

Article

Lexical Variation and Change in the Bavarian, Franconian and Swabian Dialects of Bavaria: A Combined Real- and Apparent-Time Analysis

Lars Bülow ^{1,*} , Philip C. Vergeiner ¹ and Oliver Schallert ² ¹ Department of German Studies, Ludwig-Maximilians-University of Munich, 80799 Munich, Germany² Department of German and Comparative Studies, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91054 Erlangen, Germany

* Correspondence: lars.buelow@germanistik.uni-muenchen.de

Abstract

This study provides the first systematic combined real- and apparent-time analysis of onomasiological lexical variation and change in the Bavarian, Franconian, and Swabian dialects of Bavaria (Germany). Drawing on two large-scale indirect surveys—the Maurer questionnaires (collected in 1934), and the DaBay dialect app (2025–ongoing)—we examine five lexical concepts (ANT, APPLE CORE, ROOSTER, POTATO, and LADYBUG) using descriptive mapping and Generalised Additive Mixed Models (GAMMs). The results reveal three main patterns: (i) stability in long-established heteronyms (in APPLE CORE), (ii) real-time convergence toward Standard German (in ANT, POTATO, and LADYBUG), and (iii) concept-specific areal realignments driven by competing dialect-internal and standard-influenced variants (in ROOSTER). Dialect geography consistently conditions lexical choices, while social factors play a comparatively minor role. Together, the findings demonstrate that lexical variation and change in Bavaria is systematic, regionally differentiated, and empirically tractable through large-scale historical and contemporary survey data.

Keywords: lexical variation and change; Bavarian; Swabian; Franconian; Maurer survey; dialect app

1. Introduction

Adopting a combined variationist real- and apparent-time approach, this exploratory study investigates lexical variation and change as an expression of social, cognitive, and geographical factors in the dialects of Bavaria (Germany). More specifically, it focuses on so-called formal onomasiological variation in lexis, that is, on the choices “of one word over another that expresses the same concept” (Geeraerts, 2010, p. 823). Although many studies and dialect atlas projects have examined lexical variation both within and across dialects from a synchronic point of view, the investigation of onomasiological lexical change itself still represents a clear gap in dialectal lexical research (cf. Stöckle et al., in press).

We compare data from the first half of the 20th century, namely the so-called *Maurer questionnaires* from Bavaria, with data from the first half of the 21st century, drawn from the *Dialektapp Bayern*¹ (DaBay, ‘Dialect App Bavaria’). The Maurer questionnaires from Bavaria were gathered in 1934 and are being used here for the first time for a quantitative study on lexical variation and change (see for the Alemannic Maurer questionnaires, e.g., Lameli & Schönberg, 2023; Auer & Strobel, 2025a, 2025b). Maurer’s survey, methodologically based on the Wenker questionnaires for the *Sprachatlas des Deutschen Reiches* (=WA, 1888–1923,



Academic Editor: Alexandra Lenz

Received: 10 February 2026

Revised: 11 May 2026

Accepted: 1 July 2026

Published: 16 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

‘Language Atlas of the German Empire’) (cf. [Fleischer, 2017](#)), was an indirect elicitation of 90 lexical items using a written questionnaire distributed to over 5000 locations in Bavaria (see Section 3.1). The DaBay dialect app, launched in May 2025, collects data from dialect speakers throughout Bavaria by means of a progressive web app. The DaBay dialect app enables an indirect collection of linguistic data across all levels of analysis and represents the modern dialects in different regions of Bavaria (see Section 3.2). The two datasets are comparable in that both were collected indirectly across the entire territory of the Federal State of Bavaria (Germany) and share a subset of lexical items. A comparison of these two datasets is timely and innovative because even though relatively good data on the lexicon of the various dialects in Bavaria are available (e.g., [Gütter, 1971](#); [Fritz-Scheuplein et al., 2009](#)), little is known about change at the level of the lexicon, especially in comparison to sound change. Thus, for the dialects in Bavaria there are very few studies that empirically examine lexical variation and lexical change using variationist methods (but see [Pickl, 2013](#) for Bavarian Swabia).

On the one hand, this neglect of lexical variation and change in variationist research on Bavarian is surprising, as German dialectology—since its earliest stages, including various dialect dictionary and atlas projects (e.g., *Bayerisches Wörterbuch* (=BWB, 1985–) ‘Bavarian Dictionary’; *Sprachatlas von Bayerisch-Schwaben* (=SBS, 1996–2009) ‘Language Atlas of Bavarian Swabia’; *Sprachatlas von Niederbayern* (=SNiB, 2003–2010) ‘Language Atlas of Lower Bavaria’)—has traditionally focused on the lexicon, also in Bavaria (see Section 2). On the other hand, early qualitative analyses, for instance, of data from the *Deutscher Wortatlas* (=DWA, 1951–1980, ‘German Word Atlas’), collected between 1939 and 1942, revealed that lexical distributions and phonological isoglosses correspond only to a very limited extent (e.g., [Hildebrandt, 1983](#)). These phonological isoglosses, however, continue to serve as the basis for defining the areal structure of the German dialect landscape. At least for the DWA data, there are no major correlations. What becomes apparent instead is that each lexical item not only has its own history but also its own specific geographical distribution. Consequently, it has been argued that the lexicon itself does not reflect dialect regions; rather, individual lexical variants can merely be assigned to them ([Hildebrandt, 1983](#), p. 1332).

Even though we have a rich source of data on the dialects of Bavaria in the field of lexicography, it is striking that lexical variation and change has not yet been systematically studied from a variationist perspective over time, even though studies on phonetics-phonology, morphology, and syntax suggest that the structure of the Bavarian, Franconian and Swabian dialects in Bavaria is significantly changing (cf. overviews in [Harnisch, 2019](#); [G. Koch, 2019](#); [Streck, 2019](#)). Previous research on variation and change on these levels of the linguistic architecture has demonstrated that convergence toward the standard, shifting dialect boundaries, and the relevance of social factors can be observed in these regions. The present explorative study extends this line of inquiry to the lexical level by examining whether comparable patterns of variation and change can be observed over the past hundred years. To this end, it adopts a combined real- and apparent-time approach, addressing a notable gap in the literature, as systematic comparisons between historical and contemporary lexical datasets are largely missing.

Against this background, the study compares the Maurer data with newly collected data from 2025 via the DaBay dialect app and addresses the following open research questions:

- Can we identify language change in real-time? If so, which patterns of ongoing change emerge? For instance, which variants of lexical variables have been lost, remained stable, or even expanded in use? Are there new variants emerging, e.g., through borrowing from other varieties such as the standard language?

- How has geographical variation in lexis evolved over the past century? Do the patterns of variation and change differ across and within the Bavarian, Franconian, and Swabian dialect regions?
- Is lexical change influenced by social factors such as speakers' age or gender?

To address these questions, the analysis is restricted to lexical concepts that were elicited in both datasets, thereby ensuring comparability across the two time periods. Specifically, we examine five concepts (ANT, APPLE CORE, ROOSTER, POTATO, and LADY-BUG) that are covered by both Maurer's survey (conducted in 1934) and the DaBay dialect app (2025–ongoing). Building on previous research for the levels of phonetics/phonology, morphology, and syntax (e.g., Bülow, 2019; Bülow et al., 2019; Vergeiner & Bülow, *in press*), we hypothesise a general shift towards more regiolectal or standard-language lexical forms and a loss of locally restricted forms (cf. J. E. Schmidt & Herrgen, 2011, p. 212 for these dynamics). The present study therefore tests whether a similar tendency can also be identified in lexical variation and change, both in real-time and in apparent-time. Furthermore, the analysis is guided by general principles of variationist sociolinguistics, which assume that linguistic variation is probabilistically structured and systematically conditioned by both social and geographical factors (e.g., Labov, 1972; Chambers & Trudgill, 1998). Accordingly, any observed patterns are interpreted in relation to established variationist constraints such as dialect region, age, and gender. The resulting dataset, comprising between 2358 (for POTATO) and 3087 (for APPLE CORE) responses per item (see Section 3), will be examined using Generalised Additive Mixed Models (GAMMs) to model social and geographical constraints (cf. Wirtz et al., 2025) and provide a detailed account of lexical variation and change in real- and apparent-time.

In the following, Section 2 provides a rough overview of the current state of research on lexical variation and change in the dialects of Bavaria. Section 3 then outlines the research design and methodological approach in detail and also introduces the Maurer survey alongside the DaBay survey. Section 4 presents the empirical findings, which are discussed and succinctly summarised in Section 5.

2. Previous Research

Lexical change is a complex and multi-layered process in which the form and meaning aspects of a word interact with each other. Much has already been written on this relationship (cf., e.g., Geeraerts, 2010; Pickl, 2013; Lötscher, 2017). Following Lötscher (2017, p. 45), lexical innovations usually originate in onomasiological change (see also P. Koch & Oesterreicher, 1996, p. 77). In such processes, speakers introduce innovative expressions or variants to refer to existing concepts. These innovations can become heteronyms through conventionalisation and entrenchment processes (cf. H.-J. Schmid, 2020). We understand a concept in a broad cognitive-semantic sense, defined by Murphy (2002, p. 270) "as the mental representation of classes of objects and entities in the world, which are used in thinking about those entities".

In the following section, we outline dialect-lexical research in Bavaria (Section 2.1) and identify areal patterns of lexical variation (Section 2.2). We then discuss the methodological challenges associated with traditional sources such as dialect dictionaries and dialect atlas data (Section 2.3) before summarising central mechanisms of lexical variation and change (Section 2.4).

2.1. Foundations

Since its beginnings in the first half of the 19th century, dialectology has aimed to document dialect lexis. Early efforts are reflected in the compilation of so-called idiotika and dialect dictionaries (e.g., von Sonnleithner, 1824; von Schmid, 1831; Schmeller,

1827–1837/1872–1877; see also Löttscher, 2019, pp. 680–681). However, under the strong influence of the Neogrammarian tradition, the field shifted markedly in the latter half of the nineteenth century toward the investigation of phonetic–phonological variation and change. This research focus was institutionalised through the Wenker survey (1876–1887) underlying the *Sprachatlas des Deutschen Reiches* (=WA, ‘Language Atlas of the German Empire’), which generated large-scale written dialect data from more than 50,000 school locations (cf. Fleischer, 2017, p. 149). As a result, lexical variation remained comparatively underexplored (cf. Wiesinger, 1988, p. 601).

Once the potential of the indirect methodology became evident, other dialectologists adapted it in the first half of the 20th century to investigate lexical variation. Notable examples include the *Deutscher Wortatlas* (=DWA, ‘German Word Atlas’, 1939–1942, cf. Mitzka, 1938, 1939) and Friedrich Maurer’s surveys in 1934 (in Bavaria) and 1941 (in Baden-Württemberg and Alsace) (see Section 3.1). Although Maurer’s surveys achieved broad regional coverage, they received far less scholarly attention. While the DWA was published in 22 volumes, Maurer’s survey has been almost completely forgotten, a point to which we return in Section 3.1.

The DWA subsequently became the empirical basis for numerous studies on the areal distribution of dialect lexis, focusing on the heteronymy of individual concepts (e.g., *Brennnessel* ‘nettle’, *Hagebutte* ‘rose hip’) and thematic fields (kinship, diseases, and season names) (cf. Hoffmann, 1956; Debus, 1958; Tallen, 1963). These studies consistently demonstrate that lexical data do not delineate dialect regions as systematically as phonetic–phonological features. It was therefore argued that lexical variables and their variants (heteronyms) can only be used in a very limited and unsystematic way for dialect-geographical studies (cf. Löttscher, 2019, p. 694).

