according to Wiesinger (1983a): Low German (LG), West Central German (WCG), East Central German (ECG), West Upper German (WUG) and East Upper German (EUG)[4]. Below, some frequency counts of the analyzed data are given:

Total count of monosyllabic lexemes without morpheme boundaries: 23.233.

Counts per area: **LG:** 8812; **WCG:** 4966; **ECG:** 2511; **WUG:** 2193; **EUG:** 4751.

Frequency counts of CV structure:

LG: CVC (2406); CV: (1091); CV:C (1076); CVCC (907); CVV (683); CVVC (531); VC (501); CCV:C (384); CCVV (216); V:C (190); CCV: (185); CCVVC (134); CCVC (132); VVC (88); CV:CC (66); CV (50); CVCCC (29); V:CC (29); CCVCC (22); CVVCC (21); V (20); VV (20); VCC (12); V: (7); VVCC (5); CCV:CC (3); CCVVCC (2); CCCV:C (1); CVCCCC (1).

WCG: CVC (1085); CV: (631); CVCC (586); CVV (491); CVVC (441); CV:C (405); VC (287); CCV:C (151); CCVVC (128); CCV: (127); CCVV (117); CCVC (110); VVC (64); CV:CC (56); VCC (47); CVVCC (41); V:C (40); CV (33); CCVCC (30); VV (22); V (20); V: (19); CVCCC (9); CCCVC (8); VVCC (7); CVCCCC (4); V:CC (3); CCV:CC (2); CCCV:C (1); CCVVCC (1).

ECG: CVC (612); CV: (364); CVCC (363); CV:C (266); CVV (150); VC (149); CVVC (131); CCVV (80); CCV:C (79); CCVVC (60); CCV: (59); CCVC (32); VVC (29); CV (27); CV:CC (19); V:CC (18); V: (13); V:C (13); VCC (12); V (9); VVCC (7); VV (6); CVCCC (4); CCVCC (2); CVVCC (2); VCCC (2); CCCVC (1); CV:CCC (1); V:CCC (1).

WUG: CVC (408); CV: (313); CVCC (271); CVVC (230); CV:C (163); VC (142); CVV (129); CCVV (92); CCV: (57); CCVVC (56); CCV:C (52); VV (50); CCVC (41); CCVCC (35); CV:CC (27); VCC (22); VVC (20); V (19); CVVCC (14); CVCCC (13); V:C (11); V: (10); CV (9); CCV:CC (3); C (1); CCCV:C (1); CCCVCC (1); CV:CCC (1); V:CC (1); VVCC (1).

EUG: CVC (812); CVVC (679); CV: (559); CVCC (497); CVV (373); CV:C (298); VC (281); CCVVC (190); CCVV (144); CCV:C (125); CV:CC (90); CCV: (85); CVVCC (84); CCVCC (78); V: (75); CCVC (69); V:C (58); CV (57); VVC (53); CVCCC (23); V (23); VV (20); VVCC (19); VCC (17); CCV:CC (15); V:CC (14); CCVVCC (5); CCCV:C (2); CCCVC (2); CVVCCC (2); CVCCCC (1); VCCC (1).

3. Results

3.1. CV Structure and the North–South Increase in Clustering Probability

The above counts of the CV syllable structure reveal that large clusters with three or four consonants are among the rarer items in all areas. In order to test which syllable types were responsible for creating the North–South increase and the specific pattern of each of the three vowel groups (V, V:, VV), the percentage of bi- and triconsonantal clusters was calculated in Figure 2. As an example, (CC)V(CC) means that for the onset, all monosyllables with a CCV pattern were considered for calculation (CCV, CCVC, CCVCC, CCVCCC) and for the coda, the corresponding VCC pattern (VCC, CVCC, CCVCC, CCCVCC).

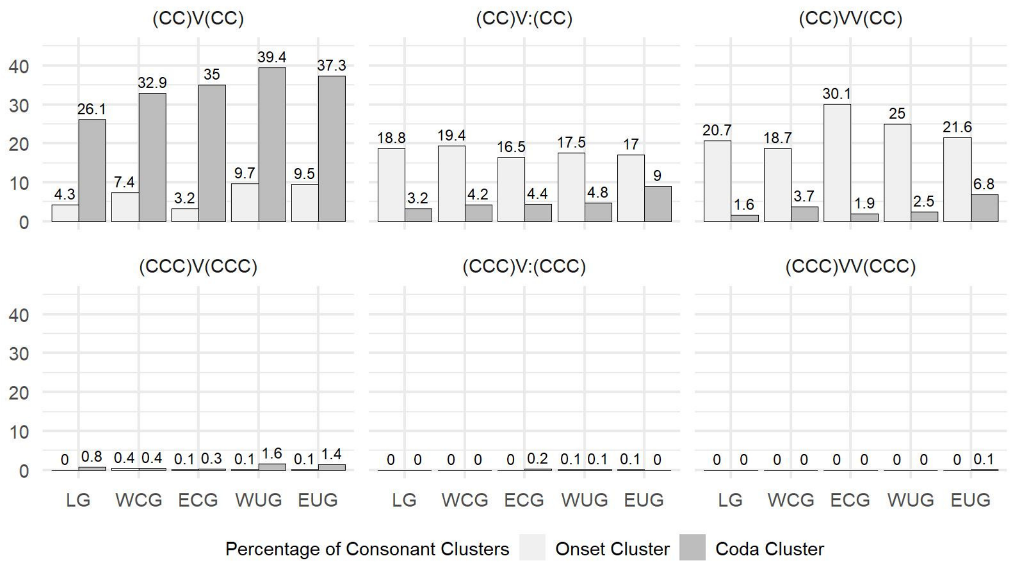

Figure 2. Percentage of onset (lighter gray) and coda (darker gray) CC or CCC consonant clusters for different vowel types (V, V:, VV) in the dialects of Low German (LG), West Central German (WCG), East Central German (ECG), West Upper German (WUG) and East Upper German (EUG).

The results given in Figure 2 illustrate that the biconsonantal clusters are mainly responsible for the spatial pattern and the particular preferences of the vowel groups. Nevertheless, both phenomena are also (faintly) visible for the triconsonantal clusters.

3.2. Vowel Percentages

As previously mentioned, the main focus of this work is put on the vowels behind the three vowel groups, V, V: and VV[5]. Since no comparative dialectal vowel frequency counts are available in the literature, the Standard German ones presented in the introduction are taken for comparison when discussing the analyses of the PhonD2 data.

The percentages of the vowels are presented separately for the three groups (V, V:, VV). For a better comparison, only vowels found in all areas were considered for calculation. To properly display the data, all values below a threshold of 3% were taken together as a single group "Other". The complete vowel frequency counts of each area can be accessed in Appendix A (Tables A1–A3). Summing up the general tendencies of Standard German vowel frequencies from the introduction, [ə] is the most frequent followed by unrounded non-back vowels [a, ɪ, i, ɛ, e], back vowels [u, ʊ, o, ɔ] and rounded front vowels [y, ʏ, ø, œ].

3.2.1. Short Vowels

Figure 3 reveals the same general pattern as in Standard German: [ɪ], [a] and [ɛ] are among the most frequent vowels followed by [ʊ] and [ɔ] and the rounded front vowels are only in Low German above the 3% threshold. In addition to this, there are some dialect-specific preferences such as East Upper German having increased percentages of [ɑ] while having the lowest percentage of [a]. This is caused by the *a-Verdumpfung*, a sound change process that transformed MHG <a> to [ɔ] or [ɑ] like in (7) and (8) below (see Koch 2019, p. 283; Wiesinger 1990, p. 450; Streck 2019, p. 214)[6].

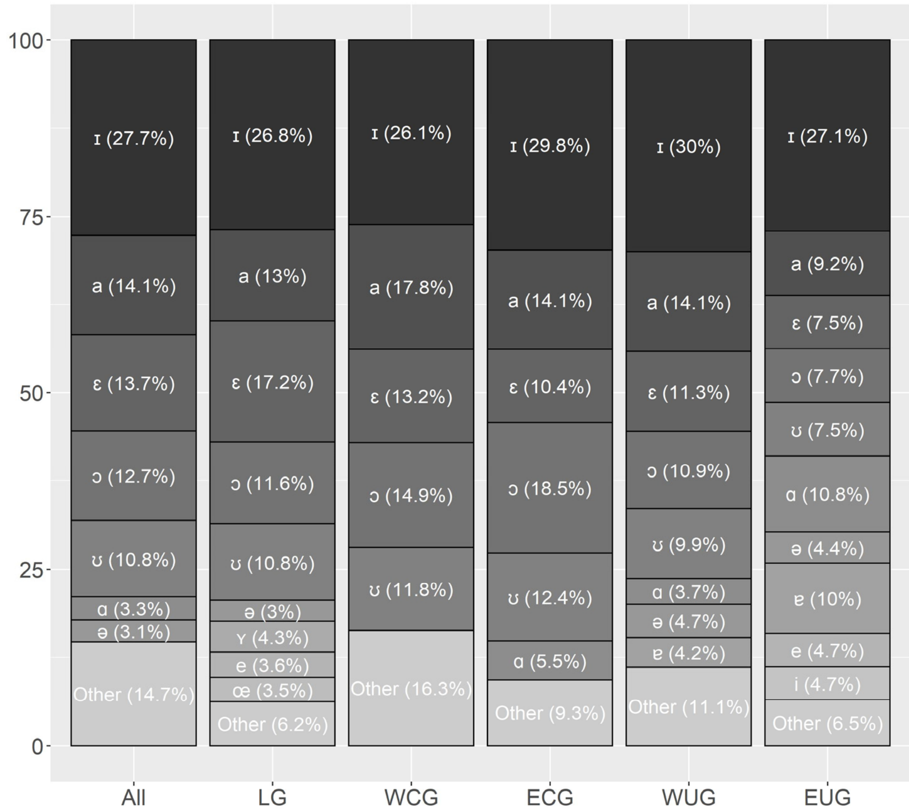

Figure 3. Percentage of short vowels in monosyllabic lexemes without morpheme boundaries for all areas (All), Low German (LG), West Central German (WCG), East Central German (ECG), West Upper German (WUG) and East Upper German (EUG).

(7) nɑxt (Nacht, 'night', Gelting, Central Bavarian);
(8) gʀɑŋk (krank, 'sick', Schachach, Central Bavarian).

Furthermore, East Upper German has the lowest percentage of [ʊ]. This can be attributed to MHG <uo>, as in *muoȝ* (NHG *muss*), which is realized with a diphthong instead of a short vowel (see Koch 2019, p. 285; and Wiesinger 1990, p. 447).

(9) ɛɐ muɐs (er muss, 'he has to', Oberviehbach, Central Bavarian);
(10) ɛɐ moʊs (er muss, 'he has to', Pielenhofen, North Bavarian).

3.2.2. Long Vowels

Figure 4 shows a clear tendency for [o:] and [a:] to be the most frequent long vowels, while [u:], [ɛ:] and [y:] are less frequent. This stands in contrast to the previously raised assumption that unrounded non-back vowels are more frequent than back vowels. East Central and East Upper German have a much higher percentage of [ɑ:] than the other areas. The Bavarian *a-Verdumpfung* has already been described in Section 3.2.1. MHG diphthongs <ei>, <ou> and <öu> are realized as [ɑ:] in Erzgebirgian and Vogtlandian dialect and MHG <â> became [ɑ:] in Eastern Thuringian and North Upper Saxon (Siebenhaar 2019, pp. 419 and 420); see examples (11)–(13) below.

(11) glɑ:p (glaube, '(I) believe', Borstendorf/Erzgebirge, Upper Saxon);
(12) kɑ:m (kamen, '(they) came', Großballhausen, Thuringian);
(13) ɑ:mt (Abend, 'evening', Altenburg/Erzgebirge, Upper Saxon).

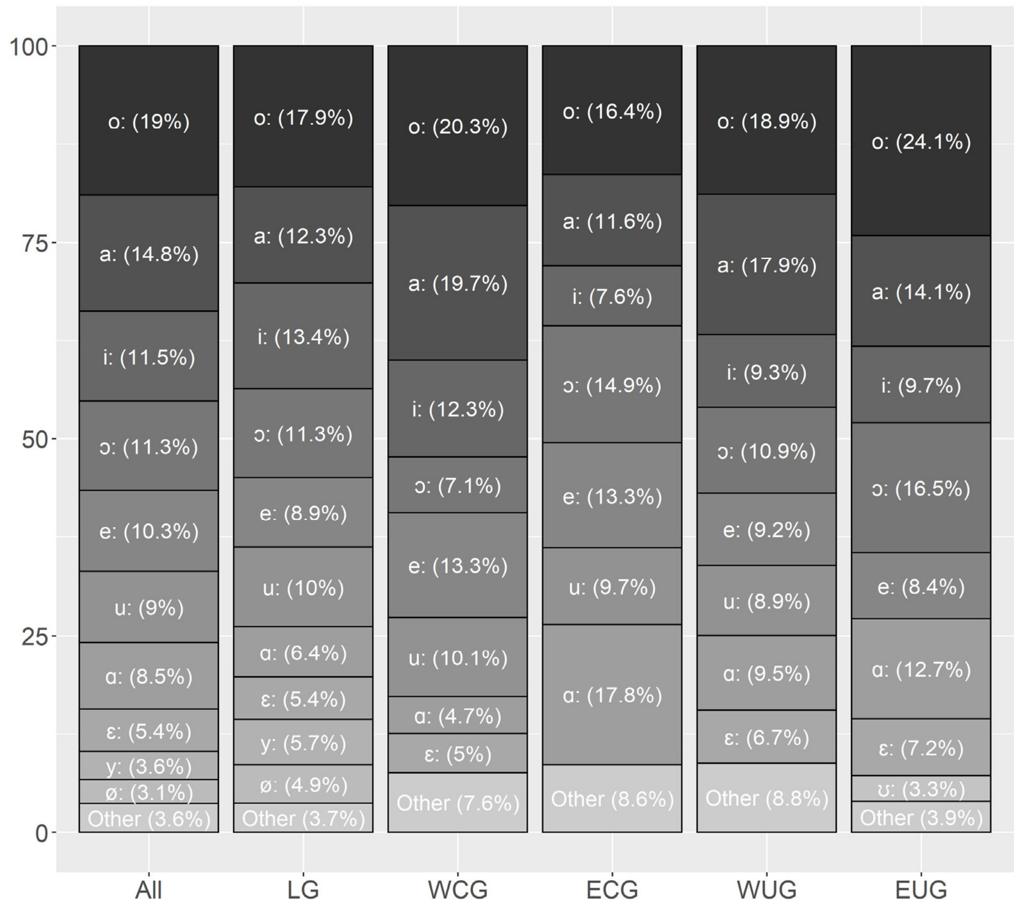

Figure 4. Percentage of long vowels in monosyllabic lexemes without morpheme boundaries for all areas (All), Low German (LG), West Central German (WCG), East Central German (ECG), West Upper German (WUG) and East Upper German (EUG).

Low German and West Central German have the highest percentages of [iː]. For Low German, this traces back to MLG <î>, which did not undergo New High German Diphthongization like in the examples (14)–(16) below (Elmentaler and Voeste 2019, p. 71).

(14) tiːt (Zeit 'time', Adorf, Westphalian);
(15) gliːk (gleich, 'in a monent', Hohnwacht, North Saxon);
(16) bliːf (bleibe, 'Stay!', Osterhagen, Eastphalian).

In contrast, the high percentage of [iː] in West Central German is related to the Central German Monophthongization of MHG <ie> (Paul 2007, p. 78).

3.2.3. Diphthongs

The percentages of diphthongs in Figure 5 reveal that as in Standard German, [aɪ] is the most frequent diphthong in all five dialect areas. [aʊ] takes the second place, except in East Upper German where it has a very low percentage. This is caused by its realization as [aɔ] or as [aː] before labial sounds in MHG <ou> as in (17)–(19) below (Wiesinger 1990, p. 452; Streck 2019, p. 214).

(17) haɔs (Haus, 'house', Maxweiler, Central Bavarian);
(18) glaːp (glaube, '(I) believe', Gelting, Central Bavarian);
(19) baːm (Baum, 'tree', Pöttmes, Central Bavarian).

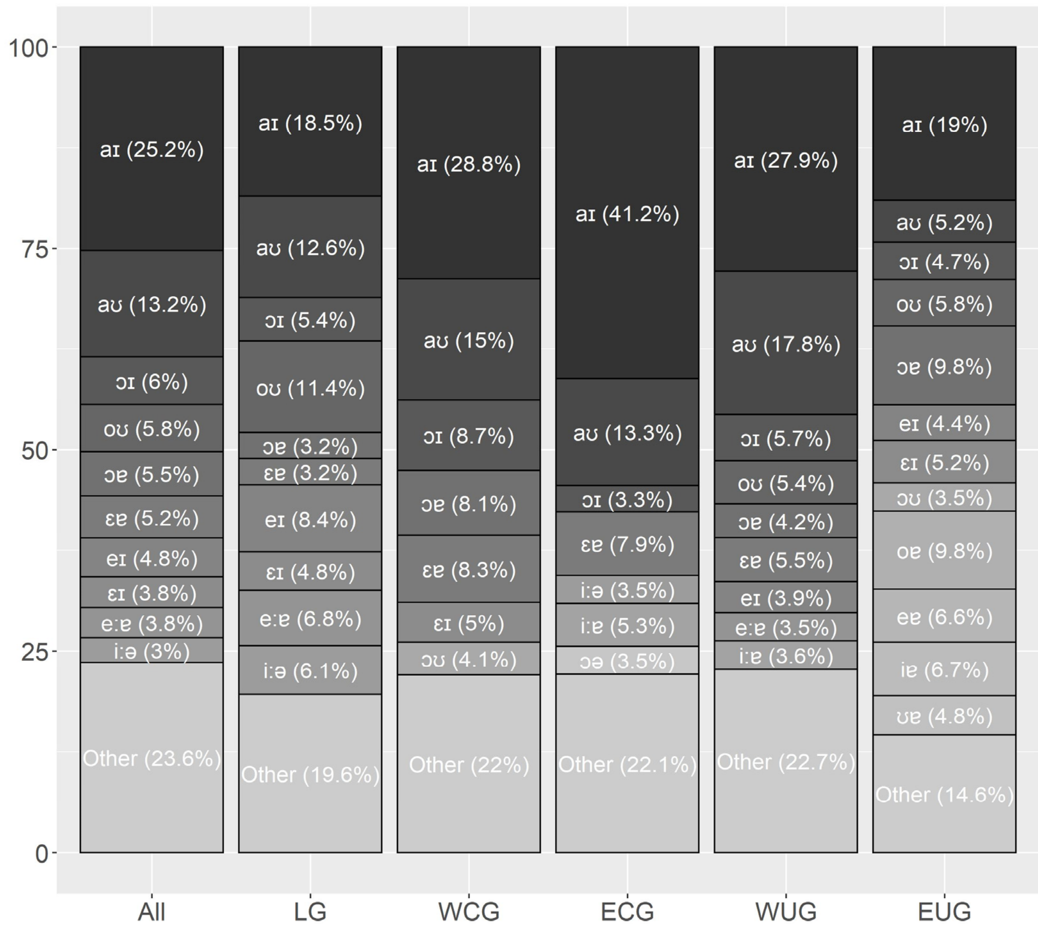

Figure 5. Percentage of diphthongs in monosyllabic lexemes without morpheme boundaries for all areas (All), Low German (LG), West Central German (WCG), East Central German (ECG), West Upper German (WUG) and East Upper German (EUG).

There is also a much larger percentage of diphthongs featuring [ɐ] in East Upper German. This comes from MHG <ei>, which is realized as [ɔɐ] or [oɐ] in North Bavarian, as in the examples (20) and (21) (Wiesinger 1990, p. 450) and also corresponds to a lowered percentage of [aɪ].

(20) glɔɐt (Kleid, 'dress', Eichenhofen, North Bavarian);
(21) tsvɔɐ (zwei, 'two', Pöttmes, Central Bavarian).

MHG <ie> did not undergo Monophthongization in this area and is preserved as [ɪɐ] or [iɐ] as in (22) and MHG <uo> became [ʊɐ] or [uɐ] as in example (9) in Section 3.2.1 (Wiesinger 1990, p. 447).

(22) jɪɐts (jetzt, 'now', Ulbering, Central Bavarian).

Low German displays a higher percentage of [oʊ] than the other areas, which comes from the preservation of MLG <ou> like in (23) and (24) (see Goltz and Walker 1990, p. 41).

(23) fʁoʊ (Frau, 'woman', Hammah, North Saxon);
(24) gloʊf (glaube, '(I) believe', Holmkjer, North Saxon).

3.2.4. Summary

All three vowel types, short and long monophthong and diphthong, display different vowel percentages in the five dialect areas, although they mainly adhere to the pattern worked out for Standard German. All dialectal differences can be explained by underlying historical sound change processes that did not affect all areas alike.

3.3. Consonant Clustering and Particular Vowels

In the Introduction, the questions were raised whether particular vowels also display individual preferences for consonant clustering (Question 2) and if yes, it has to be ruled out that this is not just positively correlated with vowel frequency (Question 3). To address Question 2, the focus is put on vowels and diphthongs from the Standard German inventory since they are shared by all dialect areas. The two reduction vowels [ə] and [ɐ] were left out because in the data consisting of monosyllabic lexemes, they can be considered as weakened forms of [ɛ] due to fast speech.

The consonant clustering probabilities of the selected vowels and diphthongs were calculated in the same way as in Lameli (2022, pp. 262–66) and Lameli and Link (forthcoming): instead of visualizing clustering probabilities with Finite State Automata for each area and vowel, the results were reduced to the relevant information, namely onset and coda clustering probabilities, which are bundled as bar plots in order to make them easily comparable (otherwise, this work without Appendix A would have required a total of 160 automata).

The consonant clustering probabilities of the short vowels [a, ɛ, ɪ, ɔ, ʊ] and the corresponding long vowels [a:, ɛ:, i:, o:, u:] and [e:] and the diphthongs [aɪ], [aʊ] and [ɔɪ] are given in Figure 6.

3.3.1. Short and Long Vowels

In Figure 6, all vowels show the same pattern as described in Lameli and Link (forthcoming): short vowels clearly prefer coda clustering (dark gray bars) over onset clustering (light gray bars) while long vowels are rather balanced with a slight preference for onset clusters. And also, the increase from North to South observed by Lameli (2022, pp. 262–66) is visible; most clearly for [ɛ], [ɪ], [ʊ] and [e:].

If the clustering probability is correlated with vowel frequency (Question 3), the expectation for the short vowels would be that [ɪ] should have the highest probabilities in all areas, followed by [ɔ] in West and East Central German and [ɛ] in Low German and West and East Upper German. But in contrast, [ɪ] even has the lowest coda clustering probabilities of West Central German and the second lowest of East Central German. [ɔ] also tends to have comparably lower probabilities in all areas. [ɛ] has the lowest coda clustering probabilities of Low German, but the second highest in East

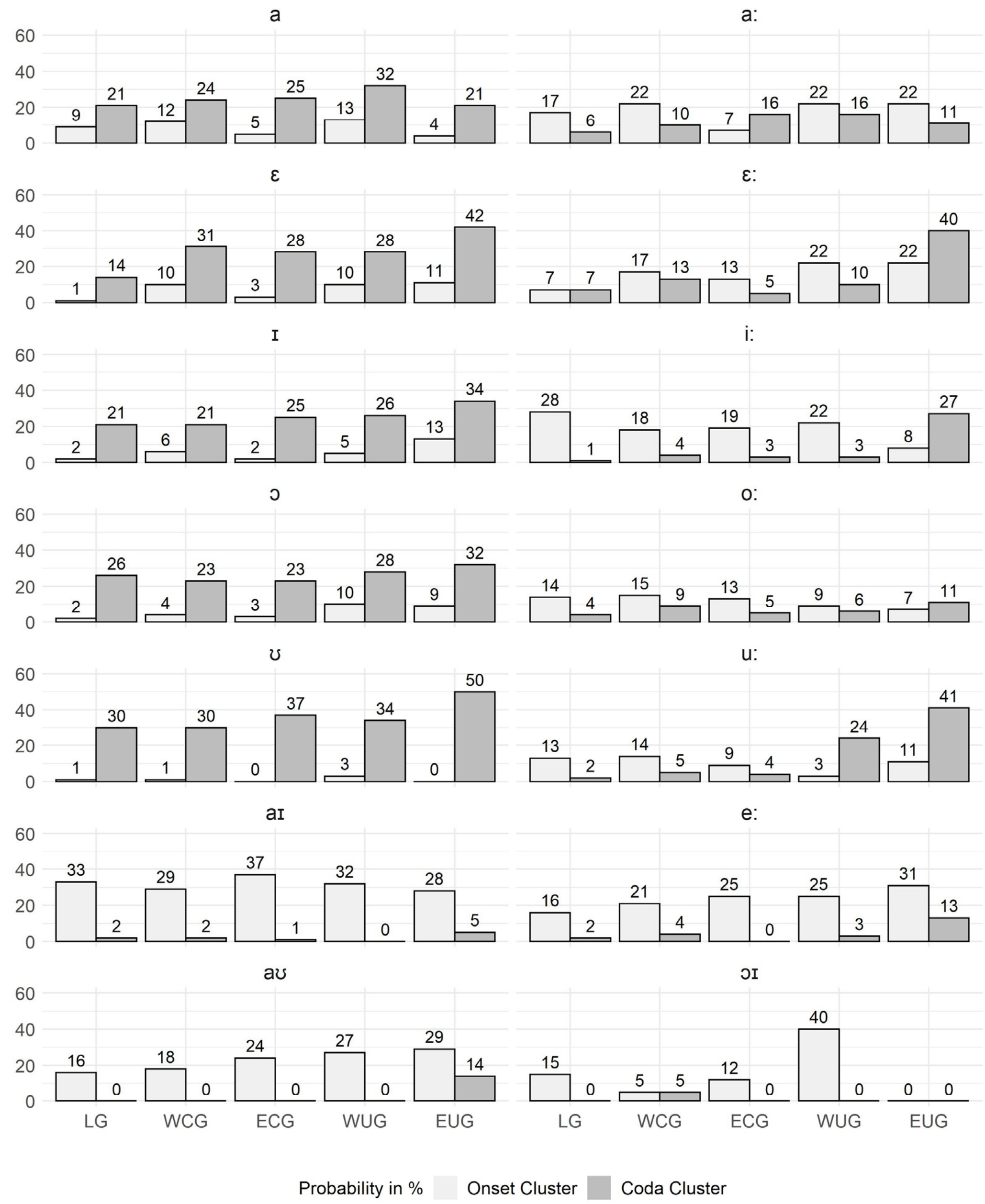

Figure 6. Consonant clustering probability for the vowels and diphthongs in the five dialect areas Low German (LG), West Central German (WCG), East Central German (ECG), West Upper German (WUG) and East Upper German (EUG).

Upper German [ʊ], which is a less frequent vowel, has the highest coda clustering probabilities in all areas (except a 1% difference in West Central German).

German [oː] is the most frequent long vowel, followed by [aː] in Low German and West Central and West Upper German and [ɑː] and [ɔː] in East Central and East Upper German. However, none of the areas has its highest clustering percentage among the long vowels (neither for onset nor for coda) in [oː]. In Low German, East Central German and West Upper German, [iː] is the long vowel with the highest onset clustering probabilities, while West Central German prefers [aː] and East Upper German [ɛː]. For coda clustering, [ɛː] has the highest probabilities in Low German and West Central German and East Central German has [aː] and both Upper German areas have [uː] as maximum.

When calculating the Pearson's product–moment correlation between the coda clustering probability of the short vowels [a, ɛ, ɪ, ɔ, ʊ] and their percentage (see Figure 3), the result is a slightly negative correlation (−0.42; $p = 0.033$). In contrast, the correlation between the onset clustering probability of the long vowels [aː, ɛː, iː, oː, uː] and their percentage given in Figure 4 was found to be not significant (0.034; $p = 0.87$).

Accordingly, there is no positive correspondence between the frequency and consonant clustering probability for long and short vowels so that Question 3 has to be answered in the negative for these two vowel types. In contrast, Question 2 can be affirmed: foremost, the clustering probability is determined by the vowel type (short ~ high coda and long ~ balanced) and dialect area (North–South increase). Inside this framework, preferences of the individual vowels become visible: [ʊ] is the most extreme case with the highest probability for coda clusters in all areas but nearly no onset clusters. [uː] tends to have the lowest clustering rates of all the long vowels (with the two Upper German areas as an exception). [aː], [ɛː] and [oː] show the balanced pattern that is characteristic for long vowels while [iː] and [eː] have a strong onset clustering preference like the diphthongs.

3.3.2. Diphthongs

As found in Lameli and Link (forthcoming), all three diphthongs in Figure 6 have a clear preference for onset clustering. [aɪ], which is the most frequent one, also has the highest onset clustering rates, except for West Upper German with a higher rate for [ɔɪ].

(25) t͡svɔɪ (zwei, 'two', Burgrieden, Swabian);
(26) flɔɪʃ (Fleisch, 'meat', Bempflingen, Swabian).

[ɔɪ] in (25) and (26) traces back to MHG <ei>, which either became [ɔɪ] or [aɪ] in Swabian or the previously described [ɔɐ] in Bavarian (Streck 2019, p. 214). However, [ɔɪ] is still much less frequent than [aɪ] (see Figure 4). [ɔɪ] has no phonotactic clusters at all in East Upper German, but Figure A6 in Appendix A reveals that it does have morphonotactic clusters (16% onset and 4% coda).

In contrast to the monophthongs, diphthongs show a significant positive correlation between their onset clustering probability and their frequency (0.55; $p = 0.032$).

Accordingly, there is a frequency effect working here, so that Question 3 has to be confirmed at least for diphthongs while they also do not display any clear individual preferences (Question 2) like the short and long vowels.

3.4. *Consonant Clustering and MHG Vowels*

In order to check whether the consonant clustering probabilities of the vowels in the five dialect areas are related to the original MHG vowels (Question 4), the data from the previous analysis were separated into groups of underlying MHG <a>, <â>, <e>, <ê>, <i>, <î>, <o>, <ô>, <u> and <û> and the umlauts and diphthongs <æ>, <œ>, <ü>, <iu>, <ei>, <ie>, <ou> and <uo> (independently of the actual vowels that were realized by the speakers). MHG <ä>, <ö> and <öe> were left out since the first was lacking in data and the latter two are only represented each by a single lemma, namely *Köln* and *Heu*. MHG <üe> has not been included since none of its 150 instances features onset or coda clusters so that its bar plot would be a row of ten zeros. An overview over the lemmata belonging to each MHG vowel is also given in Appendix A.