2.2. Word Geography and Areal Structures in the Dialects of Bavaria

Despite the limited suitability of lexical data for delineating dialect areas, a number of studies have attempted to identify areal structures on the basis of lexical variation (cf. Kranzmayer, 1960; Wiesinger, 1988). However, as Wiesinger (1988, p. 605) demonstrates for the Bavarian dialects, the lexicon does not systematically reflect the phonetic–phonological north–south division into North, Central, and South Bavarian. This is particularly noteworthy given that such phonetic–phonological isoglosses constitute the primary basis for the traditional classification of German dialects (Wiesinger, 1983). Crucially, this lack of alignment is not confined to the lexicon: other levels of linguistic architecture, including morphology and syntax, likewise fail to correspond systematically to these boundaries (cf. Vergeiner, 2023; Vergeiner et al., 2025). Comparable mismatches have also been observed in other dialect areas, including the Swabian and East Franconian dialect regions (cf. Harnisch, 2019; Streck, 2019).

At the same time, the absence of clear correspondence does not imply that lexical variation is spatially unstructured. Rather, lexical variants tend to form diffuse and often overlapping areal patterns. Instead of sharp isoglosses, transitional zones are frequently observed, in which multiple heteronyms coexist. For instance, Wiesinger (1988) identifies a transition area between Central and North Bavarian characterised by extensive overlap of lexical variants, while Klausmann (2014, pp. 81–87) describes similar patterns in the Bavarian–Swabian contact zone, particularly in the Lechrain region. These findings suggest that lexical geography is best understood not in terms of discrete dialect boundaries but as a gradient and concept-dependent system.

Further insights into the structuring of lexical variation are provided by regional dialect atlas projects. Drawing on data from the *Sprachatlas von Bayerisch-Schwaben* (=SBS, ‘Linguistic Atlas of Bavarian Swabia’), Pickl (2013) shows that the distribution of lexical variants is

shaped not only by geographical factors but also by semantic–pragmatic properties such as referential stability and expressivity. As a consequence, many lexical items—especially those referring to plants and animals—exhibit highly localised and fragmented patterns, often without clearly identifiable boundaries. This observation underscores the importance of concept-specific developments in lexical variation and change.

Finally, existing studies show that the degree of lexical diversity varies considerably across regions. While Central Bavarian dialects are often relatively homogeneous, typically displaying only a small number of competing variants per concept, the East Franconian and Alemannic-Swabian areas tend to show substantially higher degrees of heteronymic diversity (cf. Wiesinger, 1988, p. 702; Pickl, 2013). These differences have been attributed to a range of factors, including geographical fragmentation, historical settlement patterns, and varying degrees of socio-political integration. Taken together, these findings highlight that lexical variation in the dialects of Bavaria is spatially structured, but for many lexical concepts in a different way than phonetic–phonological variation.

2.3. Challenges and Limitations of Traditional Dialect Lexis Documentation

Traditional data sources of dialect lexis—most notably dialect dictionaries and regional dialect atlas projects—provide valuable empirical material but pose significant methodological challenges for the study of lexical variation and change (cf. Stöckle & Wahl, 2025). These challenges concern both the spatial representativeness of the data and their suitability for variationist and diachronic analyses.

Dialect dictionaries are typically organised according to political rather than linguistic boundaries, resulting in uneven regional coverage. Moreover, they document lexical items without systematically mapping their spatial distribution, which makes it difficult to reconstruct patterns of heteronymy or identify broader areal structures (Lötscher, 2019, p. 680).² Their usefulness for large-scale analyses is further limited by the fact that so far, many major projects remain incomplete, preventing comprehensive and directly comparable coverage across dialect regions (cf. the map in König et al., 2019, p. 138).

Regional dialect atlas projects, while providing spatially explicit data, are likewise often structured along administrative or political units, which may not coincide with linguistic boundaries. In addition, traditional atlas methodology usually records only a single lexical item per concept from one informant at each location (Geeraerts, 2010, p. 825), thereby presupposing internal homogeneity within local speech communities. This assumption is incompatible with variationist approaches, which conceptualise linguistic variation as probabilistic and socially structured. Furthermore, atlas data are necessarily selective with respect to the lexical items investigated, raising questions about the representativeness of individual maps for broader patterns of variation. These limitations are compounded by issues of semantic ambiguity, as lexical items may differ not only in form but also in meaning, complicating cross-regional comparison (Geeraerts, 2010, p. 826).

Finally, both types of sources are limited in their ability to capture language change. Dialect dictionaries generally lack temporal depth, while traditional regional dialect atlas surveys—often based on so-called NORM or NORFs, i.e., ‘nonmobile, older, rural males/females’ (cf. Chambers & Trudgill, 1998)—provide only a snapshot of highly localised and typically conservative dialect variants. As a result, neither approach readily allows for the systematic investigation of diachronic developments or the application of variationist frameworks such as an apparent-time analysis.

Taken together, these limitations show that both dialect dictionaries and regional dialect atlas projects offer only a partial perspective on lexical variation and change. While they provide indispensable descriptive foundations, they need to be complemented by approaches that enable the systematic analysis of spatial, social, and temporal dynamics.

2.4. Lexical Variation and Change in Real-Time

The present geographical structure of the dialect lexicon in Bavaria is the result of long-term processes of variation and change. Migration movements, cultural transformations, and various forms of language contact have contributed to lexical variation that is both historically layered and geographically differentiated (cf. [Kranzmayer, 1960](#); [Wiesinger, 1988](#); [Lötscher, 2019](#)). From a diachronic perspective, lexical differentiation arises through processes of innovation, diffusion, and loss, which operate unevenly across dialect regions and lexical domains (cf. [Pickl, 2013](#)).

Lexical innovations are frequently triggered by changes in material culture, such as the introduction of new objects or practices, but may also arise from expressive or associative motivations (cf. [Lötscher, 2017](#)). These innovations can spread across regions and become conventionalised, giving rise to competing heteronyms. At the same time, processes of lexical loss reduce variation, for instance when traditional referents disappear or when more localised forms are replaced by supra-regional or standard-oriented variants. In this respect, dialect–standard contact plays a central role: numerous studies have shown that convergence toward the standard language affects all levels of the linguistic system, including the lexicon (cf. [Auer, 2018](#); [Bülow et al., 2019](#); [Lenz, 2019](#)). However, such processes do not proceed uniformly but interact with regional, social, and concept-specific factors.

Despite a relatively well-developed theoretical understanding of these mechanisms, empirical evidence on lexical change in Bavarian dialects remains limited. Most existing studies are based on synchronic data and rely on traditional sources such as dialect dictionaries and traditional regional dialect atlas projects (e.g., [Pickl, 2013](#)), which are not well suited to capturing variationist patterns or diachronic dynamics (see Section 2.3). As a result, it remains unclear to what extent the processes of innovation, diffusion, and convergence can be observed in real-time and how they are constrained by geographical and social factors.

Against this background, an exploratory study combining historical and contemporary data is necessary to provide a first variationist account of lexical change in the dialects of Bavaria and to identify the factors that may constrain it. In particular, such an approach requires the comparison of datasets with extensive supra-regional coverage, which make it possible to analyse lexical variation across the entire dialect landscape in Bavaria in a systematic and comparable way. The Maurer survey and the DaBay dialect app provide precisely this type of large-scale, indirectly elicited data, thereby offering a unique opportunity to investigate lexical change over a time span of approximately one hundred years.

In the following section, we outline the research design, data, and methods employed to carry out this comparison across five concepts across the dialects of Bavaria and to model lexical variation and change in real- and apparent-time.

3. Research Design and Method

This study combines a real-time comparison with an apparent-time comparison. In real-time, we compare data from the Maurer survey (1934) with data from the DaBay survey (2025). Both surveys consist of data from dialect speakers of different age groups (younger vs. older speakers), which means that the apparent-time heuristic can be applied to both the Maurer survey and the DaBay survey. Combined real-time and apparent-time comparisons have already been employed in dialect studies (e.g., [Streck, 2012](#); [Schwarz, 2015](#)), albeit only rarely; however, to our knowledge, they have not yet been used to investigate lexical variation and change in Bavarian dialects. A further innovation, to the best of our knowledge, is that both real-time comparison points make it possible to apply the apparent-time heuristic.

In the following, in Section 3.1, we will first explain the Maurer survey in more detail, before addressing the DaBay survey in Section 3.2. After that, we will explain our methods, items, and sample in Section 3.3. The statistical approach will be outlined in Section 3.4.

3.1. Maurer Survey

The Maurer survey takes its name from Friedrich Maurer (1898–1984), a German linguist specialising in historical linguistics and dialectology. After completing his PhD under Otto Behaghel, he became a full professor at the University of Erlangen-Nuremberg in 1931, where he oversaw the *Ostfränkisches Wörterbuch* ('East Franconian Dictionary', today WBF, 2013–), founded in 1913 as a spinoff of the *Bayerisches Wörterbuch* ('Bavarian Dictionary', today BWB). In 1937, he assumed the chair of Germanic Linguistics at Freiburg University, a position he held until his retirement in 1966. Maurer co-founded the *Institut für Deutsche Sprache* in Mannheim (now the 'Leibniz Institute for the German Language') in 1964, initiated work on the *Historischer Südwestdeutscher Sprachatlas* ('Historical South-West German Language Atlas'), the first historical linguistic atlas of German varieties (Kleiber et al., 1979a, 1979b) as well as the *Südheßisches Wörterbuch* ('South Hessian Dictionary'). Yet despite these scholarly accomplishments, Maurer remains a problematic figure: a committed German nationalist and member of the Stahlhelm Brigade in the 1920s, he soon joined the NSDAP and later the NS-Dozentenbund (Klee, 2005, p. 396). Like many lower-level NSDAP party members and opportunists, he returned to his Freiburg professorship after World War II and became a key figure in its rebuilding.

Work on the Maurer survey began in 1934, when questionnaires were distributed to all schools in the Federal State of Bavaria, covering more than 5000 locations. Maurer's adoption of the indirect survey method may have derived from his earlier engagement with the Wenker materials: in 1926 he demonstrated their suitability for syntactic analysis, an aspect previously overlooked. In 1941, during Maurer's work at Freiburg University, a slightly revised version was circulated in Baden and Alsace, reaching over 1000 locations (Strobel, 2021, p. 159). We refer to this dataset as the 'Upper Rhine survey'. Maurer drew solely on the Upper Rhine survey—among various additional sources—for his monograph *Zur Sprachgeschichte des deutschen Südwestens* (Maurer, 1942). The questionnaires of the Upper Rhine survey were rediscovered in 2017 in the archives of the Institute for Germanic Linguistics at Freiburg and subsequently employed in the DFG-funded project *Alemannisch variativ* ('Variation in Alemannic', 2021–2025, PIs Peter Auer and Alfred Lameli), which examined apparent- and real-time dialect change in the Upper Rhine Alemannic region of southwestern Germany.

As discussed by Strobel (2018, pp. 3–4), the questionnaires used in the southwestern regions differ in several respects from those distributed in Bavaria. We label the latter the 'Bavarian survey'. Crucially, the Bavarian survey predates the southwestern one—conducted in 1941—by seven years. Appended to the *Bayerischer Schulanzeiger* ('Bavarian School Gazette'), it was sent to all school locations in Bavaria and includes approximately 500,000 translations from more than 5000 sites. The questionnaire contains 90 items and relatively detailed sociodemographic queries; it was primarily intended to elicit lexical and phonological features. Five full sentences additionally permit the investigation of morphosyntactic variation.

The questionnaires were completed by men and women from different age groups (see Section 3.3). Furthermore, the majority of participants were local teachers, farmers and craftsmen. Note that the socio-demographic information was not completed in all cases. In some cases, the first name (which we use to indirectly determine gender), age or occupation are missing. Of the 1310 Maurer questionnaires analysed here, gender information is missing for 323 questionnaires, and age information for 74.