3.4.1. MHG Short and Long Monophthongs

The consonant clustering probabilities for MHG short and long monophthongs are given in Figure 7. When comparing the results of the MHG vowels with the modern ones, the MHG short vowels also have a coda clustering preference and the probability increase from North to South (Lameli 2022, pp. 262–66; Lameli and Link forthcoming) is observable as well, especially for <a>, <e> and <u>.

However, the MHG long vowels have very high onset clustering probabilities and <ê>, <î> and <ô> are completely lacking coda clusters. This strongly contrasts with the balanced clustering preference known for modern long vowels. The only exception is <â> with a higher coda clustering probability, which can be related to the lemma *Abend* as in the examples (27)–(28) below.

(27) aːmt (Abend MLG âvent, 'evening', Wolgast, Mecklenburg-Vorpommersch);
(28) oːmt (Abend MHG âbent, 'evening', Clausthal-Zellerfeld, Upper Saxon).

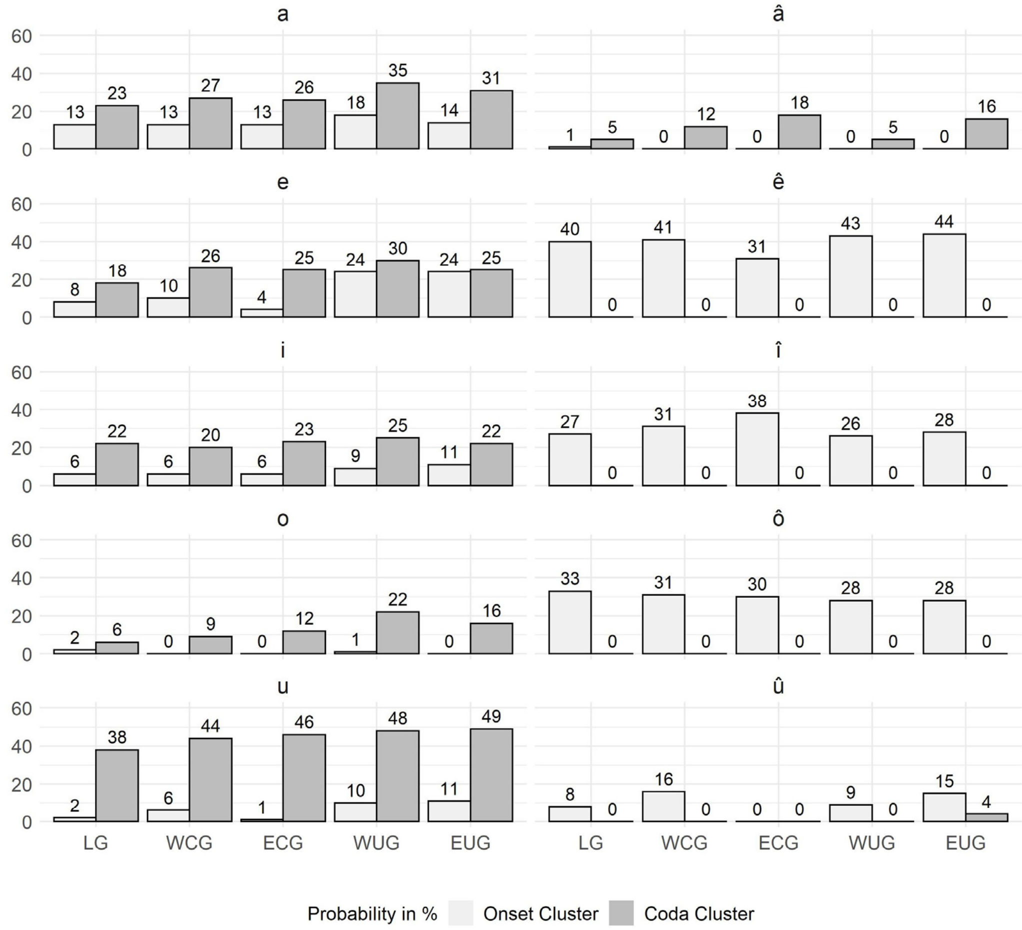

Figure 7. Consonant clustering probability for the MHG short and long monophthongs.

In the morphologically mixed data analyzed in Figure A7 in Appendix A, there are also other lemmas involved as in (29)–(31).

(29) ghapt (gehabt MHG gehât, 'had', Pielenhofen, North Bavarian);
(30) bʁɑːxt (gebracht MHG gebrâht, 'brought', Wildpoldsried, Swabian);
(31) vaːʃt (warst MHG vârest, '(you) were', Vielbrunn, Rhine Franconian).

(29)–(31) became monosyllabic due to schwa syncope and (27) and (28) also underwent the assimilation of [b] and [n].

The strong onset preference of the MHG long vowels is insofar surprising since it is the typical clustering pattern of the modern diphthongs (see Figure 6). This can be related to the New High German Diphthongization, which caused the MHG long vowels <î>, <iu> and <û> to become [aɪ], [ɔʏ] and [aʊ] in the High German Area (except Alemannic; see Streck 2019, p. 222). However, this does not explain the rather balanced clustering probability of modern long vowels.

A possible reason might be New High German Vowel Lengthening in open syllables and monosyllabic words (Paul 2007, pp. 80, 82). Since this also caused the lengthening of words with a short vowel plus coda cluster, it can be related to the balanced clustering preference of modern monosyllabic lexemes with a long vowel.

3.4.2. MHG Umlauts and Diphthongs

Figure 8 displays the consonant clustering probabilities for the MHG umlauts and diphthongs. The umlauts have very low values but prefer coda clusters like the short vowels. The MHG diphthongs <ei> and <ou> show a preference for onset clusters like the modern diphthongs. MHG <ie> and <uo> both underwent New High German Monophthongization and became NHG long vowels /i:/ and /u:/ and in the morphologically mixed data in Figure A8 in Appendix A, they have more balanced clustering patterns like the modern long vowels. In the following, it will be further explored to what amount a transition from MHG short vowels to NHG long vowels took place.

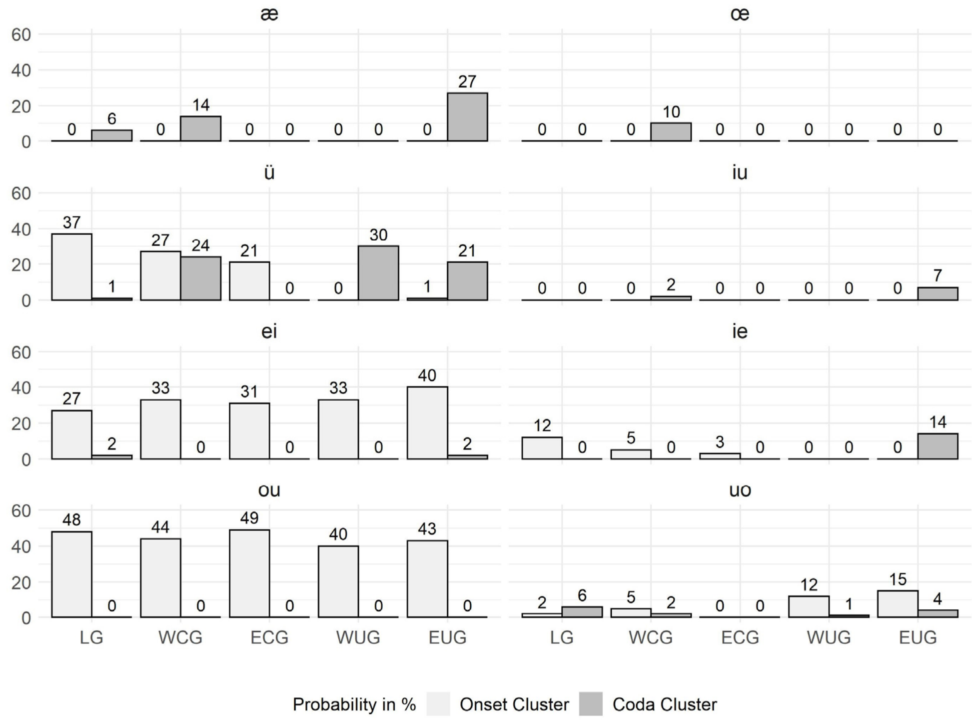

Figure 8. Consonant clustering probabilities for the MHG umlauts and diphthongs.

3.4.3. MLG/MHG-NHG Vowel Group Transition I: CV Percentages

For a better understanding of the previously described balanced clustering probabilities of modern long monophthongs, all monosyllabic lexemes with a long vowel were extracted for each of the five areas and the percentage of their CV structures was calculated.

LG: CV: (35.5%); CV:C (33.8%); CCV:C (11.9%); V:C (6.1%); CV:CC (5.2%); CCV: (4.8%); V:CC (2.3%); CCV:CC (0.2%); V: (0.1).

WCG: CV: (39.7%); CV:C (30.4%); CV:CC (14.6%); CCV:C (5.6%); CCV: (5.4%); V:C (3.7%); CCV:CC (0.6).

ECG: CV: (53.1%); CV:C (26.9%); CV:CC (6.9%); CCV: (6.5%); CCV:C (2.9%); V:C (2%); V:CC (0.8%); CV:CCC (0.4%); V:CCC (0.4).

WUG: CV: (44.3%); CV:C (18.6%); CCV: (14.2%); CV:CC (10.7%); CCV:C (8.7%); V:C (1.6%); CCV:CC (1.2%); CV:CCC (0.4%); V: (0.4).

EUG: CV: (36%); CV:C (21.5%); CV:CC (16.8%); V:C (9.8%); CCV:C (8.7%); CCV:CC (2.8%); CCV: (2.3%); V:CC (1.3%); V: (0.6).

The data reveal that the double-clustering CCV:CC has very low percentages ranging from 0.2% (LG) to 2.8% (EUG). However, types with either onset (CCV:C, CCV:) or coda clustering (CV:CC, V:CC, CV:CCC) are much more frequent. CV:CC ranges from 5.2% (LG) to 16.8% (EUG) and CCV:C is in a similar range from 2.9% (ECG) to 11.9% (LG). The observation that CV:CC is the third most frequent type (after CV: and CV:C) in most areas (except LG and WUG) supports the hypothesis that an MLG/MHG-NHG vowel group transition from short to long vowels has occurred. But to confirm this hypothesis, it is necessary to precisely investigate which MLG/MHG words with a short vowel underwent lengthening in which areas and which historical sound change processes are responsible for particular spatial patterns.

3.4.4. MLG/MHG-NHG Vowel Group Transition II: Words

Figure 9 gives an overview of all words from MHG (and from MLG for Low German) that became lengthened and monosyllabic in the respective dialect areas. Interestingly, the hierarchical clustering[7] groups the data according to known dialect classifications (Wiesinger 1983a): Upper and Central German form a larger High German group, which is finally paired with Low German.

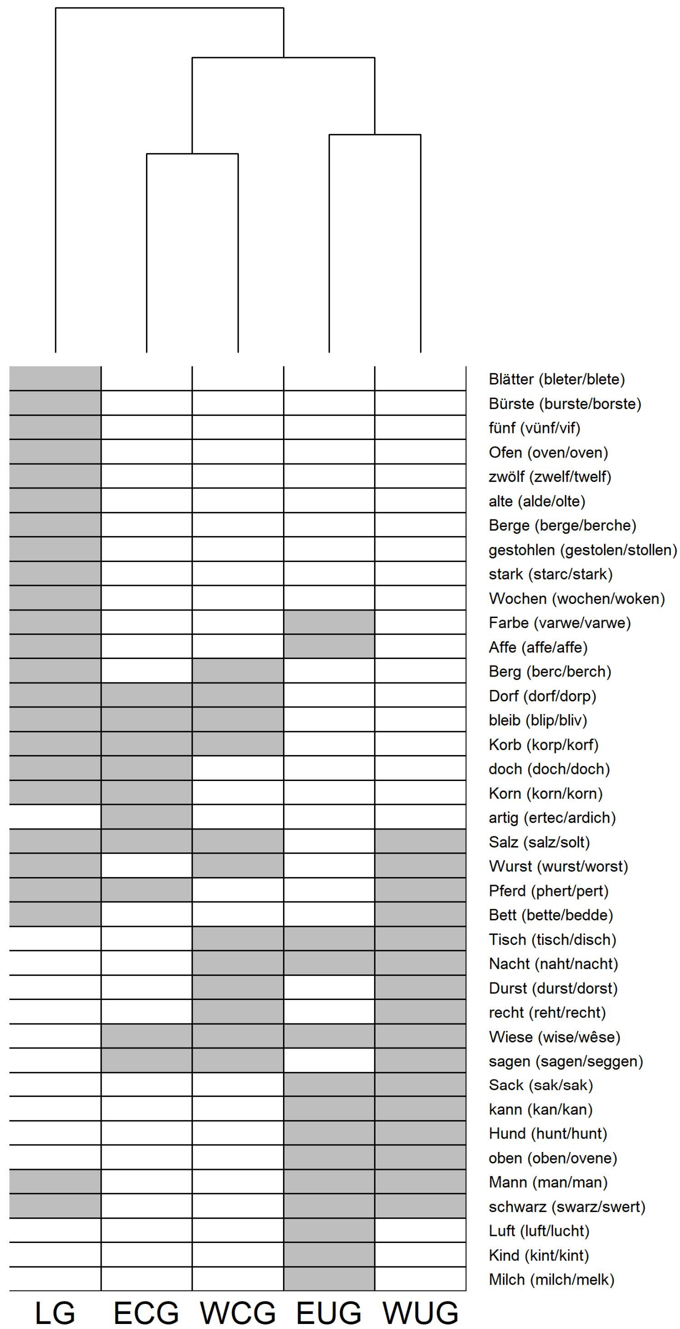

Figure 9. MHG words (and MLG words for Low German) with a short vowel that became lengthened and monosyllabic in the five dialect areas. The first word in brackets is MHG and the second is MLG.

Liberman (1997, p. 99) notes that vowel lengthening (a transition from short to long vowels) was caused by the MHG schwa apocope. This also corresponds to the observation that the largest numbers of vowel group transitions in Figure 9 are found in the regions with schwa apocope, namely Low and Upper German (see Knöbl and Nimz 2014, p. 101). However, this process does not always cause new consonant clusters when it goes in hand with the deletion of consonants such as *r* in examples (32), (40), (41) and (42), or *n* due to nasalization in Swabian (39) or the assimilation of [b] or [f] with [n] to [m] like in (33)–(36).

(32) bɛːç (Berge, 'mountains', Mirow, Mecklenburg-Vorpommersch dialect);
(33) fiːf (fünf, 'five', Wangerooge, North Saxon);
(34) oːm (Ofen, 'oven', Adorf, Westphalian);
(35) oːm (oben, 'above', Schachach, Central Bavarian);
(36) oːl (alte, 'old', Bennin, Mecklenburg-Vorpommersch dialect);
(37) vɛːk (Wochen, 'weeks', Jever, North Saxon);
(38) ɔːp (Affe, 'monkey', Wolgast, Mecklenburg-Vorpommersch dialect);
(39) mãː (Mann, 'man', Schömberg, Swabian);
(40) peːt (Pferd, 'horse', Brockhausen, Westphalian);
(41) faːp (Farbe, 'color', Reelkirchen, Westphalian).