3.2. DaBay Survey

To investigate lexical variation in contemporary dialects, we draw on data from DaBay, a web application launched in 2025 to collect dialect data from speakers across Bavaria. Comparable dialect apps have been deployed in numerous regions, as crowdsourcing offers a resource-efficient, effective, and comprehensive means of gathering up-to-date linguistic data (for an overview of such projects, cf. [Leemann, 2021](#)). Although this approach entails certain limitations—such as reduced control over informant selection—app-based datasets have repeatedly proven valuable for analysing dialect variation. In particular, they are well suited for detecting social variation, as they include more diverse speaker groups than traditional (direct) dialect surveys. As [Vergeiner \(in press\)](#) argues, through this broader user base, app-derived data might offer a more realistic picture of everyday dialect use. Moreover, compared to traditional surveys, which often capture highly conservative dialect forms ([Streckenbach, 2020](#)), app-collected data enhance the ability to identify ongoing language change (cf. [Vergeiner, in press](#)).

As of December 2025, DaBay comprises more than 1150 items across various question formats, including written items (both open-ended and multiple-choice) as well as tasks requiring participants to submit audio recordings. The questions cover all linguistic levels—phonetics/phonology, morphology, syntax, lexis, and pragmatics—as well as language attitudes and perceptions. Since previous experience has shown that participants tend to be most interested in lexical items, the largest share of the questions pertains to this level. Comparisons of older data sources, such as the DWA, with preliminary app results (published within DaBay) already point to substantial lexical change over recent decades. The present study provides the first quantitative comparison between DaBay data and an older data source.

3.3. Method, Items, and Sample

In what follows, we focus on five lexical items which were chosen because they were surveyed in both the Maurer survey and the DaBay survey with similar methods.³ These lexical concepts are APPLE CORE, i.e., ‘the remains of the eaten apple that is thrown away’ (FM 5), ANT (FM 29), ROOSTER (FM 71), POTATO (FM 1), and LADYBUG (FM 28).⁴ All five concepts were queried through open-ended questions. In the Maurer survey participants were asked to translate a written standard-language stimulus, e.g., *Ameise*, into their local dialect, e.g., *Agameise*, *Ametze*, etc. (see Section 4.2). In the DaBay survey, participants were shown an image depicting the respective concept and asked to enter the corresponding term in their dialect. Of course, these different elicitation techniques may have produced some differences, as the standard-language stimulus, for example, may have encouraged some respondents to use more standard forms. However, our results suggest that this effect was not substantial. The different realisations provided by participants were then grouped into lexical variants by aggregating forms that differed only in terms of spelling and/or regular sound changes (for example, the variant *Ameise* for the concept ANT combines realisations such as *Amoasn*, *Omoas*, *Amois*, and *Omas*, which reflect different dialectal reflexes of the vowels *â* and *ei* in the corresponding Middle High German form). The final set of variants was constructed to ensure the greatest possible comparability between the two datasets. Table 1 shows the number of (valid)⁵ answers per item and per survey.

Table 1. Number of (valid) answers per item and per survey.

	APPLE CORE	ANT	ROOSTER	POTATO	LADYBUG
Maurer	1271	1269	1280	1290	1114
DaBay	1816	1561	1388	1068	1564

In total, 13,621 answers were annotated, 6224 for the Maurer survey and 7397 for the DaBay survey. The Maurer questionnaire was completed by more male ($n = 849$) than female participants ($n = 138$). Although a total of 1310 Maurer questionnaires were analysed, it was not possible to assign a gender in 323 cases (see Section 3.1). In contrast, the gender ratio in the DaBay survey is fairly balanced. Since most participants in DaBay responded to only a subset of the questions, these data come from 3161 participants, 47% male and 53% female.

While both datasets include both younger and older speakers, the majority of participants are between 40 and 70 years of age. The age structure of the two datasets is shown in Table 2. A χ^2 -test shows that there is a significant difference in proportions of age groups between the Maurer and DaBay surveys ($\chi^2 = 139.84$, $df = 7$, $p < 0.001$), which has a small-to-medium effect (Cramér's $V \sim 0.18$). In particular, there is a difference of more than 450 participants in the 51–60 and 61–70 age groups. When the data is aggregated into two groups, namely ≤ 50 and > 50 , the difference is still statistically significant ($\chi^2 = 78.84$, $df = 1$, $p < 0.001$), but with only a small effect size ($\phi \sim 0.13$).⁶ This age-group division is also used in the subsequent analysis, as it yields two groups with roughly equal numbers of speakers in both datasets. We did not create additional age groups because this would result in too few speakers per group and region, and because using two groups facilitates our statistical analysis, which requires binary (or continuous)⁷ variables.

Table 2. Age group of the participants per survey.

Age Group	Maurer Survey	DaBay Survey
≤ 20	95	148
21–30	121	385
31–40	231	442
41–50	318	509
51–60	254	705
61–70	129	658
71–80	69	267
> 80	19	47

Let us now take a look at the regional origin of the participants. To provide a first, albeit rough, overview, Table 3 shows the number of participants for the seven administrative regions (<https://en.wikipedia.org/wiki/Regierungsbezirk>, accessed on 30 November 2025) of Bavaria (<https://en.wikipedia.org/wiki/Bavaria>, accessed on 30 November 2025) (note that in Bavaria, the administrative regions at least partially overlap with dialect regions, see also Figure 1). There is a clear surplus of DaBay participants, particularly in Upper Bavaria (in particular Munich and its surrounding area), Lower Bavaria and the Upper Palatinate, meaning the number of answers is not evenly distributed geographically. Overall, the difference between the two surveys is highly significant ($\chi^2 = 707.56$, $df = 6$, $p < 0.001$), with a medium effect size (Cramér's $V \sim 0.398$). However, both regional origin and age structure are controlled for in the statistical analysis. Thus, these imbalances pose no serious problems for further analysis.

Table 3. Regional origin of the participants per district and per survey.

Region	Maurer Survey	DaBay Survey
Upper Bavaria	231	1468
Lower Bavaria	156	687
Upper Palatinate	143	368
Central Franconia	239	174
Upper Franconia	169	134
Lower Franconia	132	102
Bavarian Swabia	240	228



Figure 1. Political districts and dialect regions in Bavaria.

3.4. Statistical Approach

Our statistical approach combines descriptive exploration with inferential modelling. For each variable, we first provide the variants and their frequencies in both datasets, followed by an overview of regional variation. For the spatial visualisation, the data are aggregated at the district level (see Figure 1), i.e., the level below the administrative regions shown in Table 3 (*Landkreise* ‘counties’ and *kreisfreie Städte* ‘independent cities’). We selected this level for data aggregation because it allows for a geographically fine-grained representation while remaining more manageable than a depiction at the level of individual municipalities (of which there are more than 2000, compared to only about 100 districts). Note, however, that districts are used solely for visualisation; the statistical analyses are based on the coordinates of individual municipalities, which provide a higher level of spatial precision. As an overlay, the dialectal divisions of Bavaria are shown based on the *Kleiner Bayerischer Sprachatlas* (=KBSA, 2006, ‘Concise Language Atlas of Bavaria’, Renn & König, p. 18). Areas with horizontal stripes indicate transitional zones.

To analyse apparent-time differences, we divide the sample into two age groups: older speakers (>50 years) and younger speakers (18–50 years).

Our statistical approach is based on Generalised Additive Mixed Models (GAMMs). GAMMs extend Generalised Linear Mixed Models (GLMMs) by incorporating non-linear smooth functions of predictor variables. This makes them particularly well suited for modelling geographical variation, as spatial coordinates can be included as smooth terms (cf. e.g., Wirtz et al., 2025, p. 77). In addition, GAMMs can accommodate multiple predictors—both linear and non-linear—as well as random effects. For further details on GAMMs and their application in variationist sociolinguistics, see, e.g., Wieling et al. (2011) and Wieling (2018).

Since it is not feasible to compute GAMMs for each individual variant, we focus on the main variants, defined as those occurring most frequently in both corpora. To model and predict the (non-)realisation of these variants, we employ logistic GAMMs. To account for

geographical variation, each model includes smooth terms for the interaction of longitude and latitude at the research locations. In addition, we include dataset (Maurer vs. DaBay), age group, and gender as categorical predictors. Model selection is based on Akaike's Information Criterion (AIC) (cf. Akaike, 1974). All Generalised Additive Mixed Models were fitted in R (R version 4.2.2) using the *mgcv* package (version 1.9-4, Wood, 2006).

4. Results

4.1. Apple Core

Our first lexical variable concerns the lexical concept APPLE CORE. Table 4 summarises the attested variants and their frequencies in the Maurer and DaBay data.

Table 4. Variants for APPLE CORE.

Variant	Frequency/Maurer Data	Frequency/DaBay Data
<i>(Apfel-)Butz(en)</i>	85.5% (1087 responses)	88.6% (1609 responses)
<i>(Apfel-)Gri(e)bs(ch)</i>	11.5% (146 responses)	3.0% (54 responses)
<i>Kernhaus/-gehäuse</i>	1.3% (17 responses)	2.0% (37 responses)
<i>Krotz(en)</i>	0.5% (6 responses)	0.9% (16 responses)
<i>Rest</i>	—	1.4% (25 responses)
<i>Strunk</i>	—	1.3% (24 responses)
<i>(Apfel-)Kern</i>	—	0.8% (15 responses)
<i>(Apfel-)Grutzl</i>	0.6% (7 responses)	0.2% (4 responses)
others	0.6% (8 responses)	1.8% (32 responses)

As Table 4 demonstrates, *(Apfel-)Butz(en)* is by far the most frequently used variant in both datasets. The second-most-common form is *(Apfel-)Gri(e)bs(ch)*, although it appears considerably more often in the Maurer data than in the DaBay data. Less frequent variants include *Kernhaus/-gehäuse*, *Krotz(en)*, *Rest*, *Strunk*, *(Apfel-)Kern*, and *(Apfel-)Grutzl*. Notably, the forms *Rest*, *Strunk*, and *(Apfel-)Kern* are attested exclusively in the DaBay dataset. These items do not originally denote the concept APPLE CORE: *Rest* is a general term meaning 'remainder', *Strunk* typically denotes the apple stalk, and *(Apfel-)Kern* refers to the seeds. Their usage in the DaBay data may therefore indicate that some speakers had difficulty retrieving a specific (dialectal) lexical item for APPLE CORE.

Figure 2 illustrates the geographical distribution of the variants. *(Apfel-)Butz(en)* predominates throughout the entire research area, with the exception of several Franconian locations—most notably in Upper Franconia—where *(Apfel-)Gri(e)bs(ch)* is common in both datasets. The variant *Krotz(en)* is restricted to western Lower Franconia. The most apparent difference between the datasets is that in the DaBay data the variants *Rest*, *Strunk*, *(Apfel-)Kern*, and *Kernhaus/-gehäuse* occur more frequently in (northern) Franconia. Apart from this, however, the regional patterns are highly similar across datasets, providing little evidence of change.

Figure 3 depicts the distribution of the answers by age and gender. Only minor social differences emerge. In the DaBay data, younger speakers report *(Apfel-)Butz(en)* slightly less frequently and instead rely more on forms such as *Rest*, *Strunk*, *(Apfel-)Kern*, and *Kernhaus/-gehäuse*. In the Maurer data, male speakers use *(Apfel-)Butz(en)* somewhat more often, whereas female speakers more frequently employ *(Apfel-)Gri(e)bs(ch)*. Given the relatively small number of female respondents in the Maurer dataset, this pattern should be interpreted with caution, however.