But to explain the increase in the coda clustering probability of the long vowels from MLG/MHG to NHG, focus is set on sound change processes leading to monosyllabic lexemes with a long vowel plus coda cluster. From all CV:CC syllables in the data in Section 3.4.3, 45.1% are found in Upper German, 25.5% in Low German, 22.1% in West Central German and 7.2% in East Central German. High percentages correspond to regions with schwa apocope, but there are other phenomena involved as well.

(a) Lengthening instead of *r*-vocalization in Low German

Niekerken (1963, p. 169f) notes that vowel lengthening occurs in Low German in contexts where High German dialects have *r*-vocalization as in (42), (44) and (46) below (the additional examples (43), (45) and (47) from Bavarian were added for comparison since they have the same syllable structure as the Low German examples).

(42) byːst (Bürste, 'brush', Huddestorf, North Saxon);
(43) bɪɐst (Bürste, 'brush', Oberau, Central Bavarian);
(44) vuːʃt (Wurst, 'sausage', Bennin, Mecklenburg-Vorpommersch dialect);
(45) vuɐʃt (Wurst, 'sausage', Oberviehbach, Central Bavarian);
(46) døːst (Durst, 'thirst', Wüllen, Westphalian);
(47) dʊɐʃt (Durst, 'thirst', Langenbruck, Central Bavarian).

(b) Monosyllabic lengthening before consonant clusters (*Schwerschlussdehnung*)

Seiler and Würth (2014, p. 149) describe that Bavarian and Thuringian have monosyllabic lengthening before consonant clusters (*Schwerschlussdehnung*), like in the examples (48)–(52).

(48) lʊːft (Luft, 'air', Oberau, Central Bavarian);
(49) kiːnt (Kind, 'child', Ramsau, Central Bavarian);
(50) zoːlts (Salz, 'salt', Wieda, Thuringian);
(51) huːnt (Hund, 'dog', Pöttmes, Central Bavarian);
(52) rɛːçt (recht, 'right', Maxweiler, Central Bavarian).

Liberman (1997, p. 102) points out that this phenomenon is called *Nominativdehnung* in the older literature, since it marks a contrast between singular (long vowel) and plural (short vowel) in otherwise identical forms. However, lengthening is not restricted to nouns, but also occurs in predicative adjectives as in (52) from Wenker sentence 35 *Das war recht von ihnen*, 'That was right of them' (p. 102).

(c) Open syllable lengthening + schwa syncope

The morphologically mixed data analyzed in Figure A9 in Appendix A also reveal that new morphonotactic coda clusters also emerge when Open Syllable Lengthening occurs in combination with schwa syncope like in (53)–(55). (56) also illustrates how Central Bavarian *l*-vocalization (see Koch 2019, p. 284; and Wiesinger 1990, p. 463) introduces a diphthong [ɔɪ] instead of a long vowel.

(53) stoːln (gestohlen, 'stolen', Eversen, North Saxon);
(54) koːln (Kohlen, 'coal', Linz, Upper Saxon);
(55) kʃtoːln (gestohlen, 'stolen', Pielenhofen, North Bavarian);
(56) kʃtɔɪn (gestohlen, 'stolen', Ramsau, Central Bavarian).

3.4.5. Summary

For long vowels, Figure 1 displays the balanced clustering probabilities for onset and coda. While the onset clustering is quite similar among the five areas, the coda clustering probability increases from North to South (see also Lameli and Link forthcoming). When comparing the clustering probabilities of the data grouped by particular vowels (Figures 5 and 6), the balanced pattern is mainly present although tendencies for different clustering preferences for each vowel become visible. But when the data are grouped according to underlying MHG (or MLG) vowels (ignoring the vowel that was uttered; Figures 7 and 8), it is striking that the MHG long vowels have the same strong onset preference as NHG diphthongs (with <â> being the only exception). One explanation for this is New High German Diphthongization, which caused MHG long vowels <î>, <iu> and <ô> to become NHG diphthongs [aɪ], [ɔʏ] and [aʊ]. Furthermore, MHG diphthongs <ie> and <uo> (which underwent New High German Monophthongization and became long vowels [i:] and [u:]) have a more balanced pattern. However, it was found that further historical sound change processes were involved as well, namely lengthening in the context of *r*-deletion and Open Syllable Lengthening (both paired with schwa apocope) and monosyllabic lengthening before consonant clusters in Central Bavarian. The latter also explains why East Upper German is the only area in which the coda clustering probability of the long vowel is higher than that for onset clustering (see Figure 1).

4. Discussion

Considering the findings of Lameli (2022, pp. 262–66) and Lameli and Link (forthcoming), four questions were raised in the introduction:

1. Which syllable structures (CV) shape the pattern found by Lameli and Link (forthcoming)?
2. Are some vowels more inclined to have consonant clusters than others?
3. Do more frequent vowels also have a higher probability for clustering?
4. Can differences in clustering be traced back to MHG/MLG vowels?

Question 1 was answered in 3.1 where it was found that the spatial North–South increase as well as the preferences of the three vowel groups (V, V:, VV) are dominated by syllable types with biconsonantal clusters in the onset and/or coda. However, this is not surprising in terms of markedness: clusters with three or more clusters are more marked than biconsonantal ones in the onset (*Silbenanlautgesetz*; Hall 2000, p. 213) as well as in the coda (*Silbenauslautgesetz*; Hall 2000, p. 214).

It is of great interest to go deeper into the consonant clusters below the level of the CV structure. Such an analysis is given in Harnisch (1987, pp. 255–59), who examined the relation between vowel quantity and particular consonant clusters for the dialect of Ludwigsstadt (Thuringian–East Franconian transition area). However, performing this for the data studied in this work would have extended the scope of a single paper. Thus, this is going to be performed in a second work that is currently under preparation.

Considering Question 2, the findings of Section 3.3 could confirm that particular vowels do have clustering preferences of their own. Nevertheless, these preferences are ruled by the characteristic pattern of the vowel group (V, V:, VV) described by Lameli and Link (forthcoming): monosyllabic lexemes with a short vowel in the nucleus prefer coda clustering while those with a diphthong prefer onset clusters and those with a long vowel have a balanced pattern.

Question 3 was tested by calculating the correlation between the clustering probabilities found in Section 3.3 and the vowel percentages from Section 3.2. The monophthongs (V, V:) had no positive correlations so that a frequency effect can be ruled out for them. In contrast, the diphthongs were found to be positively correlated with frequency.

For Standard German, Beedham (1995, p. 146) describes a tendency for particular consonant + vowel and vowel + consonant sequences to be predictors of whether a verb has strong or weak inflection. It would be highly interesting to test the five dialect areas for this tendency as well but in order to gain valuable insights, a much broader database would be necessary: Beedham (1995) is based on 169 strong and 1467 weak verbs whereas the Wenker sentences from the PhonD2-Corpus have a type count of 139 verbs (for the full corpus), from which 73 are strong and 66 are weak.

For answering Question 4, the data were grouped according to the underlying MHG (and MLG) vowels in Section 3.4. The analyses of the clustering probability in Figures 7 and 8 reveal an interesting phenomenon: instead of the previously described threefold pattern, a twofold pattern of consonant clustering preference becomes visible. Words with an MHG (and MLG) long vowel have the same strong onset preference as the modern diphthongs. This raised the question of how the balanced clustering preference of the modern long vowel did emerge. The analysis of the CV structure in Section 3.4.3 illustrated that there are only a small number of CCV:CC syllables, but a high number of CV:CC and CCV:C, which lead to the assumption that there must have been a transition from MHG (and MLG) short vowels to modern long vowels. In Section 3.4.4, three historical sound change processes that have caused such a transition were identified and validated with a large number of examples from the PhonD2 Corpus.

In Lameli and Link (forthcoming), the strong coda cluster preference of monosyllabic words with a short vowel was explained with regard to the compensation of the syllable weight (Auer 1991, p. 8), which was not necessary for long vowels and diphthongs counting as heavy due to being bimoraic. This compensation of the syllable weight actually corresponds more closely to the twofold pattern of the clustering preference found when analyzing the data based on the underlying MHG/MLG vowels in Section 3.4. Ryan (2016, p. 721) notes that most languages with a weight-sensitive stress system have a binary distinction but more fine-grained distinctions are also found (p. 728). A possible explanation for the threefold clustering pattern of the modern dialects, namely the tendency for long vowels towards balanced probabilities, could be in terms of the Weight-by-Position parameter indicating whether the coda is treated as moraic or not in a language (Hayes 1989, p. 258). In the context of Monosyllabic Lengthening before consonant clusters (*Schwerschlussdehnung*), Seiler and Würth (2014, p. 151) suggest that Thuringian does not have Weight-by-Position and consonant clusters are not considered for calculating the syllable weight. Furthermore, long monophthongs can also be considered as lighter than diphthongs: Ryan (2016, p. 730) notes that vowels with a longer duration correlate with a greater syllable weight and according to Delattre (1964, p. 91), diphthongs have averagely longer durations than the long and short monophthongs.

In addition to the compensation of the syllable weight, Lameli and Link (forthcoming) also assumed that the preference for coda clustering is helpful for marking word boundaries in monosyllabic words, which supports speech perception (Caro Reina 2019, p. 267). With regard to the MHG/MLG twofold vs. the modern threefold pattern, there is probably a tendency towards the strengthening of the phonological word in monosyllabic lexemes, which becomes interesting in the context of the syllable vs. word language typology framework. Szczepaniak (2007) suggests that German was once a language that put a strong emphasis on the syllable (Old High German as a syllable language), which changed to a focus on the phonological word (New High German as a word language). The stage where this process was consolidated is Early New High German, which corresponds to the twofold clustering preference of MHG/MLG becoming threefold in the five dialect areas (Szczepaniak 2007, p. 331). Furthermore, Szczepaniak (2007, p. 329) emphasizes that such a change from syllable to word language goes in hand with optimizing the structure of the phonological word by either word final consonant epenthesis (which leads to consonant clustering) or by word-related restrictions on phoneme distribution.

But since the analyses of this work are focused on monosyllabic lexemes, it is not known whether the boundary strengthening observed for them also takes place in polysyllabic words. Thus, future work does not only have to consider the consonant clusters themselves (which is going to be the case in a second paper) but also explore polysyllabic words under these aspects.

Funding: This research was funded by the Deutsche Forschungsgemeinschaft (grant number 432304149).

Data Availability Statement: The data analyzed in this work are freely available at https://dsa.info/PhonD2/ (accessed on 17 July 2024). All analysis scripts used in this work can be obtained on request from the author.

Conflicts of Interest: The author declares no conflicts of interest.

Abbreviations

MHG	Middle High German
MLG	Middle Low German
NHG	New High German
LG	Low German
WCG	West Central German
ECG	East Central German
WUG	West Upper German
EUG	East Upper German

Appendix A

Total count of monosyllabic lexemes (with and without morpheme boundaries): 28.570.
Counts per area: **LG:** 10.774; **WCG:** 5937; **ECG:** 3105; **WUG:** 2700; **EUG:** 6054.
Frequency counts of CV structure:

LG: CVC (2543); CV:C (1433); CVCC (1297); CV: (1099); CVVC (951); CVV (744); CCV:C (523); VC (502); CCVVC (242); CCVV (240); V:C (201); CCV: (185); CV:CC (145); CCVC (143); VVC (135); CCVCC (116); CV (50); CVCCC (40); CVVCC (40); V:CC (30); CCV:CC (28); VV (24); V (20); VCC (15); CCVVCC (10); V: (7); VVCC (5) CCVCCC (3); CCCV:C (1); CV:CCC (1); CVCCCC (1).

WCG: CVC (1438); CVCC (799); CV: (632); CVVC (541); CVV (512); CV:C (496); VC (295); CCV:C (176); CCVVC (140); CCVC (139); CCVV (137); CCV: (128); VVC (80); CV:CC (79); VCC (65); CVVCC (53); V:C (42); CCVCC (34); CV (33); VV (22); V (20); V: (19); CVCCC (16); CCCVC (8); VVCC (7); CCV:CC (6); CCCV:C (4); CVCCCC (4); V:CC (4); CCCVCC (3); CCCVVC (1); CCVVCC (1); CV:CCC (1).

ECG: CVC (676); CVCC (506); CV:C (387); CV: (365); CVVC (214); CVV (188); VC (150); CCV:C (113); CCVVC (93); CCVV (86); CCV: (59); VVC (45); CCVC (41); CV:CC (33); CV (27); VCC (20); V:CC (19); V:C (17); V: (13); V (9); CVCCC (8); CVVCC (7); VV (7); VVCC (7); CCVCC (5); CCV:CC (3); VCCC (2); CCCVC (1); CV:CCC (1); V:CCC (1).

WUG: CVC (580); CVCC (365); CV: (317); CVVC (255); CV:C (210); CVV (172); VC (144); CCVV (102); CCVVC (75); CCV: (60); CCV:C (59); CCVCC (53); VV (50); CCVC (46); CV:CC (34); VCC (34); CVVCC (23); VVC (23); CVCCC (21); V (19); CCCVCC (13); V:C (11); V: (10); CV (9); CCV:CC (7); CCVVCC (2); V:CC (2); C (1); CCCV:C (1); CV:CCC (1); VVCC (1).

EUG: CVC (1026); CVVC (829); CVCC (724); CV: (561); CV:C (437); CVV (408); VC (281); CCVVC (263); CVVCC (186); CCV:C (171); CCVV (156); CV:CC (140); CCVCC (133); CCVC (90); CCV: (86); V: (75); VVC (67); V:C (61); CV (57); CVCCC (45); VCC (33); CCVVCC (31); V:CC (31); CCCVVC (26); CCV:CC (25); V (23); VVCC (21); CCCVCC (20); VV (20); CVVCCC (9); CCCV:C (6); CCCV:CC (6); CCCVVCC (3); CCCVC (2); CVCCCC (1); VCCC (1).