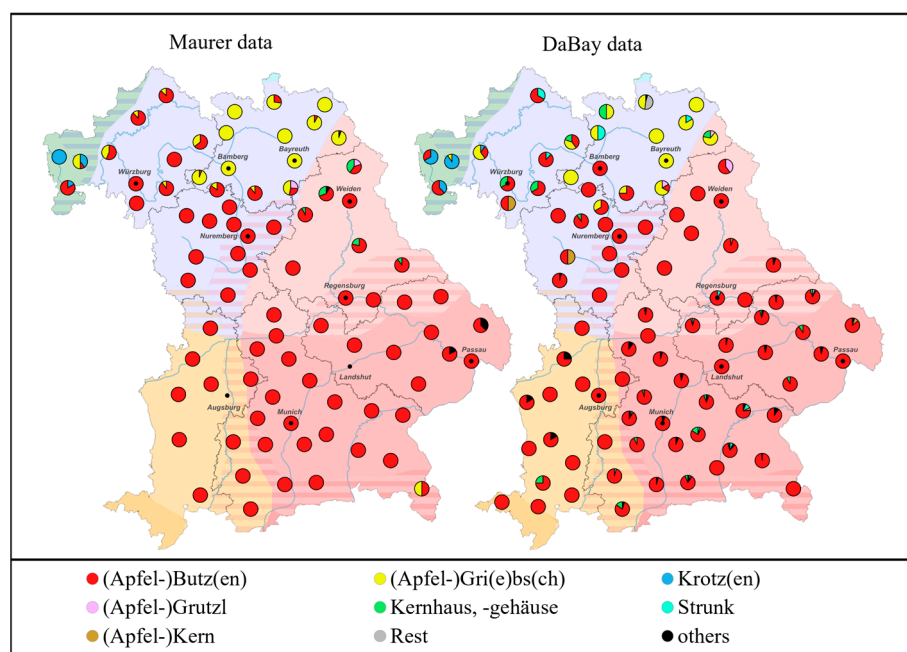


Figure 2. Geographical distribution of the main variants for APPLE CORE.

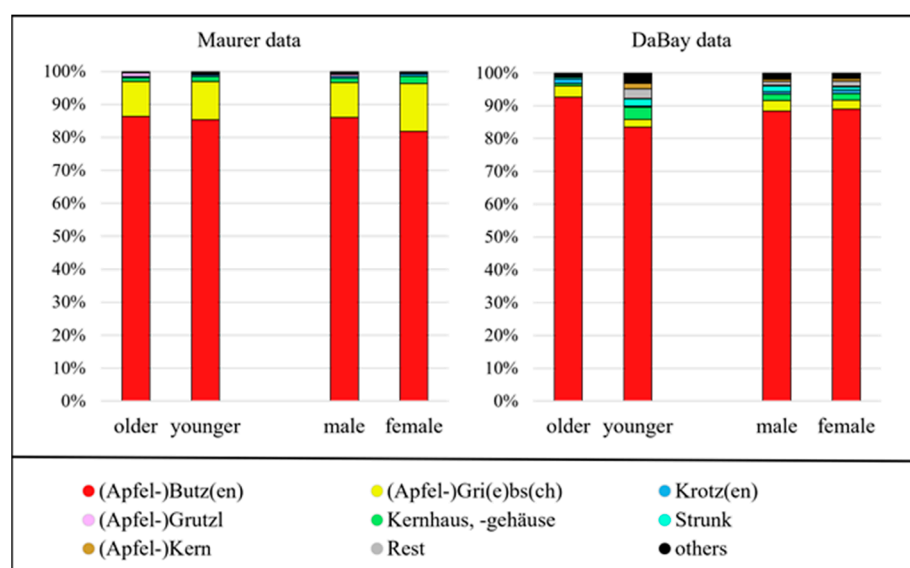


Figure 3. Social differences for the main variants of APPLE CORE.

In sum, both the real-time and apparent-time analyses suggest that the variable is largely stable. Nevertheless, the DaBay data—particularly from northern Franconia—provide some indication that speakers occasionally struggle to identify an appropriate lexical item, which may signal the very early stages of lexical change.

As indicated in Section 3.4, we focus on the main variant of each concept—in the present case *(Apfel-)Butz(en)*—for our statistical testing. Table 5 shows the results of the best-performing GAMM, which includes the datasets (Maurer data vs. DaBay data), the age groups (older vs. younger speakers), the interaction between the datasets and the age groups, and gender (male vs. female) as categorical predictors. Asterisks indicate levels of statistical significance, with * denoting $p < 0.05$, ** denoting $p < 0.01$, and *** denoting $p < 0.001$. For categorical predictors, asterisks indicate interactions.

Table 5. GAMM summary for (*Apfel-*)*Butz(en)*.

Categorical Predictors	Estimate	Std. Error	z	p
(intercept)	3.281	0.279	11.766	< 0.001 ***
dataset [reference level: Maurer data]	−0.055	0.323	−0.170	0.865
age [reference level: older speakers]	−0.039	0.318	−0.122	0.903
age * dataset	−1.407	0.379	−3.710	< 0.001 ***
gender [reference level: male]	−0.032	0.172	−0.184	0.854
Smooth terms	edf	ref. df	χ^2	p
longitude, latitude	22.24	29	399.3	< 0.001 ***

Deviance explained = 45.5%, $n = 2755$, AIC = 1179.8. *** $p < 0.001$.

The model indicates that the dataset—and thus, indirectly, the time periods—has no significant main effect. The same holds for age group and gender as global predictors. However, the interaction between dataset and age is significant and shows a negative effect on the use of (*Apfel-*)*Butz(en)*. Since the Maurer data and the older speakers serve as reference levels, this finding supports our observation that an apparent-time effect is present in the DaBay data. In addition, the interaction of longitude and latitude, representing geographical variation, also turns out to be significant.

4.2. *Ant*

Our second lexical variable concerns the concept ANT. Table 6 presents the variants attested in both datasets and their frequencies.

Table 6. Variants for ANT.

Variant	Frequency/Maurer Data	Frequency/DaBay Data
<i>Ameise</i>	61.5% (781 responses)	93.9% (1465 responses)
<i>Agameise, Hagameise</i>	16.4% (208 responses)	1.5% (23 responses)
<i>Ametze, Ämetze</i>	11.0% (140 responses)	2.4% (38 responses)
<i>Seichameise</i>	8.5% (108 responses)	0.6% (9 responses)
<i>Seichametze, -ämetze</i>	0.5% (6 responses)	–
<i>Pätzameise</i>	1.1% (14 responses)	0.8% (12 responses)
<i>Bärameise</i>	0.6% (8 responses)	–
others	0.3% (4 responses)	0.9% (14 responses)

In both datasets the lexeme *Ameise*—which corresponds to the Standard German lexeme—is the most frequently used variant. However, while it accounts for more than 90% of responses in the DaBay data, it is considerably less common in the Maurer dataset, where several additional variants occur with notable frequency, including *Agameise/Hagameise*, *Ametze/Ämetze*, and *Seichameise*. Less frequent forms include, for example, *Seichametze/-ämetze*, *Pätzameise*, and *Bärameise*. All of these variants are etymologically related to *Ameise* (MHG *âmeiʒe*), though some have undergone folk-etymological reshaping (e.g., *Agameise/Hagameise*; cf. BWB 336–337) or compounding (e.g., *Seichameise*).

Figure 4 shows the spatial distribution of the variants. In the Maurer data, *Ameise* is already predominant in Upper Bavaria, Lower Bavaria, and Upper Palatinate, while *Agameise/Hagameise* is most frequent in Swabia. *Ametze/Ämetze* is concentrated in northeastern Franconia, and *Seichameise* appears primarily in southwestern Franconia. *Pätzameise* occurs mainly in western Lower Franconia. In the DaBay data, the regional differentiation is drastically reduced: *Agameise/Hagameise* is only sporadically attested in northern Swabia, *Ametze/Ämetze* is largely confined to Upper Franconia, and the area where *Seichameise* was previously common has shifted almost entirely to *Ameise*. The only form whose distribution appears relatively stable is *Pätzameise* in western Lower Franconia.

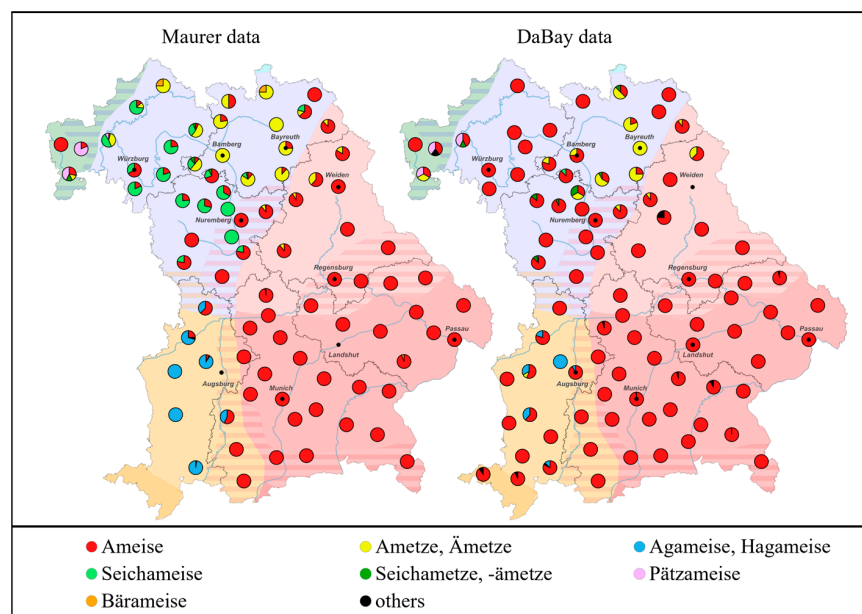


Figure 4. Geographical distribution of the main variants for ANT.

Figure 5 presents the social differences. In the Maurer data, older and younger speakers show very similar patterns. Female respondents use *Agameise/Hagameise* slightly more often, although this may reflect regionally unbalanced gender sampling. In the DaBay dataset, *Ameise* is so dominant that only little social differentiation remains.

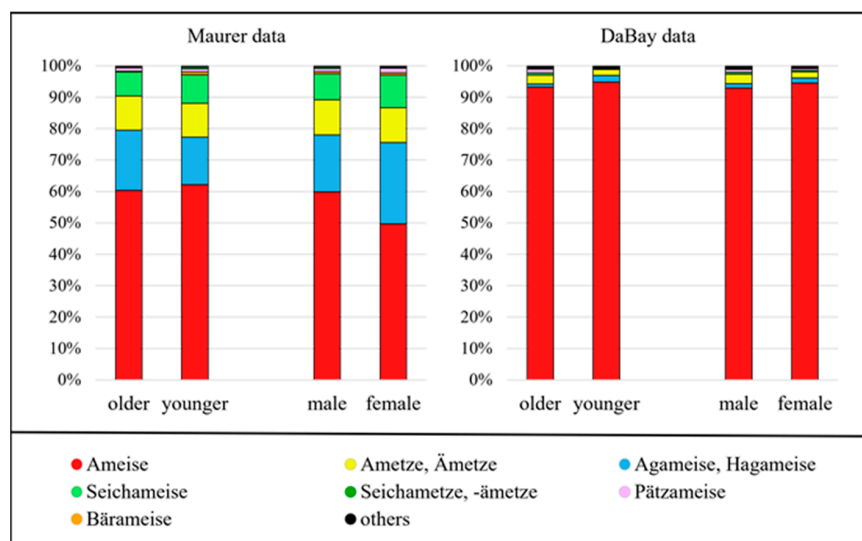


Figure 5. Social differences for the main variants of ANT.

Overall, the real-time comparison reveals a substantial shift from a regionally diverse set of dialect forms toward the widespread use of *Ameise*. Interestingly, this development is not reflected in any apparent-time differences within the Maurer dataset.

For our statistical testing, we again focus on the main variant—in this case, the lexeme *Ameise*. Table 7 presents the results of the final model, which includes the datasets, age group, gender, the interaction between dataset and age group, and the interaction between dataset and gender. Of these categorical predictors, only the effect of the datasets turns out to be significant, supporting the assumption that there is indeed a real-time difference. The effect is positive, indicating a significant increase in the use of this form. In addition, the effect of geography (longitude, latitude) is again significant.