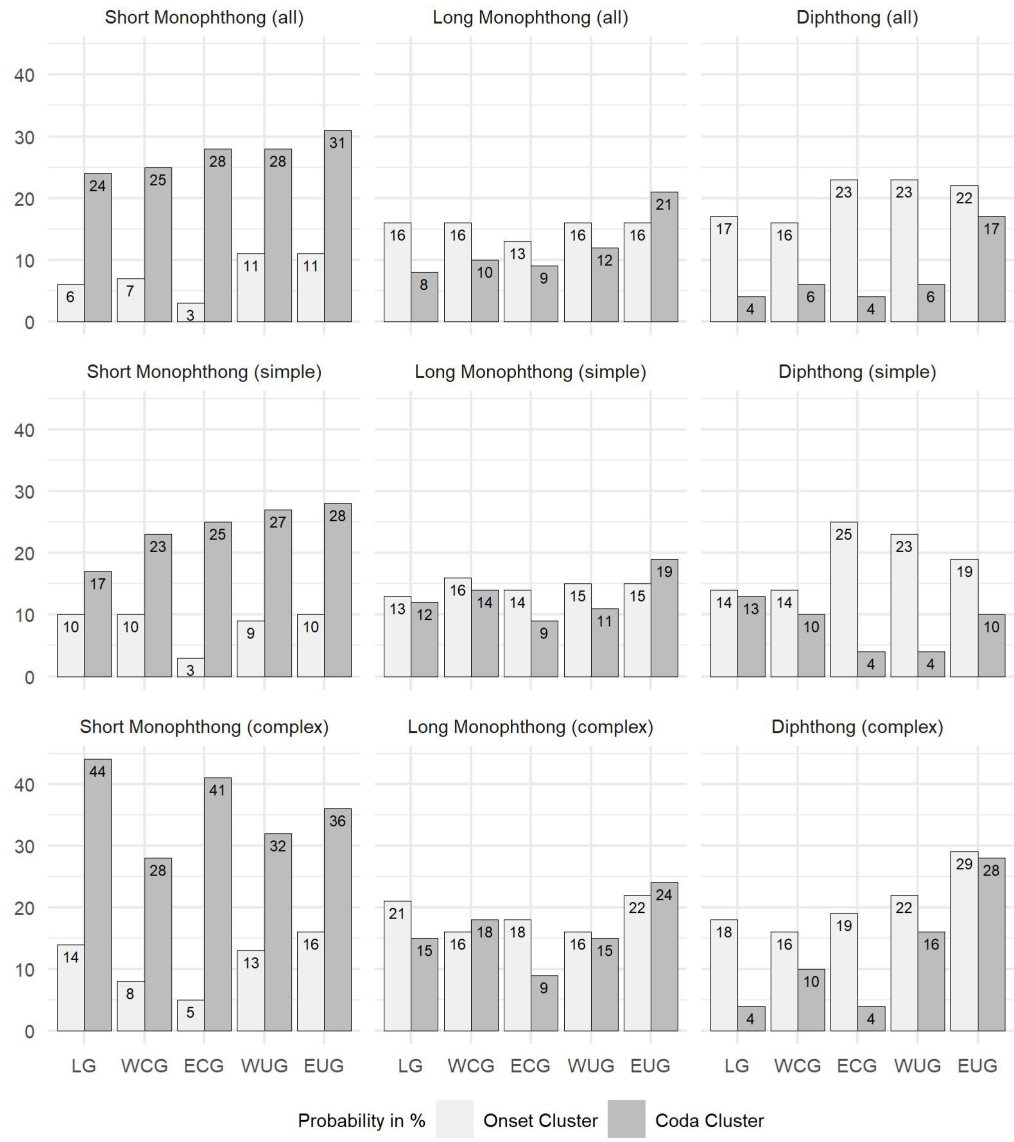

Figure A1. Probability of onset (lighter grey) and coda (darker grey) consonant clustering for different vowel types (V, Vː, VV) in dialects of Low German (LG), West Central German (WCG), East Central German (ECG), West Upper German (WUG) and East Upper German (EUG). The first row repeats Figure 6 and the second and third rows display the results for the same data split into morphologically simple and complex words.

Figure A1 has been taken from Lameli and Link (forthcoming) where it can be found as Figure 8 in the Discussion. In Lameli and Link (forthcoming), the following explanation was given for this analysis:

> "The results in Figure A1 reveal higher clustering rates for complex monosyllables; especially in the coda. This is in line with the findings of Dressler and Kononenko (2021, p. 43). However, the spatial pattern itself remains unchanged with the sole exception of the high coda-clustering rate of Low and East Central German complex monosyllabic words with a short vowel in the nucleus. The lower rates in the Upper German area can be attributed to a higher degree of segment-elision as in Swabian [kanʃ] vs. Standard German [kanst] "you can" (see also Streck 2019, p. 225)."

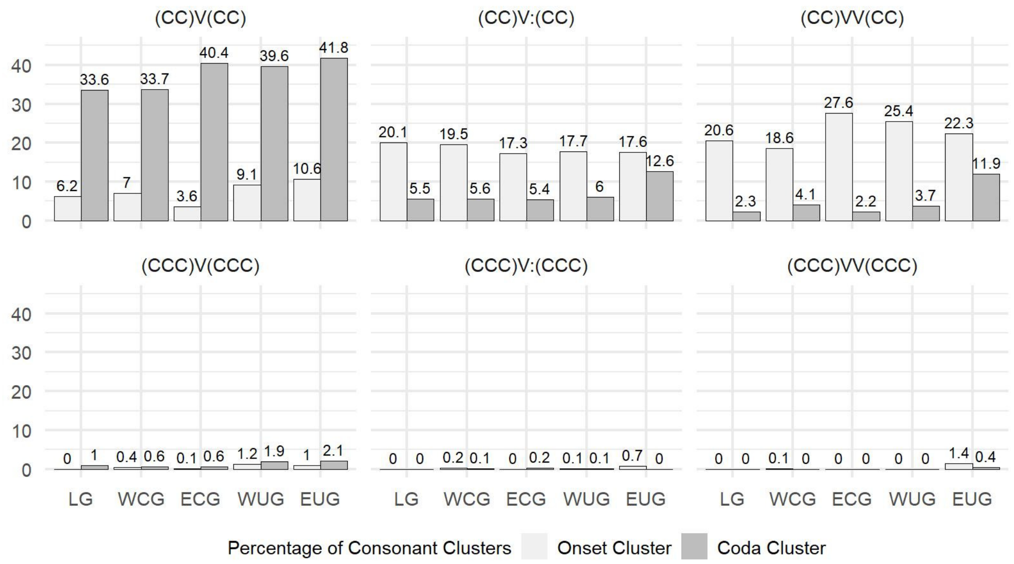

Figure A2. Percentage of onset (lighter gray) and coda (darker gray) CC or CCC consonant clusters for different vowel types (V, V:, VV) in the dialects of Low German (LG), West Central German (WCG), East Central German (ECG), West Upper German (WUG) and East Upper German (EUG). Analysis of morphologically mixed data.

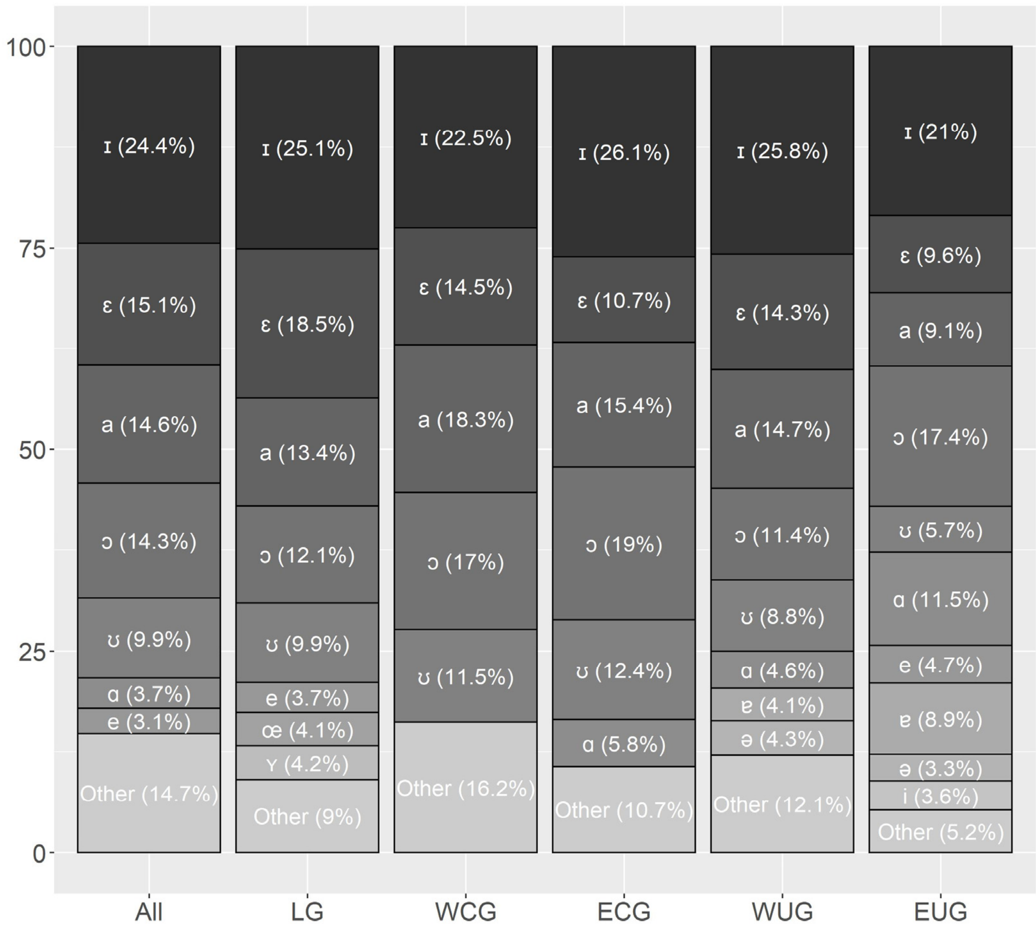

Figure A3. Percentage of short vowels in monosyllabic lexemes for all areas (All), Low German (LG), West Central German (WCG), East Central German (ECG), West Upper German (WUG) and East Upper German (EUG). Analysis of morphologically mixed data.

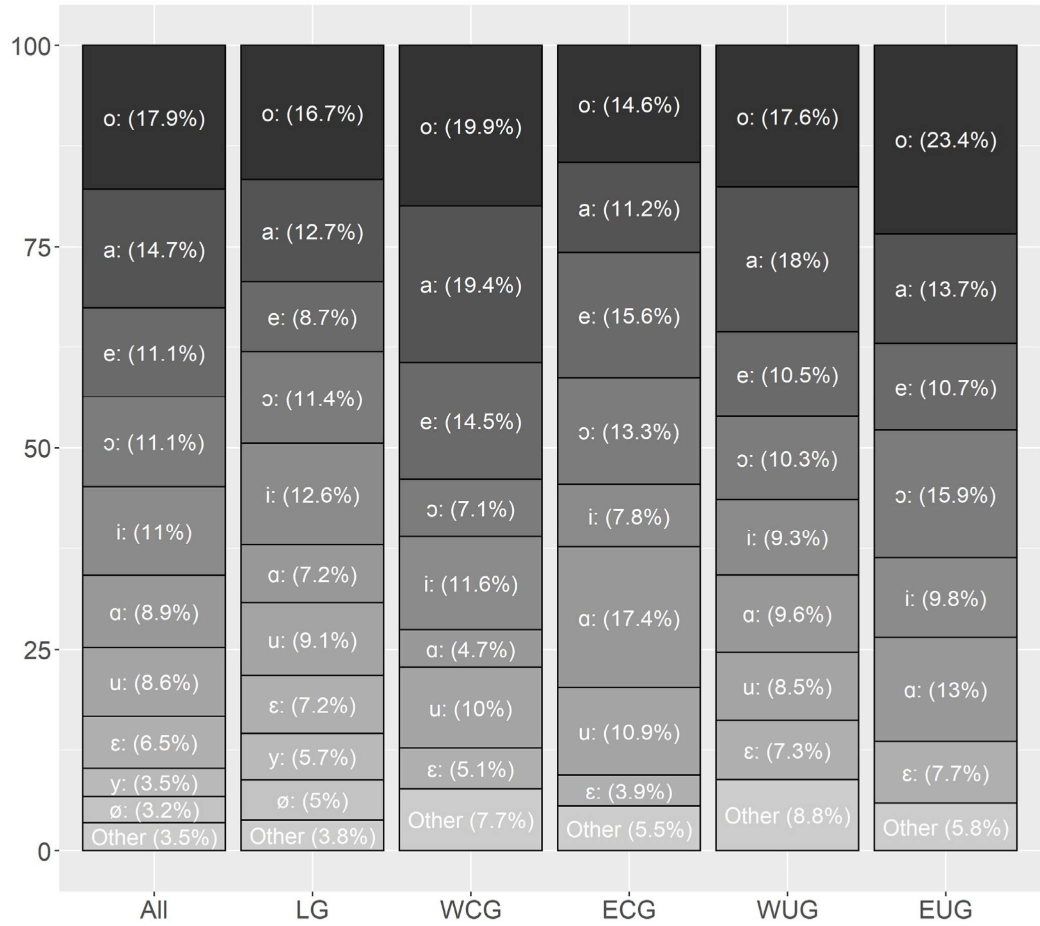

Figure A4. Percentage of long vowels in monosyllabic lexemes for all areas (All), Low German (LG), West Central German (WCG), East Central German (ECG), West Upper German (WUG) and East Upper German (EUG). Analysis of morphologically mixed data.

Figure A6 shows the same analysis as in Figure 6 but based on morphologically simple monosyllabic words. As expected from the results of A1, there is a general tendency for all areas alike to have slightly higher coda clustering rates than in Figure 6 but for EUG [ɛː], it is even 10% lower. Also, WUG [aʊ] has a 9% lower probability for onset clustering. The only exception is EUG [ɔɪ] with no clusters in Figure 6. But this is also not surprising since the 40 words responsible for this pattern are lemmata like Salz, bald, Feld and kalt, which all underwent Bavarian *l*-vocalization.

Table A1. Frequency table for short vowels.

LG-V	LG-Freq	WCG-V	WCG-Freq	ECG-V	ECG-Freq	WUG-V	WUG-Freq	EUG-V	EUG-Freq
ɪ	1182	ɪ	639	ɪ	377	ɪ	678	ɪ	224
ɛ	873	a	520	ɔ	274	a	388	ɔ	186
a	631	ɔ	482	a	223	ɛ	377	ɑ	123
ɔ	568	ɛ	413	ʊ	179	ɔ	300	ɛ	103
ʊ	465	ʊ	326	ɛ	154	ʊ	231	a	97
ʏ	198	e	83	ɑ	84	ɑ	121	ɐ	95
œ	195	ɑ	72	ə	31	ə	112	ʊ	61
e	174	ʏ	64	o	30	ɐ	107	e	50
ə	134	o	54	œ	24	œ	72	i	39
ɐ	129	œ	52	e	24	e	68	ə	35

Table A1. *Cont.*

LG-V	LG-Freq	WCG-V	WCG-Freq	ECG-V	ECG-Freq	WUG-V	WUG-Freq	EUG-V	EUG-Freq
ɑ	75	ə	41	ʏ	19	ʏ	58	œ	22
o	35	u	41	ɐ	10	i	53	o	16
i	27	i	21	i	8	o	32	ʏ	8
y	19	ɐ	19	u	6	u	28	u	6
u	16	y	18	ø	2	y	8	ø	4
ø	10	ø	12			ø	7	y	4
ã	6	ã	1			ã	2	õ	2
		õ	1			ẽ	1	ã	1
		~	1			õ	1	ũ	1

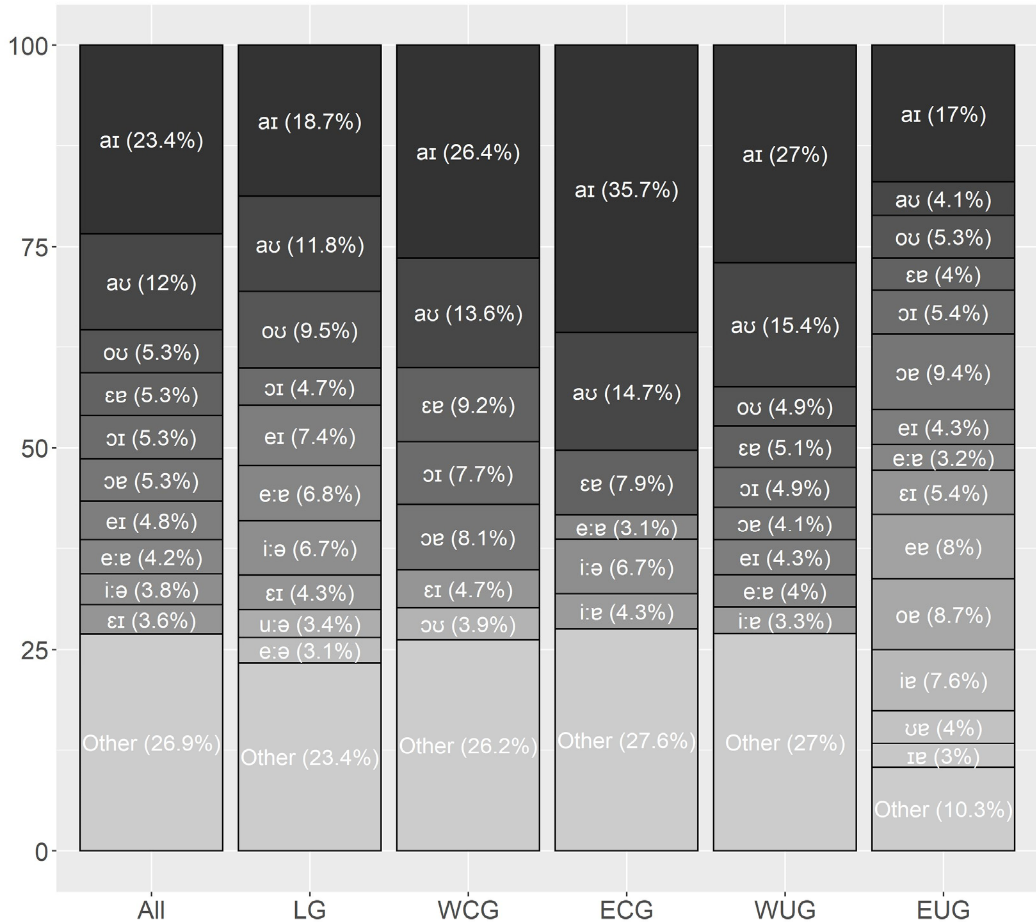

Figure A5. Percentage of diphthongs in monosyllabic lexemes for all areas (All), Low German (LG), West Central German (WCG), East Central German (ECG), West Upper German (WUG) and East Upper German (EUG). Analysis of morphologically mixed data.

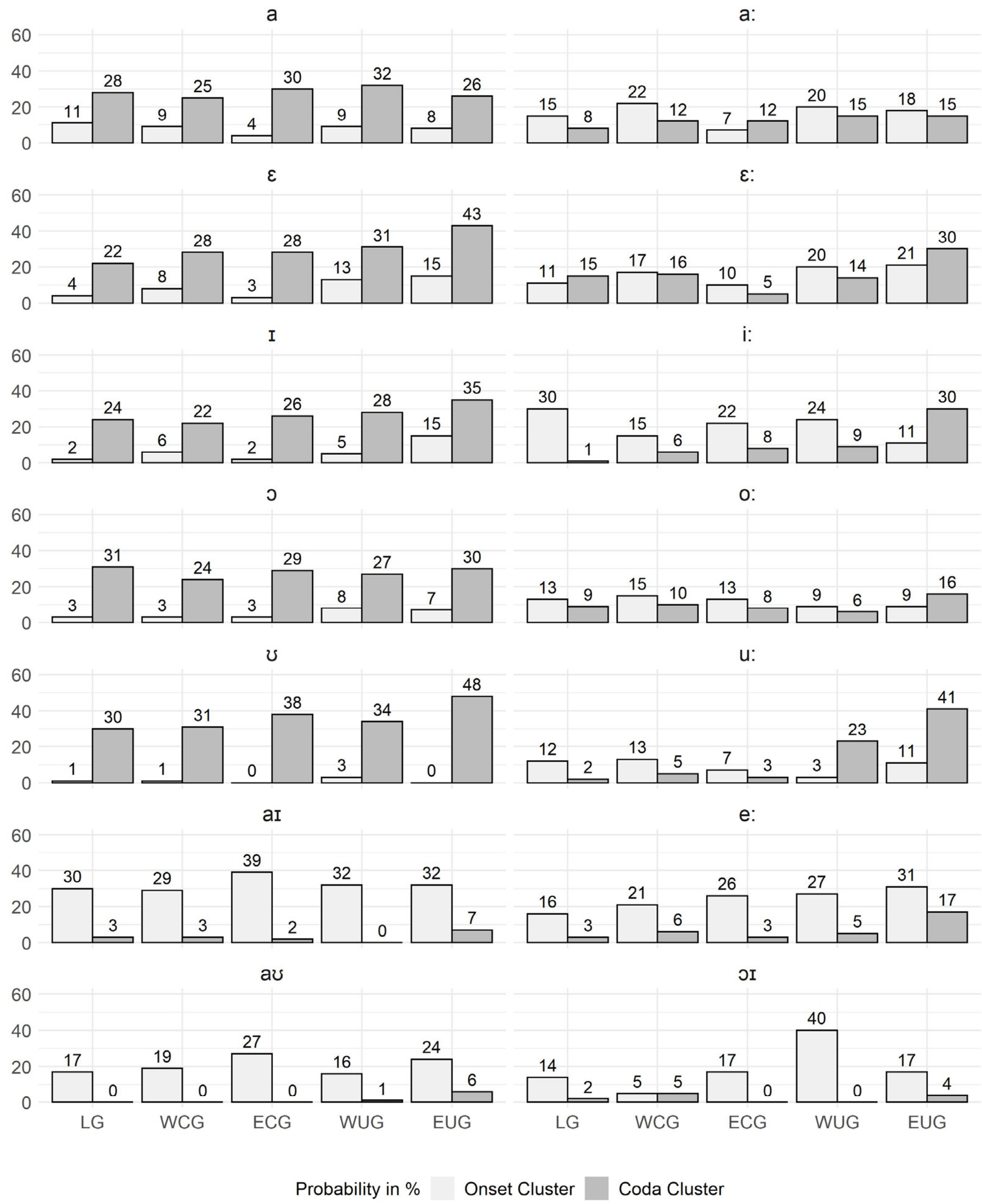

Figure A6. Consonant clustering probability for the vowels and diphthongs in the five dialect areas Low German (LG), West Central German (WCG), East Central German (ECG), West Upper German (WUG) and East Upper German (EUG). Analysis of morphologically mixed data.

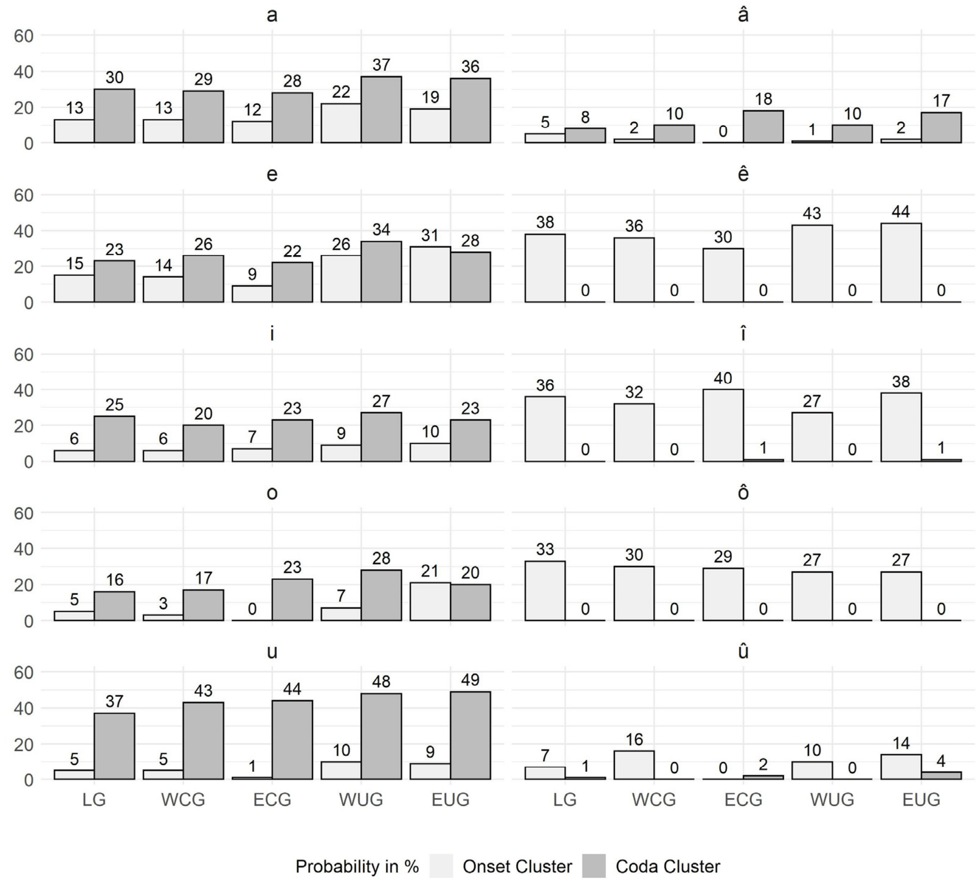

Figure A7. Consonant clustering probability for the MHG short and long monophthongs. Analysis of morphologically mixed data.

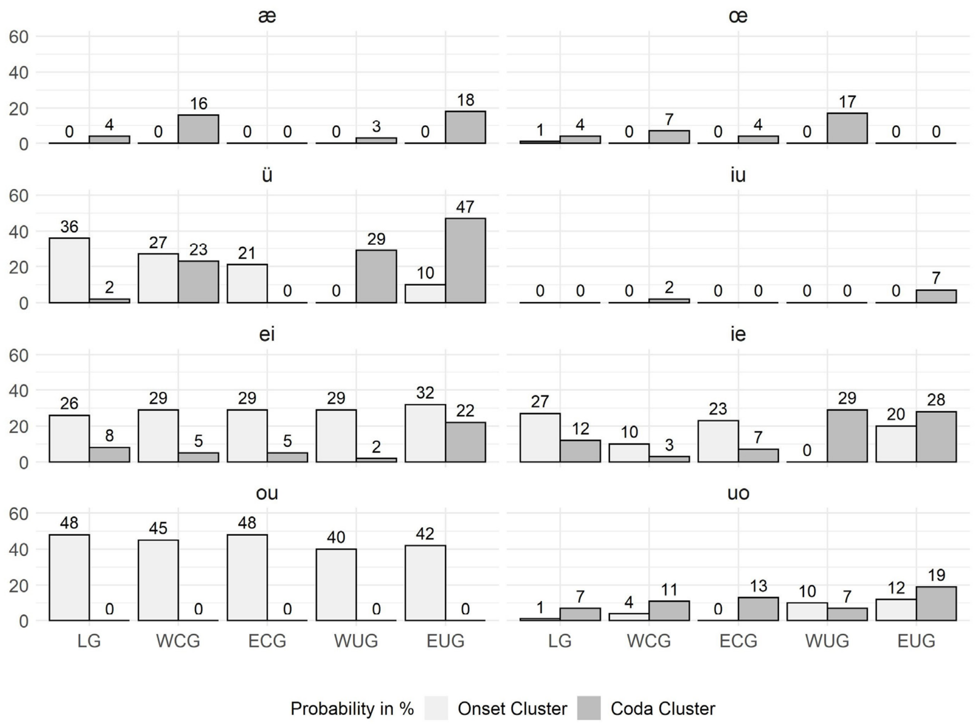

Figure A8. Consonant clustering probabilities for the MHG umlauts and diphthongs. Analysis of morphologically mixed data.

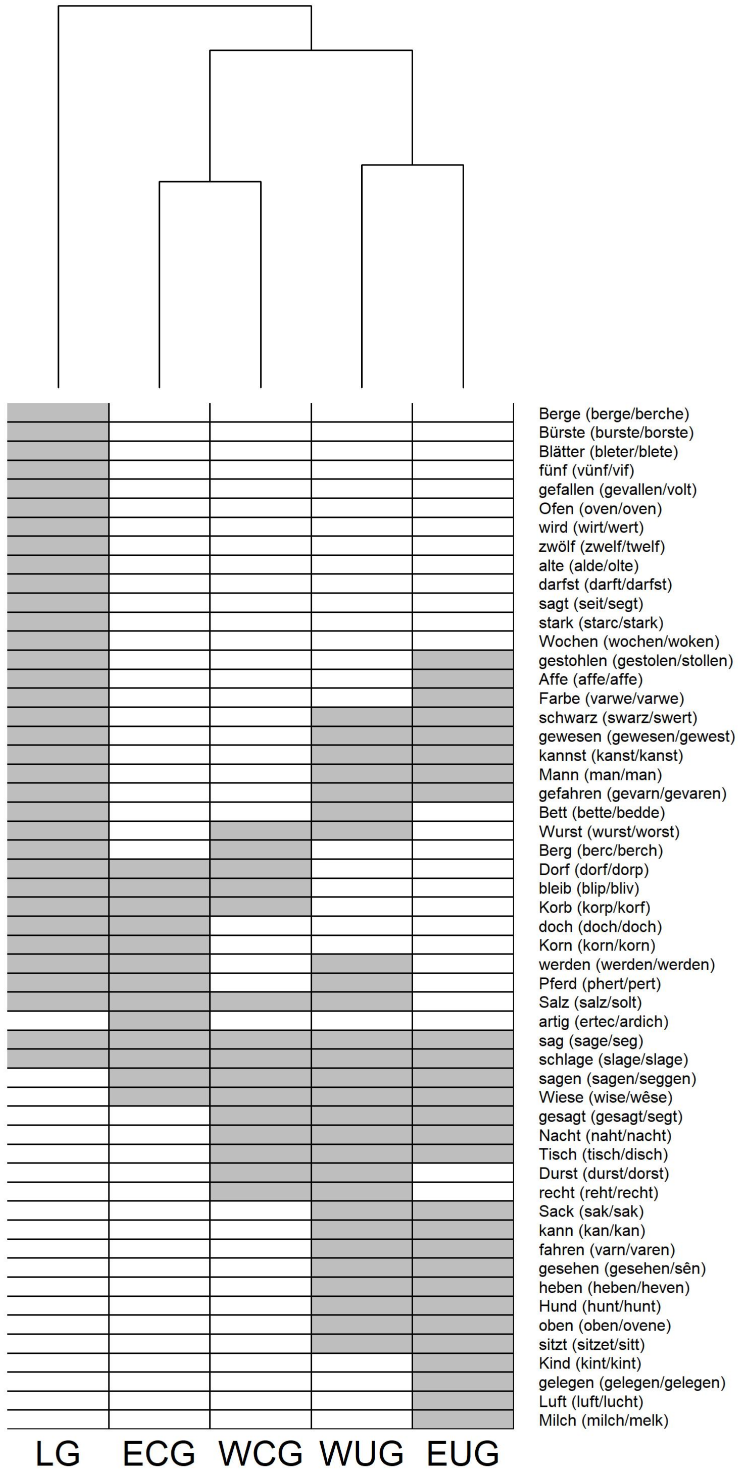

Figure A9. MHG words (and MLG words for Low German) with a short vowel that became lengthened and monosyllabic in the five dialect areas. The first word in brackets is MHG and the second is MLG. Analysis of morphologically mixed data.