Table 7. GAMM summary for *Ameise*.

Categorical Predictors	Estimate	Std. Error	z	p
(intercept)	2.680	0.283	9.446	< 0.001 ***
dataset [reference level: Maurer data]	1.384	0.292	4.728	< 0.001 ***
age [reference level: older speakers]	−0.268	0.214	−1.251	0.211
age * dataset	0.686	0.361	1.900	0.058
gender [reference level: male]	−0.323	0.317	−1.020	0.309
gender * dataset	0.768	0.419	1.833	0.067
Smooth terms	edf	ref. df	χ^2	p
longitude, latitude	24.36	29	408.7	< 0.001 ***

Deviance explained = 63.6%, $n = 2504$, AIC = 961.0. *** $p < 0.001$.

4.3. Rooster

The third lexical variable examined in this study refers to the concept ROOSTER. Table 8 lists the variants found in both datasets and reports their frequencies.

Table 8. Variants for ROOSTER.

Variant	Frequency/Maurer Data	Frequency/DaBay Data
<i>Gockel</i>	44.5% (569 responses)	57.4% (797 responses)
<i>Gickel</i>	11.0% (141 responses)	31.2% (433 responses)
<i>Göcker, Gücker</i>	13.4% (171 responses)	1.4% (19 responses)
<i>Gicker</i>	13.0% (167 responses)	1.9% (27 responses)
<i>Gecker</i>	4.6% (59 responses)	1.2% (17 responses)
<i>Gockeler</i>	2.4% (31 responses)	0.3% (4 responses)
<i>Hahnengockel</i>	4.9% (63 responses)	0.1% (1 response)
<i>Hahn</i>	0.6% (8 responses)	4.4% (61 responses)
<i>Hahner</i>	2.2% (28 responses)	0.3% (4 responses)
others	3.4% (43 responses)	1.8% (25 responses)

Across both datasets, *Gockel* is the most frequent variant, accounting for more than half of all responses in the DaBay data and slightly less in the Maurer dataset. Note that *Gockel* is an onomatopoeic form, as are *Gickel*, *Göcker*, *Gücker*, *Gicker*, *Gecker*, and *Gockeler*, which differ from *Gockel* only in their stem vowel and/or ending. While these other variants occur relatively frequently in the Maurer data, they are reported only at very low frequencies in the DaBay data—with the exception of *Gickel*. Another relevant change concerns *Hahn*: while scarcely attested in the Maurer data, it appears far more frequently in the DaBay data, likely reflecting the influence of Standard German.

Figure 6 visualises the geographical distribution. In the Maurer data, clear dialect-regional patterns emerge: In the Bavarian and Alemannic regions, *Gockel* predominates, except for Lower Bavaria and adjacent areas, where *Gickel* is more common. In Franconia, forms such as *Gicker*, *Göcker*, *Gücker*, and *Gecker* prevail. The DaBay data, by contrast, show clear patterns of change: *Gockel* has spread into most parts of Franconia, and *Hahn* is already sometimes used there. By contrast, *Gickel* appears relatively stable in the southeastern regions.

Figure 7 illustrates the social distribution. Both datasets show only minimal differences by age group or gender. The most noteworthy observation is that younger speakers in the DaBay dataset use *Hahn* more frequently, consistent with the hypothesis that this Standard German variant might currently be diffusing into the dialect.

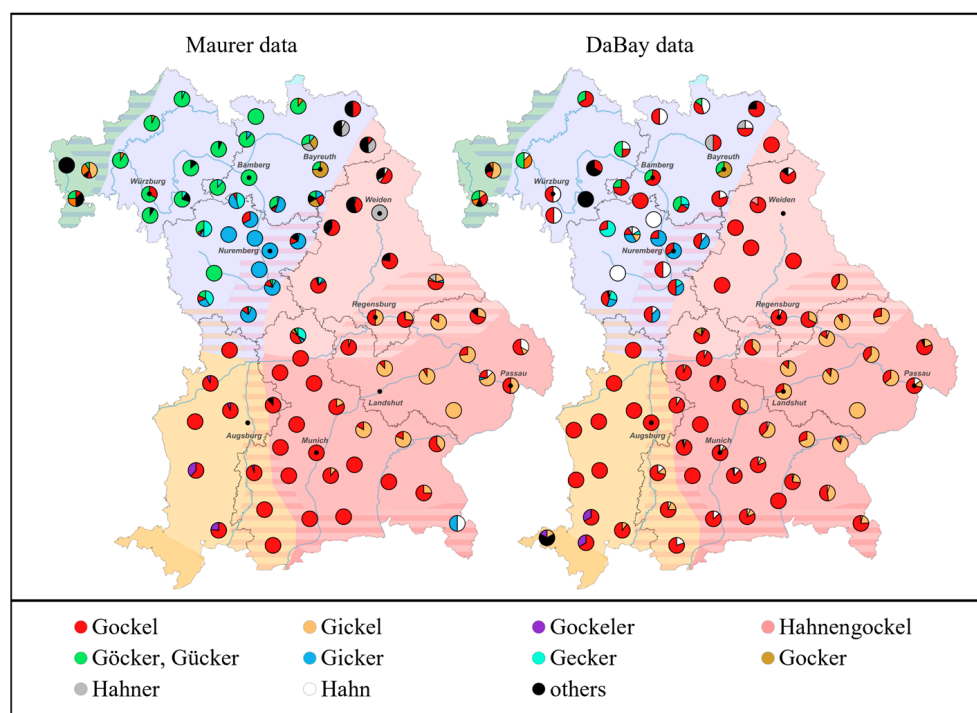


Figure 6. Geographical distribution of the main variants for ROOSTER.

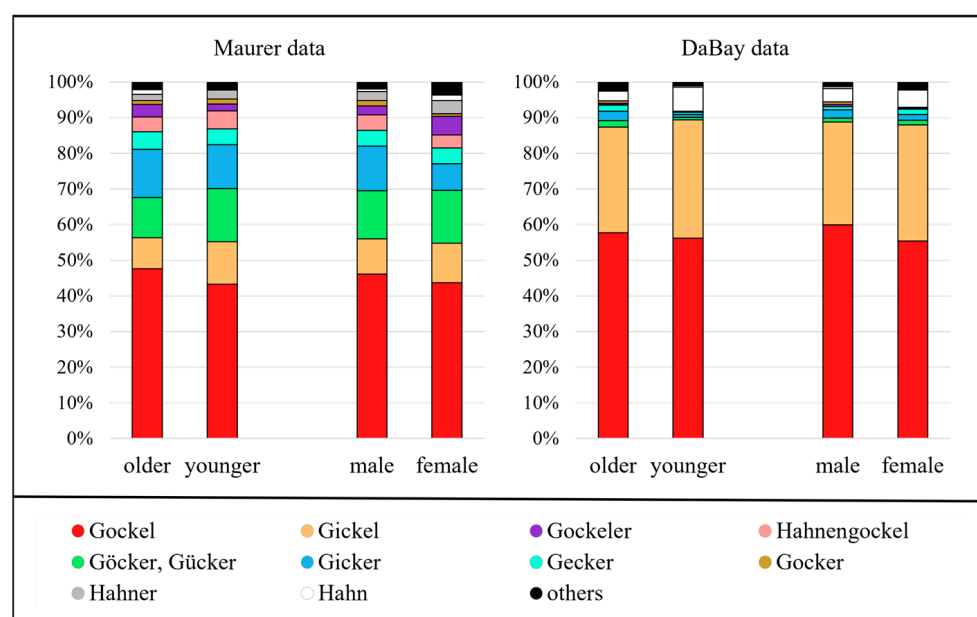


Figure 7. Social differences for the main variants of ROOSTER.

In sum, the real-time comparison suggests an ongoing change in which *Gockel* has spread into Franconia, while *Hahn* is gaining ground—possibly as a result of dialect–standard contact.

Our statistical modelling focuses on the main variant *Gockel* for this variable. The best-performing model (see Table 9) includes the datasets, gender, and age as categorical predictors. Of these, only the effect of the datasets—and thus the real-time effects—are significant. The positive effect indicates an increasing use of *Gockel* in the more recent dataset. Moreover, the effects of geography once again prove to be statistically significant.

Table 9. GAMM summary for *Gockel*.

Categorical Predictors	Estimate	Std. Error	z	p
(intercept)	−0.446	0.124	−3.608	< 0.001 ***
dataset [reference level: Maurer data]	1.122	0.154	7.279	< 0.001 ***
age [reference level: older speakers]	0.093	0.116	0.798	0.425
gender [reference level: male]	−0.236	0.130	−1.824	0.068
Smooth terms	edf	ref. df	χ^2	p
longitude, latitude	26.39	29	735.5	< 0.001 ***

Deviance explained = 39.4%, $n = 2344$, AIC = 2026.5. *** $p < 0.001$.

4.4. Potato

The fourth variable examines lexical variation for the concept POTATO. The variants are listed in Table 10.

Table 10. Variants for POTATO.

Variant	Frequency/Maurer Data	Frequency/DaBay Data
<i>Erdäpfel</i>	61.2% (790 responses)	46.7% (499 responses)
<i>Erdbirne(n)</i>	22.4% (289 responses)	4.7% (50 responses)
<i>Kartoffel(n)</i>	6.4% (83 responses)	41.0% (438 responses)
<i>Grundbirne(n)</i>	5.2% (67 responses)	3.7% (40 responses)
<i>Bodenbirne(n)</i>	2.9% (38 responses)	0.8% (9 responses)
<i>Potacke</i>	1.4% (18 responses)	2.5% (27 responses)
<i>Bumser</i>	0.4% (5 responses)	0.5% (5 responses)

Several clear differences emerge. While *Erdäpfel* is the most common variant in both datasets, it is substantially more frequent in the Maurer data. Similarly, *Erdbirne(n)*, *Grundbirne(n)*, and *Bodenbirne(n)* occur more often in the Maurer dataset. By contrast, the lexeme *Kartoffel(n)*—which corresponds to Standard German—is considerably more frequent in the DaBay data. These differences again suggest a change driven by contact with the standard language.

Figure 8 depicts the geographical distributions. In the Maurer dataset, *Erdäpfel* predominates in most eastern and southern areas. *Grundbirne(n)* is concentrated in Lower Franconia, *Erdbirne(n)* in Central Franconia and adjacent regions (except for the area around Nuremberg, where *Potacke* occurs), and *Bodenbirne(n)* mainly in parts of Swabia and neighbouring Upper Bavaria. *Kartoffel(n)* is rare and mostly attested in Upper Bavaria. While the overall patterns remain recognisable in the DaBay data, *Kartoffel(n)* is far more common—especially in Upper Bavaria, Lower Bavaria, Swabia, and parts of Franconia.

The social differences are shown in Figure 9. In the Maurer dataset, older and younger speakers behave similarly. Female respondents use *Erdäpfel* slightly more often and *Erdbirne(n)* slightly less frequently, though this may reflect a bias in regional sampling. In the DaBay data, younger (and female) speakers use *Kartoffel(n)* more often, supporting the interpretation that the spread of *Kartoffel(n)* is still ongoing.

In sum, the real-time comparison—and possibly also the apparent-time comparison for the DaBay data—indicates a language change toward the standard-conforming variant *Kartoffel(n)*.

Our statistical model (see Table 11) focuses on the variant *Erdäpfel*, which is still the main variant in both datasets. In this case, the model with the lowest AIC value includes only the datasets and age group as predictors, with the effect of the datasets being significant. The negative effect indicates a decreasing use of this variant in the more recent data. Geography is again relevant, as shown by the significant interaction of longitude and latitude.

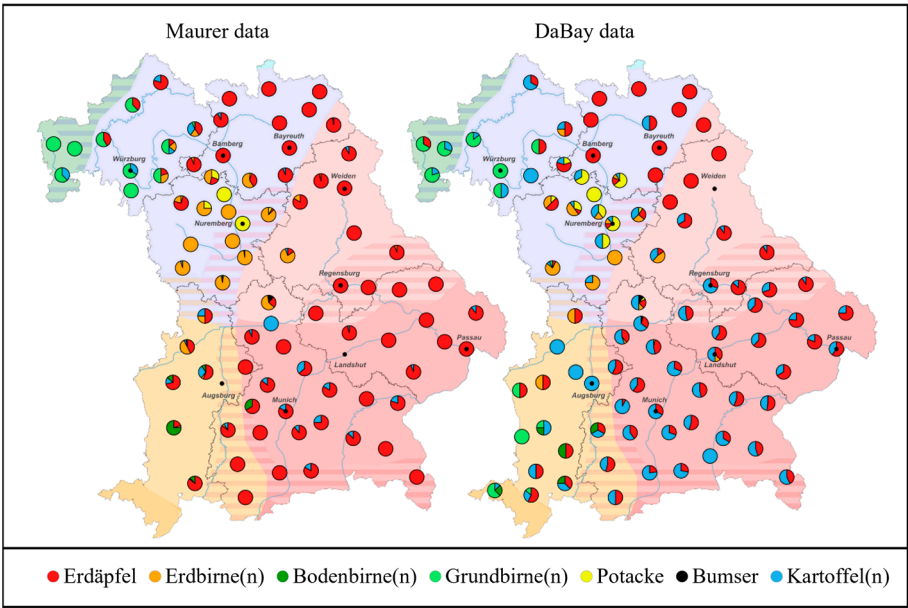


Figure 8. Geographical distribution of the variants for POTATO.

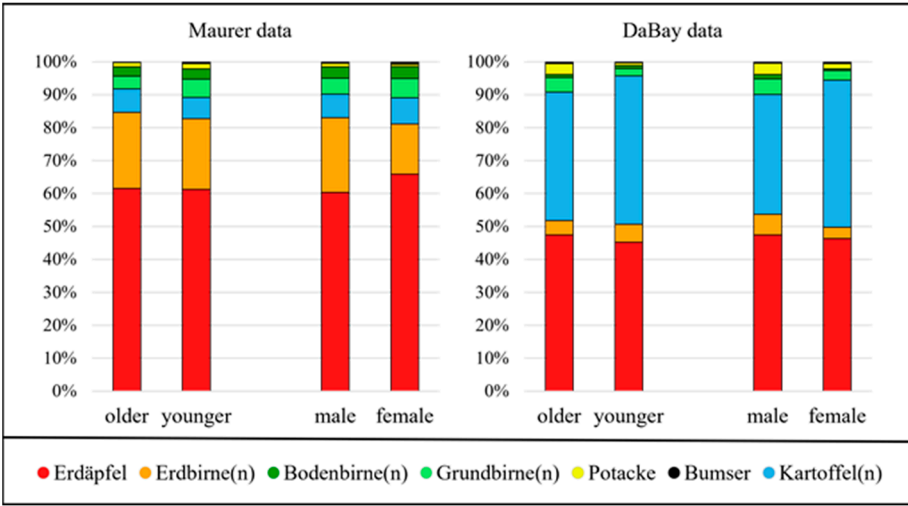


Figure 9. Social differences for the variants of POTATO.

Table 11. GAMM summary for *Erdäpfel*.

Categorical Predictors	Estimate	Std. Error	z	p
(intercept)	1.159	0.135	8.615	< 0.001 ***
dataset [reference level: Maurer data]	−1.624	0.161	−10.083	< 0.001 ***
age [reference level: older speakers]	−0.193	0.117	−1.649	0.099
Smooth terms	edf	ref. df	χ²	p
longitude, latitude	25.21	29	427.2	< 0.001 ***

Deviance explained = 32.3%, *n* = 2027, AIC = 1951.1. *** *p* < 0.001.

4.5. Ladybug

Our final variable concerns the concept LADYBUG. Table 12 presents the main variants and their frequencies.

Table 12. Variants for LADYBUG.

Variant	Frequency/Maurer Data	Frequency/DaBay Data
<i>Frauenkäfer</i>	24.8% (276 responses)	2.8% (44 responses)
<i>Marienkäfer</i>	18.0% (200 responses)	84.3% (1318 responses)
<i>Herrgottskäfer</i>	11.3% (126 responses)	0.3% (4 responses)
<i>Junikäfer</i>	4.4% (49 responses)	6.1% (95 responses)
<i>Muttergotteskäfer</i>	8.8% (98 responses)	—
<i>Herrgottsmuckerl</i>	5.9% (66 responses)	—
<i>Johanniskäfer</i>	3.7% (41 responses)	—
<i>Himmelpetzerl</i>	3.4% (38 responses)	—
<i>Himmelskäfer</i>	2.5% (28 responses)	—
<i>Herrgottsvogel</i>	2.2% (25 responses)	—
<i>Maikäfer</i>	0.2% (2 responses)	3.7% (58 responses)
others	14.8% (165 responses)	2.9% (45 responses)

In the Maurer data, *Frauenkäfer*, *Marienkäfer*, and *Herrgottskäfer* are the most common forms, although several additional variants—such as *Muttergotteskäfer*, *Junikäfer*, and *Herrgottsmuckerl*—also occur with notable frequency. By contrast, the DaBay dataset presents a markedly different picture, with *Marienkäfer*, the Standard German form, overwhelmingly predominating. Apart from *Marienkäfer*, only *Junikäfer*, *Maikäfer*, and *Frauenkäfer* appear with notable frequency. Interestingly, the variant *Maikäfer* is hardly attested in the Maurer data. At least in Standard German, *Maikäfer* refers to a different species (COCKCHAFER), and the relatively high frequency of this term in the DaBay data may indicate that some participants had difficulty identifying the appropriate dialectal term. A similar explanation may apply to *Junikäfer*, which is used not only for the concept of LADYBUG but also for another type of beetle (JUNE BEETLE).

Figure 10 shows the spatial distribution of the different variants. The Maurer dataset exhibits considerable regional differentiation: *Frauenkäfer* is widespread across the Bavarian dialect area, and *Herrgottskäfer*, *Herrgottsmuckerl*, and *Herrgottsvogel* are typical of Franconia, alongside variants such as *Herrgottspetzerl*, *Johanniskäfer*, and *Junikäfer*. In Swabia, *Muttergotteskäfer* and *Herrgottskäfer* are common. *Marienkäfer* occurs most often in Upper Bavaria, Swabia, and Lower Franconia. In sharp contrast, the DaBay data show near-uniform dominance of *Marienkäfer* throughout the research area; only in (Central) Franconia does *Junikäfer* occur with greater frequency. Although *Junikäfer* was already attested in this particular region in the Maurer data, it appears to have spread further, alongside the clear diffusion of *Marienkäfer* throughout Bavaria.

The social patterns are shown in Figure 11. In the Maurer dataset, no systematic age or gender differences are evident—except that younger speakers use *Junikäfer* somewhat more frequently. In the DaBay data, the age differences suggest that *Marienkäfer* is continuing to diffuse, whereas the use of *Junikäfer* appears to be declining.

To summarise, the real-time comparison indicates extensive lexical change: most traditional dialect variants for LADYBUG have been replaced by *Marienkäfer*. As this variant corresponds to the standard lexeme, the change can plausibly be attributed to dialect–standard contact.

Table 13 shows the main outcomes of our statistical modelling for *Marienkäfer* which can be regarded as the main variant for this concept (although it is most frequent only in the DaBay data and ranks second in the Maurer data). The best-performing model includes the datasets, age group, the interaction between age group and dataset, and gender as categorical predictors. Among these predictors, the dataset shows a significant effect, pointing to real-time change, with the positive estimate indicating an increase over time. Moreover, the interaction between age group and dataset is significant, and the positive effect supports the assumption that there is an apparent-time change in the DaBay

data, with *Marienkäfer* still diffusing. As in the other models, the effects of geography are also significant.

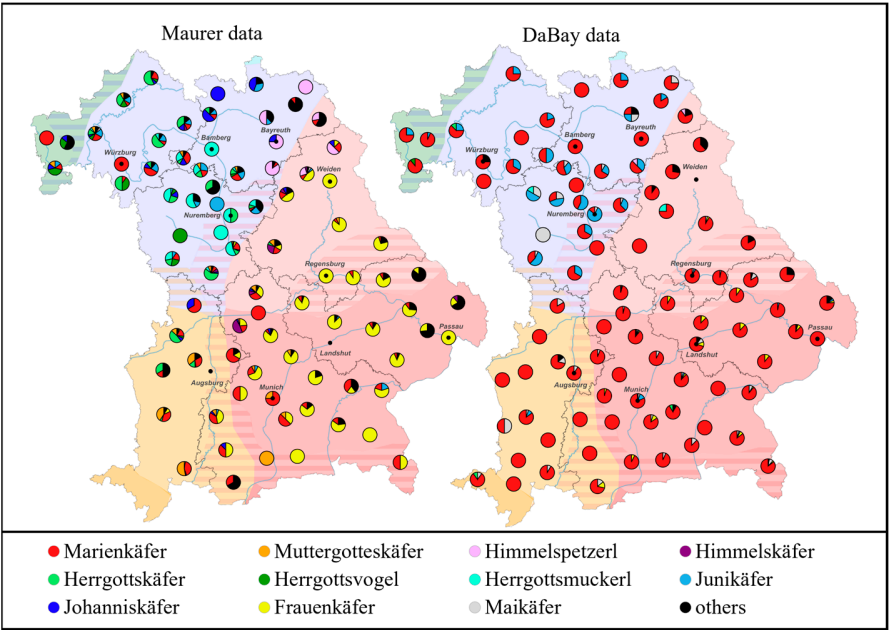


Figure 10. Geographical distribution of the variants for LADYBUG.

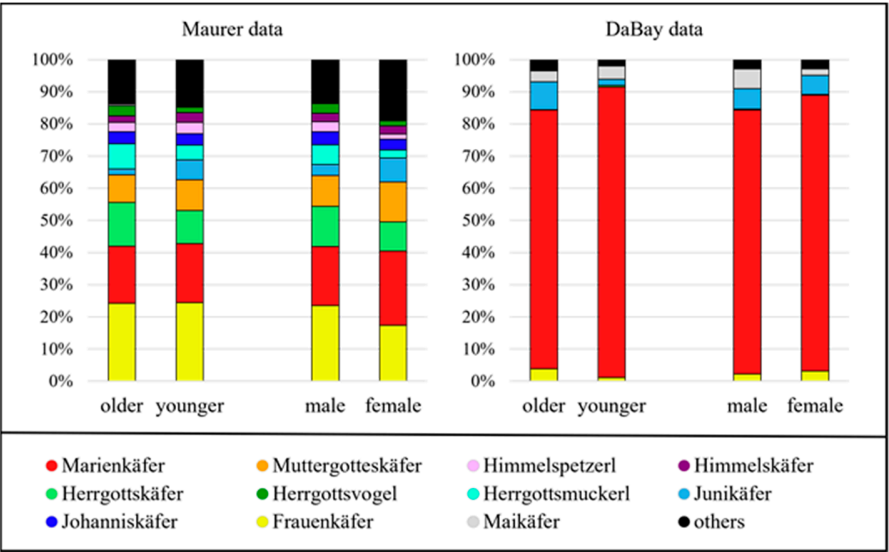


Figure 11. Social differences for the variants of LADYBUG.

Table 13. GAMM summary for *Marienkäfer*.