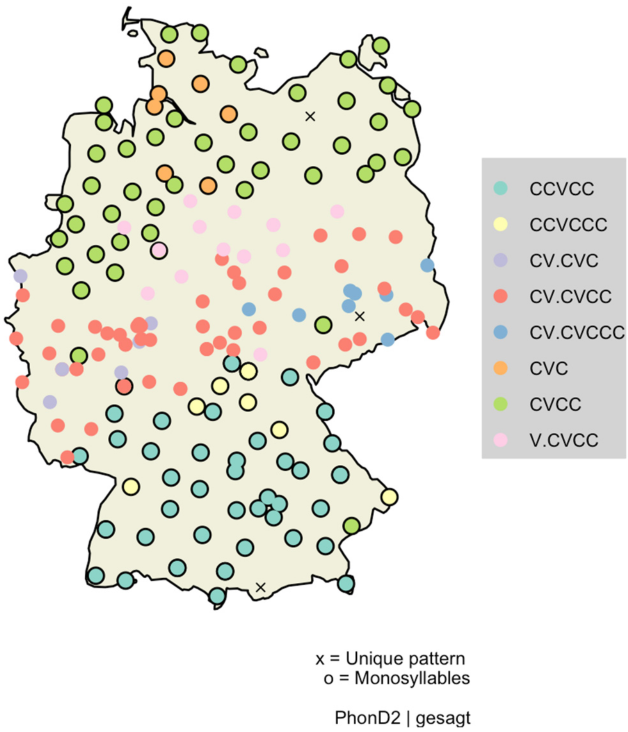

Figure A10. Map of the lemma *gesagt* from the PhonD2-website. All monosyllabic realizations of the lemma are marked with a black circle. In the South, there are the CV structures CCVCC and CCVCCC. (https://www.dsa.info/PhonD2/0aakagi42.html Last accessed: 17 July 2024).

Table A2. Frequency table for long vowels.

LG-V	LG-Freq	WCG-V	WCG-Freq	ECG-V	ECG-Freq	WUG-V	WUG-Freq	EUG-V	EUG-Freq
oː	596	oː	315	ɑː	175	aː	241	oː	200
aː	455	aː	307	eː	156	oː	235	ɔː	136
iː	450	eː	229	oː	146	eː	141	aː	117
ɔː	408	iː	183	ɔː	133	ɔː	138	ɑː	111
uː	324	uː	158	aː	112	ɑː	128	eː	92
eː	310	ɔː	112	uː	109	iː	125	iː	84
ɑː	257	ɛː	80	iː	78	uː	114	ɛː	66
ɛː	257	ɑː	74	ɛː	39	ɛː	98	ɐː	42
yː	205	øː	42	yː	16	ɐː	46	ʊː	22
øː	180	yː	41	øː	15	yː	31	uː	10
œː	95	œː	14	ɪː	10	øː	29	ɪː	9
ɐː	63	ʊː	14	œː	7	ʊː	28	øː	4
ɪː	27	ɪː	10	ʊː	7	ɪː	18	œː	4
ʏː	14	ʏː	5	əː	4	œː	12	~ː	3
ʊː	13	~ː	2	ʏː	3	ʏː	7	õː	3
ãː	2	ẽː	1	~ː	1	ãː	5	ãː	2
~ː	1	ũː	1			õː	4	ẽː	2
								~ː	2
								ĩː	1
								yː	1

Table A3. Frequency table for diphthongs.

LG-V	LG-Freq	WCG-V	WCG-Freq	ECG-V	ECG-Freq	WUG-V	WUG-Freq	EUG-V	EUG-Freq
aɪ	364	aɪ	354	aɪ	207	aɪ	325	aɪ	169
aʊ	230	aʊ	183	aʊ	85	aʊ	186	uɐ	97
oʊ	185	ɛɐ	124	ɛɐ	46	ɛ6	62	ɔɐ	94
eɪ	144	ɔɐ	109	i:ə	39	ɔɪ	59	oɐ	87
e:ɐ	133	ɔɪ	104	i:ɐ	25	oʊ	59	eɐ	80
i:ə	131	ɛɪ	63	e:ɐ	18	eɪ	52	iɐ	76
ɔɪ	91	ɔʊ	53	eɪ	15	ɔɐ	49	aɔ	69
ɛɪ	83	eɪ	38	ɔə	15	e:ɐ	48	ɛɪ	54
u:ə	67	i:ɐ	38	ɔɪ	15	i:ɐ	40	ɔɪ	54
e:ə	61	i:ə	34	u:ə	15	i:ə	28	oʊ	53
ɔɐ	58	e:ɐ	25	ɛ:ɐ	12	ʊɐ	27	eɪ	43
o:ə	56	uɐ	25	ɔɐ	9	iə	24	aʊ	41
yə	56	ɪɐ	21	oʊ	9	ʊə	22	ɛɐ	40
ɛɐ	51	u:ə	19	u:ɐ	8	u:ɐ	20	ʊɐ	40
ɔə	51	u:ɐ	19	ɑɐ	6	ɛ:ɐ	18	e:ɐ	32
ø:ə	39	aɐ	18	iə	6	eɐ	18	ɪɐ	30
øɪ	39	oʊ	18	iɐ	6	u:ə	18	ɔʊ	27
ø:ɐ	37	uə	15	o:ɐ	6	e:ə	17	ae	22
eɐ	32	ø:ɐ	14	y:ə	6	ɪɐ	17	oɪ	18
i:ɐ	32	ʊɐ	14	aɐ	5	o:ɐ	15	ɐɪ	15
o:ɐ	29	y:ɐ	14	e:ə	5	uə	15	œɪ	13
y:ə	26	e:ə	12	e:ɑ	5	ɐɪ	14	i:ɐ	13
œɐ	25	eə	12	ɛɪ	5	oə	14	u:ɐ	10
ɛ:ɐ	23	oɐ	12	ɪɐ	5	ɔʊ	14	ʊɪ	9
ɔʊ	21	ɛə	11	ø:ɐ	4	aɐ	13	o:ɐ	7
oɐ	20	ɔə	11	aɔ	4	ɛɪ	13	œe	6
œɪ	18	iɐ	10	ɔʊ	4	uɐ	13	ɑ:ɐ	6
u:ɐ	16	o:ɐ	9	aə	3	ɪə	12	ʊə	6
ɐɪ	14	ɑɪ	8	ɑɪ	3	o:ə	12	a~	5
ɪə	13	œɐ	7	eɐ	3	iɐ	11	uɪ	5
œə	11	ɛ:ɐ	7	oɐ	3	ɛə	10	aɐ	4
ɪɐ	11	eɐ	7	u:ɑ	3	ʊɪ	10	ɑe	4
uɐ	11	iə	6	ʊɐ	3	a:ɐ	9	ɑɔ	4
ʊɪ	10	yɐ	6	y:ɐ	3	ɔə	8	e:ə	4
iə	9	ɐɪ	5	ø:ə	2	oɐ	8	ɛ:ɐ	4
y:ɐ	9	uɪ	5	øɪ	2	eɑ	7	o~	4
iɐ	8	ʊɪ	4	a:ə	2	ʏə	7	yɐ	4
ɔ:ə	8	øɪ	3	ɑʊ	2	œɐ	6	øɪ	3
oɪ	8	œɪ	3	ɛə	2	eə	6	a:ɐ	3
eə	7	ɑɐ	3	iɑ:	2	œɪ	5	ɑɪ	3
ɛə	7	ɪə	3	o:ə	2	ɑɪ	5	ɪə	3
oə	7	ɔa	3	ɔ:ə	2	ɛ:ə	5	oe	3
ʊɐ	7	ʊə	3	ø:ɑ	1	uɪ	5	ɔe	3
yɐ	7	ʏɐ	3	øə	1	yə	5	ʏɐ	3
aɔ	6	ø:ə	2	øɐ	1	yɐ	5	aə:	2
ʏɐ	6	øɐ	2	ɑ:ə	1	ø:ɐ	4	ɑɐ	2
a:ə	5	a:ɐ	2	a:ɐ	1	e:ɛ	4	ae:	2
uə	5	ɑ:ɐ	2	aʏ	1	əɪ	3	a~	2
ʊə	5	iɔ	2	ɛ:ə	1	øə	3	o:ə	2
œ:ə	4	o:ə	2	ɛ:ɑ	1	ɐɔ	3	o:a	2
œʏ	4	ɔ:ə	2	eə	1	ɑʊ	3	ɔɛ	2
i:ɛ	4	əɑ:	1	i:ɑ	1	u:ɪ	3	oe:	2
ɐ:ɪ	3	ø:ɪ	1	iɔ	1	ø:ə	2	ɔ~	2
aɐ	3	ɐ:ɪ	1	ɔ:ɐ	1	ɑ:ɐ	2	ɔʏ	2
ɑɪ	3	a:ə	1	o:ɑ	1	ae:	2	uɔ	2
aʏ	3	aɔ	1	oɑ	1	ɪ:ə	2	yɪ	2

Table A3. *Cont.*

LG-V	LG-Freq	WCG-V	WCG-Freq	ECG-V	ECG-Freq	WUG-V	WUG-Freq	EUG-V	EUG-Freq
ɔ:ɐ	3	ɛ:ə	1	ʊ:ə	1	iɔ	2	øɐ	1
o:a	3	e:ɑ	1	uə	1	ɔ:ə	2	ɐ:ə	1
ʊɔ	3	ɪɛ	1	ʊə	1	uɔ	2	ɐ:ʊ	1
ʏə	3	ɪe:	1	uɑ	1	y:ɐ	2	ɐə	1
ø:a	2	o:ɑ	1	ʊɪ	1	ʏɐ	2	œə	1
øə	2	oə	1	uɔ	1	yɛ	2	œɐ	1
øɐ	2	oa	1	y:ɑ	1	øɐ	1	ɑ~	1
øʊ	2	oɑ	1	y:ɔ	1	ɐ:ɪ	1	a~	1
øʏ	2	oɪ	1	ʏə	1	œə	1	ɛ:ə	1
œ:ɪ	2	ɔɪ:	1			œʏ	1	e:a	1
a:ɐ	2	ɔʏ	1			a:ə	1	ɛɔ	1
ɑ:ɐ	2	ʊ:ɐ	1			ɑ:ə	1	i:ə	1
e:a	2	ua	1			ɑ:a	1	i:a	1
e:ɛ	2	uɛ	1			aə	1	ɔə	1
ɪ:ə	2	yə	1			ɑɐ	1	oa	1
i:a	2					a~	1	uə	1
i:ɑ	2					aʏ	1	ue:	1
iɛ	2					eˑã	1		
iʊ	2					ea	1		
ɔʏ	2					ɛa	1		
uɛ	2					eɪ:	1		
uɔ	2					o:a	1		
ø:ɪ	1					ʊ:ə	1		
ø:ʊ	1					ua	1		
aə	1					uɑ	1		
ɑɐ	1					ʊɑ	1		
ɑʊ	1								
e:ɐ:	1								
ɛa	1								
ɛɔ	1								
ɪɛ	1								
iɔ	1								
ua	1								
ʊa	1								
ʊɑ	1								
uɪ	1								
y:ø	1								
y:ɔ	1								
yɛ	1								
yɪ	1								

MHG vowels:

MHG a: acht (aht); Affe (affe); alte (alde); alter (alder); bald (balde); bestellt (bestalt); dann (danne); darf (darf); darfst (darft); fängt (vant); fahren (varn); Farbe (varwe); fest (vaste); Flasche (vlasche); ganz (ganz); ganze (ganze); ganzes (ganzes); Garten (garte); gebrannt (gebrant); gefahren (gevarn); gefallen (gevallen); gekannt (gekant); gemacht (gemacht); gerade (gerate); gesagt (gesagt); hacken (hacken); halbes (halpes); kalte (kalte); kaltes (kaltes); kam (quam); kann (kan); kannst (kanst); krank (kranc); lagen (lagen); Mädchen (maidel); machen (machen); macht (macht); Mann (man); Nacht (naht); Sack (sak); sag (sage); sagen (sagen); sah (sach); Salz (salz); schlage (slage); schwarz (swarz); stark (starc); wächst (wahset); wachsen (wahsen); war (was); warme (warme); warmes (warm); warten (warten); wartet (wartet); Wasser (waʒʒer).

MHG â**: Abend (**â**bent); abends (**â**bendes); da (d**â**); gebracht (gebr**â**ht); geh (g**â**); gehabt (geh**â**t); gehe (g**â**); gehen (g**â**n); gehst (g**â**st); geht (g**â**t); geschlafen (gesl**â**fen);**

getan (getân); habe (hân); haben (hân); habt (hât); Haken (hâke); hast (hâst); hat (hât); hatten (hâten); kamen (quâmen); Schafe (schâpe); taten (tâten); waren (wâren); warst (vârest); wart (wâret).

MHG e: artig (ertec); Berg (berc); Berge (berge); bestellt (bestelt); Bett (bette); Blätter (bleter); essen (ezzen); Feld (velt); fertig (vertic); Gänse (gens); Geld (gelt); gelegen (gelegen); gelernt (gelernt); gerne (gern); gesehen (gesehen); gestern (gester); gewesen (gewesen); heben (heben); Herz (herz); Herzen (herzen); möchte (mehte); möchtet (mehtet); Pferd (phert); recht (reht); reden (reden); schlecht (sleht); schlechte (slehte); schlechten (slehten); schnell (snel); sechs (sehs); selbst (selbest); sprechen (spehten); wenden (wenden); werden (werden); Wetter (wetter); zwölf (zwelf).

MHG ê: erst (êrste); Schnee (snê); sehr (sêr); stehen (stên); steht (stêt); zwei (zwêne).

MHG i: beißt (bit); bin (bin); bist (bist); bleib (blip); bleibt (blip); brennt (brinnt); bringe (bringe); Geschichte (geschiht); hinten (hinden); Kind (kint); Kinder (kint); Kirche (kirche); Milch (milch); Rippe (rippe); schlimm (slimp); sieht (siht); sind (sint); sitzen (sitzen); sitzt (sitzet); Tisch (tisch); trinken (trinken); werde (wirde); wieder (wider); Wiese (wise); will (wil); willst (wilt); wird (wirt).

MHG î: beißen (bîzen); bleiben (blîben); drei (drî); Eis (îs); geblieben (geblîben); gleich (glîche); hier (hîr); schneien (snîwen); schreien (schrîn); treiben (drîven); weiße (vîȝe); Wein (wîn); Zeit (zît); Zeiten (zîten).

MHG o: doch (doch); Dorf (dorf); dort (dort); gedroschen (gedroschen); gekommen (gekomen); gestohlen (gestolen); gestorben (gestorben); geworden (geworden); Holz (holz); kochen (kochen); kocht (kocht); Kohlen (kolen); komm (kom); komme (kome); kommen (komen); kommt (komt); Korb (korp); Korn (korn); morgen (morgen); noch (noch); soll (sol); sollen (scholn); sollst (sol); sollte (solde); Vögel (vogele); Wochen (wochen); wollt (wolt); wollte (wolte); wollten (wolten); worden (worden); Wort (wort).

MHG ô: böse (bôse); bösen (bôsen); Brot (brôt); Bruder (brôder); groß (grôz); höher (hôher); hoch (hôch); Ohren (ôren); rot (rôt); roten (rôten); schon (schôn).

MHG u: Bürste (burste); dürft (durft); Durst (durst); gefunden (gevunden); Hund (hunt); jungen (jungen); können (kunnen); könnt (kunnt); Luft (luft); Pfund (punt); sonst (sus); trockenen (truckenen); Wurst (wurst); zurück (zerucke).

MHG û: bauen (bûwen); Bauer (bûr); Bauern (bûre); brauche (brûche); braune (brûne); gebaut (gebûwet); Haus (hûs); Hause (hûse); laut (lût); lauter (lûter); Mauer (mûr); nun (nû).

MHG æ: hätte (hæte); hätten (hæten); hättest (hætest); Käse (kæse); mähen (mæn); nähen (næjen); täte (tæte); täten (tæten); tat (tæt); wäre (wære); wärst (wærest).

MHG œ: gehören (gehœren); größer (grœzer); hört (hœrt); schön (schœn); schöne (schœne); schönen (schœnen).

MHG ü: dünne (dünne); fünf (vünf); hervor (hervür); Stück (stück); stünde (stünde); Tür (türe); würde (würde); würden (würden).

MHG üe: Füße (vüeȝe); Kühe (küeje); müde (müede); müssen (müeȝen); müsst (müeȝt).

MHG iu: Feuer (viur); Häuser (hiuser); heute (hiute); Leute (liute); Leuten (liute); neue (niuwe); neun (niun).

MHG ei: Ei (ei); Eier (eier); eigenes (eigen); ein (ein); Fleisch (vleisch); heiß (heiȝ); heiße (heiȝe); Kleid (kleit); Kleider (kleider); klein (klein); kleine (kleine); kleinen (kleinen); Leiter (leiter); Meiste (meiste); sagt (seit); sagte (seite); Seife (seife); weiß (weiȝ); weißt (weiȝt).

MHG ie: fliegen (vliegen); fliegt (vliegt); ging (gienc); ginge (gienge); lieb (liep); liebes (liebes); nie (nie); Tier (tier); vier (vier).

MHG ou: auch (ouch); Bäume (boume); Baum (boum); Frau (vrouwe); glaube (gloube); hauen (houwen).

MHG uo: Brüder (bruoder); füttern (vuoren); früh (vruo); Fuß (vuoz); genug (genuoc); gut (guot); gute (guote); gutes (guotes); Kuchen (kuoche); Kuh (kuo); muss (muoʒ); musst (muoʒt); Mutter (muoter); suchen (suochen); tu (tuo); tun (tuon); tut (tuot).

Notes

1. If the reader is interested to discover even more such examples, they can browse the website of the PhonD2-Project. Under https://www.dsa.info/PhonD2/lemma.html (accessed: 17 July 2024), there is a list of lemmata and by clicking any of them, a map of the CV structures opens. The actual words can also be accessed by the search mask below the map. In Figure A10 in the appendix, the map for *gesagt* is given. On the website, this map can be found under https://www.dsa.info/PhonD2/0aakagi42.html (Last accessed: 17 July 2024).
2. The 40 Wenker sentences are a standard instrument of German dialectology (Chambers and Trudgill 1998, p. 15) and trace back to the dialect survey of Georg Wenker started in 1876 (see Wenker 2013).
3. Consonants: quality, coronalization, retroflex *r*, word-internal glottal stops; vowels: quality, quantity, nasality, Central Franconian tone accents.
4. **Low German:** North Saxon, Eastphalian, Westphalian, Mecklenburgisch-Vorpommersch, Middle-Pommeranian and Brandenburgian; **West Central German:** Rhine-Franconian, Central Hessian, East Hessian, North Hessian, Mozelle Franconian, Ripuarian and Low Franconian; **East Central German:** Thuringian, Upper Saxon, North Upper Saxon, South Markgravian and Schlesian; **West Upper German:** High Alemannic, Central Alemannic, Low Alemannic and Swabian; **East Upper German:** Central Bavarian, North Bavarian and East Franconian.
5. Since going into detail about the consonant clusters would exceed a single paper, a separate paper discussing them is currently under preparation.
6. This sound change process is also described in the *Kollmersches Gesetz*, according to which Old High German (OHG) /a/ has been raised to /o/ if the next syllable contained /o/ as in OHG kasto, which became Central Bavarian /khɔstn/ (Kasten "box"; see Schikowski 2009, pp. 9, 120).
7. Figure 9 was created with the function *heatmap.2* from the R package *gplots 3.1.3* (Warnes et al. 2022). Hierarchical clustering was performed with complete-linkage clustering and distances between the five areas were calculated with Euclidean distance.

References

Alber, Birgit, and Marta Meneguzzo. 2016. Germanic and Romance onset clusters—How to account for microvariation. *Theoretical Approaches to Linguistic Variation* 234: 25–51. [CrossRef]

Auer, Peter. 1991. Zur More in der Phonologie. *Zeitschrift für Sprachwissenschaft* 10: 3–36. [CrossRef]

Beedham, Christopher. 1995. Vowel + consonant and consonant + vowel sequences in the strong verbs of German and English. *Cahiers Ferdinand de Saussure* 49: 139–63.

Caro Reina, Javier. 2019. *Central Catalan and Swabian A Study in the Framework of the Typology of Syllable and Word Languages*. Berlin: Walter de Gruyter. ISBN 9783110555387.

Chambers, John K., and Peter Trudgill. 1998. *Dialectology*, 2nd ed. Cambridge: Cambridge University Press. ISBN 9780511805103.

Delattre, Pierre. 1964. Comparing the vocalic features of English, German, Spanish and French. *International Review of Applied Linguistics in Language Teaching* 2: 71–98. [CrossRef]

Dressler, Wolfgang U., and Alona Kononenko. 2021. German phonotactic vs. morphonotactic obstruent clusters: A corpus linguistic analysis. In *Experimental, Acquisitional and Corpus Linguistic Approaches to the Study of Morphonotactics*. Edited by Wolfgang U. Dressler, Basilio Calderone, Sabine Sommer-Lolei and Katharina Korecky-Kröll. Vienna: Austrian Academy of Sciences, pp. 15–51. ISBN 978-3-7001-8714-1.

Dressler, Wolfgang U., and Katarzyna Dziubalska-Kolaczyk. 2006. Proposing Morphonotactis. *Italian Journal of Linguistics* 18: 249–66.

Durrell, Martin, and Winifred V. Davies. 1990. Hessian. In *The Dialects of Modern German: A Linguistic Survey*. Edited by Charles Russ. London: Routledge, pp. 210–40. ISBN 9780415861809.

Elmentaler, Michael, and Anja Voeste. 2019. Areale Variation im Deutschen historisch: Mittelalter und Frühe Neuzeit. In *Sprache und Raum. Ein Internationales Handbuch der Sprachvariation. Band 4: Deutsch*. Edited by Joachim Herrgen and Jürgen Erich Schmidt. Berlin: De Gruyter Mouton, pp. 61–100. ISBN 978-3-11-018003-9.

Goltz, Reinhard H., and Alastair G. H. Walker. 1990. North Saxon. In *The Dialects of Modern German: A Linguistic Survey*. Edited by Joachim Herrgen and Jürgen Erich Schmidt. London: Routledge, pp. 31–58. ISBN 9780415861809.

Göschel, Joachim. 1992. *Das Forschungsinstitut für Deutsche Sprache "Deutscher Sprachatlas", Wissenschaftlicher Bericht*. Marburg: Forschungsinstitut für Deutsche Sprache.

Göschel, Joachim. 2000. Der Phonetische Atlas von Deutschland. *Јужнословенски филолог* 56: 283–88.

Gregová, Renáta. 2010. A comparative analysis of consonant clusters in English and in Slovak. *Bulletin of the Transilvania University of Braşov* 3: 79–84.

Hall, T. Alan. 2000. *Phonologie: Eine Einführung*. Berlin: Walter de Gruyter. ISBN 9783110215878.

Harnisch, Rüdiger. 1987. *Natürliche generative Morphologie und Phonologie des Dialekts von Ludwigsstadt*. Tübingen: Max Niemeyer Verlag. ISBN 9783484301900.

Hayes, Bruce. 1989. Compensatory Lengthening in Moraic Phonology. *Linguistic Inquiry* 20: 253–306.

Herrgen, Joachim, and Jürgen Erich Schmidt. 2019. *Sprache und Raum. Ein internationales Handbuch der Sprachvariation. Band 4: Deutsch*. Berlin: De Gruyter Mouton. ISBN 978-3-11-018003-9.

Knöbl, Ralf, and Madlen Nimz. 2014. Sprachräumliche Aspekte des Gebrauchs der deverbalen Modalpartikel glaub(e) zur Modulierung des Geltungsanspruchs von Äußerungen. *Социопрагматическое и семантическое пространство языка*, 91–108.

Koch, Günter. 2019. Bairisch in Deutschland. In *Sprache und Raum. Ein internationales Handbuch der Sprachvariation. Band 4: Deutsch*. Edited by Joachim Herrgen and Jürgen Erich Schmidt. Berlin: De Gruyter Mouton, pp. 279–318. ISBN 978-3-11-018003-9.

Lameli, Alfred. 2022. Syllable structure spatially distributed: Patterns of monosyllables in German Dialects. *Journal of Germanic Linguistics* 34: 241–87. [CrossRef]

Lameli, Alfred, Alexander Werth, Valeria Bunkov, and Samantha Link. 2023. PhonD2: A database on phonotactic structures of German dialects. Paper presented at the 20th International Congress of Phonetic Sciences, Prague, Czech Republic, August 7–11; pp. 3201–5.

Lameli, Alfred, and Samantha Link. Forthcoming. Correspondence of consonant clustering with vowel types in German dialects. In *Dialekte im sozialen Raum. Formen-Verwendungen-Bedeutungen*. Edited by Stephan Elspaß, Jürg Fleischer and Marlene Hartinger. Stuttgart: Steiner-Verlag.

Lasch, Agathe. 1974. *Mittelniederdeutsche Grammatik*, 2nd ed. Tübingen: Max Niemeyer Verlag. ISBN 3-484-10183-0.

Liberman, Anatoly. 1997. The reaction of monosyllables to apocope in German dialects. In *Insights in Germanic Linguistics. II: Classic and Contemporary*. Edited by Irmengard Rauch and Gerald F. Carr. Berlin: Gruyter Mouton, pp. 99–133. ISBN 9783110148541.

Maddieson, Ian. 2013. Vowel Quality Inventories. *The World Atlas of Language Structures Online (WALS)*. [CrossRef]

Menzerath, Paul. 1954. *Die Architektonik des Deutschen Wortschatzes*. Bonn: F. Dümmler.

Niekerken, Walther. 1963. Von den Formen und Wirkungen der Liquida *r* im Nordsächsischen. *Language Typology and Universals (STUF)* 16: 165–76. [CrossRef]

Pätzold, Matthias, and Adrian P. Simpson. 1997. Acoustic analysis of German vowels in the Kiel Corpus of Read Speech. *Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung Universität Kiel* 32: 215–47.

Paul, Hermann. 2007. *Mittelhochdeutsche Grammatik*. Berlin: De Gruyter Mouton. ISBN 9783484640344.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online: https://www.R-project.org/ (accessed on 17 July 2024).

Rowley, Anthony R. 1997. *Morphologische Systeme der Nordostbayerischen Mundarten in ihrer Sprachgeographischen Verflechtung*. Stuttgart: Franz Steiner Verlag. ISBN 978-3-515-06700-3.

Russ, Charles V. J., ed. 1990a. High Alemannic. In *The Dialects of Modern German: A Linguistic Survey*. London: Routledge, pp. 364–93. ISBN 9780415861809.

Russ, Charles V. J., ed. 1990b. Swabian. In *The Dialects of Modern German: A Linguistic Survey*. London: Routledge, pp. 337–63. ISBN 9780415861809.

Russ, Charles V. J., ed. 1990c. *The Dialects of Modern German: A Linguistic Survey*. London: Routledge. ISBN 9780415861809.

Ryan, Kevin M. 2016. Phonological Weight. *Language and Linguistics Compass* 10: 720–33. [CrossRef]

Schikowski, Robert. 2009. *Die Phonologie des Westmittelbairischen*. München: Ludwig Maximilians Universität München. [CrossRef]

Schmidt, Jürgen Erich. 1986. *Die mittelfränkischen Tonakzente (Rheinische Akzentuierung)*. Stuttgart: Franz Steiner Verlag. ISBN 3-515-04803-0.

Seiler, Guido, and Kathrin Würth. 2014. Monosyllabic Lengthening in German and its relation to the syllable vs. word language typology. In *Syllable and Word Languages*. Edited by Javier Caro Reina and Renata Szczepaniak. Berlin: Walter de Gruyter, pp. 140–59. ISBN 9783110343458.

Selkirk, Elisabeth. 1984. On the major class features and syllable theory. In *Language Sound Structure*. Edited by Mark Aronoff, Richard Oehrle, Frances Kelley and Bonnie Wilker Stephens. Cambridge: The MIT Press, pp. 107–36. ISBN 9780262511742.

Siebenhaar, Beat. 2019. Ostmitteldeutsch: Thüringisch und Obersachsisch. In *Sprache und Raum. Ein internationales Handbuch der Sprachvariation. Band 4: Deutsch*. Edited by Joachim Herrgen and Jürgen Erich Schmidt. Berlin: De Gruyter Mouton, pp. 407–35. ISBN 978-3-11-018003-9.

Spangenberg, Karl. 1990. Thuringian. In *The Dialects of Modern German: A Linguistic Survey*. Edited by Charles Russ. London: Routledge, pp. 265–89. ISBN 9780415861809.

Streck, Tobias. 2019. Alemannisch in Deutschland. In *Sprache und Raum. Ein Internationales Handbuch der Sprachvariation. Band 4: Deutsch*. Edited by Joachim Herrgen and Jürgen Erich Schmidt. Berlin: De Gruyter Mouton, pp. 206–45. ISBN 978-3-11-018003-9.

Szczepaniak, Renata. 2007. *Der phonologisch-typologische Wandel des Deutschen von einer Silben-zu einer Wortsprache*. Berlin: Walter de Gruyter. ISBN 9783110192742.

Trubetskoy, Nikolaï Sergeyevich. 1967. *Grundzüge der Phonologie (1939)*. Göttingen: Vandenhoeck & Ruprecht.

Warnes, Gregory R., Ben Bolker, Lodewijk Bonebakker, Robert Gentleman, Wolfgang Huber, Andy Liaw, Thomas Lumley, Martin Maechler, Arni Magnusson, Steffen Moeller, and et al. 2022. gplots: Various R Programming Tools for Plotting Data. R Package Version 3.1.3. Available online: https://CRAN.R-project.org/package=gplots (accessed on 15 May 2024).

Wenker, Georg. 2013. *Schriften zum "Sprachatlas des deutschen Reichs". Gesamtausgabe*. 3 Bände. Edited by von Alfred Lameli. Hildesheim: Georg Olms Verlag.

Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer.

Wiesinger, Peter. 1983a. Die Einteilung der deutschen Dialekte. In *Dialektologie. Ein Handbuch zur deutschen und allgemeinen Dialektforschung. 2. Halbband*. Edited by Peter Wiesinger, Werner Besch, Ulrich Knoop, Wolfgang Putschke and Herbert Ernst Wiegand. Berlin: Walter de Gruyter, pp. 807–900. ISBN 9783110095715.

Wiesinger, Peter. 1983b. Phonologische Vokalsysteme deutscher Dialekte: Ein synchronischer und diachronischer Überblick. In *Dialektologie. Ein Handbuch zur Deutschen und Allgemeinen Dialektforschung. 2. Halbband*. Edited by Peter Wiesinger, Werner Besch, Ulrich Knoop, Wolfgang Putschke and Herbert Ernst Wiegand. Berlin: Walter de Gruyter, pp. 1042–76. ISBN 9783110095715.

Wiesinger, Peter. 1990. The Central and Southern Bavarian Dialects in Bavaria and Austria. In *The Dialects of Modern German: A Linguistic Survey*. Edited by Charles Russ. London: Routledge, pp. 438–519. ISBN 9780415861809.