Categorical Predictors	Estimate	Std. Error	z	p
(intercept)	−1.622	0.154	−10.569	< 0.001 ***
dataset [reference level: Maurer data]	2.996	0.191	15.680	< 0.001 ***
age [reference level: older speakers]	0.034	0.187	0.182	0.855
age * dataset	0.825	0.251	3.291	< 0.001 ***
gender [reference level: male]	0.211	0.127	1.665	0.096
Smooth terms	edf	ref. df	χ ²	p
longitude, latitude	19.91	29	157.6	< 0.001 ***

Deviance explained = 39.0%, n = 2393, AIC = 1997.5. *** p < 0.001.

5. Discussion and Conclusions

The present study presents the first systematic analysis combining real-time and apparent-time perspectives on onomasiological lexical variation and change in the Bavarian, Franconian, and Swabian dialects of Bavaria, drawing on two large datasets collected a century apart: the Maurer questionnaires (1934) and the DaBay dialect app (2025–ongoing). Examination of onomasiological variation in five lexical concepts—APPLE CORE, ANT, ROOSTER, POTATO, and LADYBUG—shows that lexical variation and change is not only empirically tractable but also highly patterned, shaped by geographical and social factors. At a broader level, our findings highlight that dialect lexicons, often assumed to be volatile and weakly areally structured (cf. Hildebrandt, 1983; Wiesinger, 1983), reveal systematic trajectories of change when studied within a variationist real- and apparent-time framework.

Across the five lexical variables, three patterns of change emerge. First, we find evidence of relative stability when long-established heteronyms continue to co-exist over longer time periods. This applies most clearly to APPLE CORE, where (*Apfel-*)*Butz(en)* remains nearly categorical across Bavaria in both datasets. Despite minor indications of lexical uncertainty or underspecification in northern Franconia—reflected in the increased reporting of variants such as *Rest*, *Strunk*, or (*Apfel-*)*Kern* in the DaBay data—no substantial real-time change is detectable. This confirms earlier observations that not all lexical domains are equally susceptible to change and that the diffusion of new heteronyms might be contingent on concept-specific factors such as salience, cultural embeddedness, and frequency of use (cf. Geeraerts, 2010).

Second, three variables exhibit pronounced real-time convergence toward the standard language. This is especially evident in the concepts ANT, POTATO, and LADYBUG. For ANT, the widespread shift from regionally diverse dialect variants (*Agameise/Hagameise*, *Ametze/Ämetze*, *Seichameise*) toward the standard heteronym *Ameise* suggests a substantial diffusion process that has nearly levelled fine-grained dialect distinctions. Notably, this development proceeds without apparent-time effects in the Maurer data, indicating either an acceleration of change in recent decades or a supra-generational process already underway by the early 20th century but only now approaching completion. A similar pattern emerges for POTATO, where *Erdäpfel* remains frequent but shows a clear long-term decline, while *Kartoffel(n)* has expanded dramatically. The diffusion of *Kartoffel(n)* likely reflects a case of standard-driven lexical replacement, a mechanism well-documented also in phonological and morphological variables (cf. e.g., Bülow, 2019; Bülow et al., 2019). The social patterning in the DaBay data—greater use of *Kartoffel(n)* among younger speakers—strongly suggests that the change is ongoing and socially stratified, consistent with standard-oriented levelling processes observed in other Upper German-speaking regions (cf., e.g., Lenz, 2019; Bülow et al., 2019).

Another example of dialect-to-standard convergence is LADYBUG, where the formerly rich assemblage of heteronyms documented in the Maurer data has largely collapsed into a single dominant variant, *Marienkäfer*. Traditional dialect-specific variants such as *Frauenkäfer*, *Herrgottskäfer*, and *Muttergotteskäfer* have nearly vanished. This reduction in lexical diversity illustrates a rapid shift toward a supra-regional lexeme. The significant interaction between age and dataset in the statistical model indicates that the diffusion is not yet complete: younger speakers in 2025 use *Marienkäfer* more often, while *Junikäfer*, a regionally salient Franconian variant, shows (further) signs of decline. Together, these findings underscore that lexical convergence toward the standard can proceed at different speeds across concepts, likely depending on regional factors.

Third, the variable ROOSTER reveals complex areal realignments rather than uniform convergence towards the standard. Here, *Gockel* expands in real-time into Franconia, while *Gickel* remains robust in southeastern Bavaria, and *Hahn*—another standard-identical form—shows signs of diffusion among younger speakers. This competition between dialect variants and

standard-driven forms suggests that lexical change does not always follow a straightforward unidirectional path toward the standard variant. Rather, the lexicon may reflect multiple interacting diffusion processes: some dialect-internal changes from below and other, standard influenced changes from above. Importantly, the continued vitality of *Gickel* in southeastern Bavaria demonstrates that not all dialect variants are equally affected by replacement.

A further important finding concerns geographical patterning, which remains significant across all GAMMs (see Section 4). Despite widespread levelling, geography continues to exert a strong conditioning effect, underscoring the enduring relevance of spatial structure for lexical variation. That is, even in cases of strong standard influence, dialect regions do not behave homogeneously: Franconia, for instance, frequently exhibits lag or partial resistance to adopting large-scale variants, while Upper Bavaria and Bavarian Swabia often lead in adopting such variants. These patterns resonate with long-standing insights from dialect geography but add a crucial real-time dimension that has so far been underexplored in lexical research (but see for the dialects in Austria [Lenz et al., in press](#)).

Finally, social factors play a comparatively modest role, though age effects are detectable in the DaBay data for two variables, namely APPLE CORE and LADYBUG. This suggests that younger speakers are both adopters and transmitters of lexical innovations, aligning the lexicon with trends observed at other linguistic levels in Bavaria. Gender differences, by contrast, appear inconsistent and likely reflect sampling imbalances rather than systematic sociolinguistic patterning. With this in mind, it is important to acknowledge the methodological limitations of this study linked to the indirect nature of both datasets. On the one hand, unlike classical dialectological fieldwork, neither the Maurer survey nor the DaBay survey ensures a sociologically homogeneous speaker base (e.g., NORM/NORFs), nor do they guarantee that responses reflect the oldest attainable local speech forms. On the other hand, comparing the two surveys offers the opportunity to apply apparent-time heuristics and to incorporate selected social factors such as gender into the analysis. The broader social diversity of the respondent pool—with variations in age, gender, and other demographic dimensions—can be seen as an advantage, as it allows us to capture lexical dynamics beyond the traditional focus on NORM/NORF speakers.

Taken together, the results show that lexical variation and change in the dialects of Bavaria is neither random nor uniform. Instead, it proceeds through concept-specific pathways shaped by dialect–standard contact, dialect region, and social dynamics. Only larger follow-up studies can show whether semantic properties also play a role. In any case, both datasets would be suitable for such analyses. Most importantly, the study demonstrates that real-time lexical change can be captured with variationist methods and that large-scale historical survey data such as the Maurer questionnaires provide an invaluable empirical foundation for understanding long-term developments in regional lexicons (see also [Auer & Strobelt, 2025a, 2025b](#)). Future research should expand the set of lexical variables and explore how other factors such as cognitive salience and semantic domains shape the pace and direction of lexical change. For instance, concepts that are highly salient and frequently used in everyday interaction may be less susceptible to standard influence than those that are less salient and/or associated with more specialised domains. However, given that our sample consists of only five concepts, these factors could not be investigated here. Furthermore, future studies need to take into account the role of larger cities and their surrounding commuter belts. Urban centres such as Munich, Nuremberg, or Augsburg are characterised by high dialect–standard contact, increased mobility, and more heterogeneous speaker networks—all conducive to the loss of traditional, small-scale heteronyms in favour of supra-regional or standard-near variants. By doing so, we can move toward a more comprehensive theory of lexical variation and change that accounts for both the stability and the dynamics of the dialect lexicon.

Author Contributions: Conceptualization, L.B. and P.C.V.; preparation of Maurer data, L.B. and O.S.; preparation of DaBay data, P.C.V.; statistics, P.C.V.; writing—original draft preparation, L.B. and P.C.V.; writing—review and editing, L.B. and P.C.V. All authors have read and agreed to the published version of the manuscript.

Funding: The DaBay survey was supported by the Bavarian Academy of Sciences and Humanities (Grant number: Forschungsstipendium 28 August 2024).

Institutional Review Board Statement: The DaBay survey was conducted in accordance with the European Data Protection Regulation. The survey was approved by the Ethics Committee of the Faculty of Languages and Literatures at LMU Munich (no. 2026_0001_LMU1) on 4 February 2026.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the DaBay survey.

Data Availability Statement: The datasets presented in this article are not readily available because the DaBay data contain potentially identifiable information and are therefore subject to data protection regulations. Requests to access the DaBay data should be directed to philip.vergeiner@germanistik.uni-muenchen.de and lars.buelow@germanistik.uni-muenchen.de. Requests to access the Maurer data should be directed to the *Bayerische Akademie der Wissenschaften* ('Bavarian Academy of Sciences and Humanities'), which holds the rights to the Bavarian Maurer questionnaires.

Acknowledgments: We gratefully acknowledge the *Bayerische Akademie der Wissenschaften* ('Bavarian Academy of Sciences and Humanities') as the rights holder of the Bavarian Maurer questionnaires and thank the Academy for granting us permission to use the data.

Conflicts of Interest: The authors declare no conflicts of interest.

Notes

- ¹ The app is funded by the Ludwig-Maximilians-University of Munich and supported by a research grant of the Bavarian Academy of Sciences and Humanities. The project was developed by Philip Vergeiner and Lars Bülow and is overseen and coordinated by Philip Vergeiner.
- ² One notable exception would be the *Pfälzisches Wörterbuch*, which includes 420 maps on lexical and morphological variation in its area of investigation. This dictionary is accessible via the *Trierer Wörterbuchnetz*: <https://woerterbuchnetz.de/?sigle=PfWB&lemid=A00001> (accessed on 7 May 2026).
- ³ Note that most other items included in the Maurer survey are either not included in the DaBay survey or were elicited using different methods (e.g., single-choice questions with predefined response options in DaBay).
- ⁴ In brackets, the number of the respective item in the Maurer survey is given.
- ⁵ Invalid responses mainly reflect obvious misunderstandings of the question (e.g., providing the dialect word for APPLE instead of APPLE CORE, or using the word for CHICKEN instead of ROOSTER).
- ⁶ The Maurer dataset includes 765 speakers aged ≤ 50 (485 male, 83 female, 197 unknown) and 471 speakers aged > 50 (354 male, 49 female, 68 unknown). The DaBay dataset includes 1484 speakers aged ≤ 50 (696 male, 780 female, 8 unknown) and 1677 speakers aged > 50 (790 male, 881 female, 6 unknown).
- ⁷ Note that in DaBay, participants were asked to indicate their age in ten-year intervals.

References

- Akaike, H. (1974). A new look at the statistical model identification. In E. Parzen, K. Tanabe, & G. Kitagawa (Eds.), *Selected papers of Hirotugu Akaike* (pp. 215–222). Springer. [CrossRef]
- Auer, P. (2018). Dialect change in Europe—Leveling and convergence. In C. Boberg, J. Nerbonne, & D. Watt (Eds.), *The handbook of dialectology* (pp. 159–176). Wiley.
- Auer, P., & Strobel, M. B. (2025a). *Dialektgeographie von Baden und dem Elsass aufgrund der Erhebung Frdr. Maurers*. Vol. 1: Wortgeographische Karten. Albert-Ludwigs-Universität Freiburg. Available online: <https://freidok.uni-freiburg.de/data/268976> (accessed on 12 April 2026).
- Auer, P., & Strobel, M. B. (2025b). *Dialektgeographie von Baden und dem Elsass aufgrund der Erhebung Frdr. Maurers (1941)*. Vol. 2: Kommentare zu den Wortkarten. Albert-Ludwigs-Universität Freiburg. Available online: <https://freidok.uni-freiburg.de/data/268977> (accessed on 12 April 2026).
- Bülow, L. (2019). Variation und Wandel der Pluralformen von *sein* in den Dialekten Salzburgs. *Journal LIPP*, 7, 18–38. Available online: <https://lipp.ub.uni-muenchen.de/lipp/article/view/4879> (accessed on 12 April 2026).

- Bülow, L., Vergeiner, P. C., Scheutz, H., Wallner, D., & Elspaß, S. (2019). Wie regelhaft ist Lautwandel? Grundsätzliche Überlegungen anhand einer Fallstudie zu mhd. *ô* in den bairischen Dialekten Österreichs und Südtirols. *Zeitschrift für Dialektologie und Linguistik*, 86(1), 5–50. [CrossRef]
- BWB = Kommission für Mundartforschung. (1985–). *Bayerisches Wörterbuch* (J. Denz, B. Insam, A. Rowley, & H. U. Schmid, Eds.). 2 Vols. (=Bayerisch-Österreichisches Wörterbuch Part II). Oldenbourg.
- Chambers, J. K., & Trudgill, P. (1998). *Dialectology*. Cambridge University Press.
- Debus, F. (1958). *Die deutschen Bezeichnungen für die Heiratsverwandtschaft* (=Beiträge zur deutschen Philologie 19). Schmitz.
- DWA = Mitzka, W., Schmitt, L. E., Hildebrandt, R., & Gluth, K. (1951–1980). *Deutscher Wortatlas* (Vols. 1–22). Schmitz.
- Fleischer, J. (2017). *Geschichte, Anlage und Durchführung der Fragebogenerhebungen von Georg Wenkers 40 Sätzen: Dokumentation, Entdeckungen und Neubewertungen* (=Deutsche Dialektgeographie 123). Olms.
- Fritz-Scheuplein, M., König, A., Krämer-Neubert, S., & Wolf, N. R. (2009). *Wörterbuch von Unterfranken: Eine lexikographische Bestandsaufnahme* (3rd ed.). Königshausen und Neumann.
- Geeraerts, D. (2010). Lexical variation in space. In P. Auer, & J. E. Schmidt (Eds.), *Theories and methods* (=Language and space: An international handbook of linguistic variation 1) (pp. 821–837). De Gruyter.
- Gütter, A. (1971). *Nordbairischer Sprachatlas*. Robert Lerche.
- Harnisch, R. (2019). Ostfränkisch. In J. Herrgen, & J. E. Schmidt (Eds.), *Deutsch* (=Language and space: An international handbook of linguistic variation 4) (pp. 363–406). De Gruyter.
- Hildebrandt, R. (1983). Typologie der arealen lexikalischen Gliederung deutscher Dialekte aufgrund des Deutschen Wortatlases. In W. Besch, U. Knoop, W. Putzke, & H. E. Wiegand (Eds.), *Dialektologie: Ein Handbuch zur deutschen und allgemeinen Dialektforschung* (=Handbücher zur Sprach- und Kommunikationswissenschaft 1.2) (pp. 1331–1367). De Gruyter.
- Hoffmann, W. (1956). *Schmerz, Pein und Weh: Studien zur Wortgeographie deutschmundartlicher Krankheitsnamen*. Schmitz.
- KBSA = Renn, M., & König, W. (2006). *Kleiner Bayerischer Sprachatlas*. Deutscher Taschenbuch Verlag.
- Klausmann, H. (2014). *Schwäbisch: Eine süddeutsche Sprachlandschaft*. Theiss.
- Klee, E. (2005). *Das Personenlexikon zum Dritten Reich: Wer war was vor und nach 1945* (2nd ed.). Fischer Taschenbuch Verlag.
- Kleiber, W., Kunze, K., & Löffler, H. (1979a). *Historischer Südwestdeutscher Sprachatlas. Vol. 1: Text, Einleitung, Kommentare und Dokumentationen*. Francke.
- Kleiber, W., Kunze, K., & Löffler, H. (1979b). *Historischer Südwestdeutscher Sprachatlas. Vol. 2: Karten, Einführung, Haupttonvokalismus, Nebentonvokalismus, Konsonantismus*. Francke.
- Koch, G. (2019). Bairisch in Deutschland. In J. Herrgen, & J. E. Schmidt (Eds.), *Deutsch* (=Language and space: An international handbook of linguistic variation 4) (pp. 279–318). De Gruyter.
- Koch, P., & Oesterreicher, W. (1996). Sprachwandel und expressive Mündlichkeit. *Zeitschrift für Literaturwissenschaft und Linguistik (LiLi)*, 102, 64–96. [CrossRef]
- König, W., Elspaß, S., & Möller, R. (2019). *dtv-Atlas Deutsche Sprache*. Deutscher Taschenbuch Verlag.
- Kranzmayer, E. (1960). *Die bairischen Kennwörter und ihre Geschichte: Mit 5 Skizzen*. Böhlau.
- Labov, W. (1972). *Sociolinguistic patterns*. University of Pennsylvania Press.
- Lameli, A., & Schönberg, A. (2023). A measure for linguistic coherence in spatial language variation. In Y. Scherrer, T. Jauhainen, N. Ljubešić, P. Nakov, J. Tiedemann, & M. Zampieri (Eds.), *Tenth workshop on NLP for similar languages, varieties and dialects (VarDial 2023)* (pp. 133–141). Association for Computational Linguistics. [CrossRef]
- Leemann, A. (2021). Apps for capturing language variation and change in German-speaking Europe: Opportunities, challenges, findings, and future directions. *Linguistics Vanguard*, 7, 20190022. [CrossRef]
- Lenz, A. N. (2019). Bairisch und Alemannisch in Österreich. In J. Herrgen, & J. E. Schmidt (Eds.), *Deutsch* (=Language and space: An international handbook of linguistic variation 4) (pp. 318–363). De Gruyter.
- Lenz, A. N., Ziegler, T., Höll, J., Kunzmann, M., & Dorn, A. (in press). Lexikalische Dynamik in österreichischen Varietäten aus panchronischer Perspektive. In A. Schamberger-Hirt, F. Erhard, M. Schnabel, & V. Schwab (Eds.), *200 Jahre Dialektologie—Forschungsstand und Auftrag. Tagungsband zur 15. Bayerisch-Österreichischen Dialektologietagung vom 13. bis 15. September 2023 in München*.
- Lötscher, A. (2017). *Areale Diversität und Sprachwandel im Dialektwortschatz* (=ZDL-Beihefte 169). Steiner.
- Lötscher, A. (2019). Die areale Lexik im Oberdeutschen. In J. Herrgen, & J. E. Schmidt (Eds.), *Deutsch* (=Language and space: An international handbook of linguistic variation 4) (pp. 679–709). De Gruyter.
- Maurer, F. (1942). Zur Sprachgeschichte des deutschen Südwestens. In F. Maurer (Ed.), *Oberrheiner, Schwaben, Südalemannen*. Hünenburg.
- Mitzka, W. (1938). Der Deutsche Wortatlas. *Zeitschrift für Mundartforschung*, 14(1), 40–55.
- Mitzka, W. (1939). Der Fragebogen zum Deutschen Wortatlas. *Zeitschrift für Mundartforschung*, 15(2), 105–111.
- Murphy, G. L. (2002). Conceptual approaches I: An overview. In D. A. Cruse, F. Hundsnurscher, M. Job, & P. R. Lutzner (Eds.), *Lexikologie/lexicology* (pp. 269–277). De Gruyter.
- Pickl, S. (2013). *Probabilistische Geolinguistik: Geostatistische Analysen lexikalischer Variation in Bayerisch-Schwaben* (=ZDL-Beihefte 154). Steiner.
- SBS = König, W. (Ed.). (1996–2009). *Sprachatlas von Bayerisch-Schwaben*. 16 Vols. (=Bayerischer Sprachatlas: Regionalteil 1). Winter.

- Schmeller, J. A. (1872–1877). *Bayerisches Wörterbuch: Zweite, mit des Verfassers Nachträgen vermehrte Ausgabe* (G. K. Frommann, Ed.). 2 Vols. Oldenbourg. (Original work published 1827–1837).
- Schmid, H.-J. (2020). *The dynamics of the linguistic system: Usage, conventionalization, and entrenchment*. Oxford University Press.
- Schmidt, J. E., & Herrgen, J. (2011). *Sprachdynamik. Eine Einführung in die moderne Regionalsprachenforschung*. Erich Schmidt Verlag.
- Schwarz, C. (2015). *Phonologischer Dialektwandel in den alemannischen Basisdialekten Südwestdeutschlands im 20. Jahrhundert: Eine empirische Untersuchung zum Vokalismus*. Franz Steiner.
- SNiB = Eroms, H.-W. (Ed.). (2003–2010). *Sprachatlas von Niederbayern*. 7 Vols. Winter.
- Stöckle, P., Dorn, A., Gellan, A., Kunzmann, M., & Lenz, A. N. (in press). Dialects in Austria: Lexical variation and change. In A. N. Lenz, S. Elspaß, S. Newerkla, & G. Budin (Eds.), *German in Austria: Variation-Contact-Perception*.
- Stöckle, P., & Wahl, S. (2025). Introduction. In P. Stöckle, & S. Wahl (Eds.), *Lexicography and language variation* (pp. 7–11). Vienna University Press.
- Streck, T. (2012). *Phonologischer Wandel im Konsonantismus der alemannischen Dialekte Baden-Württembergs. Sprachatlasvergleich, Spontansprache und dialektometrische Studien*. Franz Steiner.
- Streck, T. (2019). Alemannisch in Deutschland. In J. Herrgen, & J. E. Schmidt (Eds.), *Deutsch (=Language and space: An international handbook of linguistic variation 4)* (pp. 206–245). De Gruyter.
- Streckenbach, A. (2020). Zur Chronologisierung sprachlicher Daten am Beispiel der Realisierungsvarianten von *nichts* in Unterfranken und am Oberrhein. In R. Langhanke, A. Kleene, A. Schaufuß, & M. Hahn (Eds.), *Dynamik in den deutschen Regionalsprachen: Gebrauch und Wahrnehmung* (pp. 55–85). Olms.
- Strobel, M. B. (2018). Analysen zur Realisierung von mhd. *ë* am Oberrhein in den Antworten zu den Abfragebegriffen *Keller, Kartoffel, Zwetschge, Maulwurf* und *Nest*—Friedrich Maurers Dialekterhebung im Elsass und in Baden. *Freiburger Arbeitspapiere zur Germanistischen Linguistik (FRAGL)*, 39, 1–29. Available online: <https://uni-freiburg.de/germanistik/wp-content/uploads/sites/166/FRAGL-39.pdf> (accessed on 12 April 2026).
- Strobel, M. B. (2021). Die Verschriftungen in der Dialekterhebung Friedrich Maurers in Baden und im Elsass als Evidenz für die Verbreitung der Standardlautung. *Zeitschrift für Germanistische Linguistik (ZGL)*, 49(1), 155–188. [CrossRef]
- Tallen, M. (1963). Wortgeographie der Jahreszeittennamen in den germanischen Sprachen. *Deutsche Wortforschung in europäischen Bezügen*, 2, 159–230.
- Vergeiner, P. C. (2023). Quantitative perspectives on the geolinguistic structures of dialect morphology in Austria. *Dialectologia et Geolinguistica*, 31, 87–115. [CrossRef]
- Vergeiner, P. C. (in press). *NORM:Fs versus Crowd: Zum Einsatz von Crowdsourcing-Apps in der Dialektologie*. Linguistik Online.
- Vergeiner, P. C., & Bülow, L. (in press). Variation und Wandel der Infinitivsyntax im Bairischen. Zum Einfluss semantischer Verbklassen. In A. Werth, M. Gillmann, & D. Höllein (Eds.), *Verbklassen. Kategorisierung zwischen Syntax und Semantik*. De Gruyter.
- Vergeiner, P. C., Bülow, L., & Elspaß, S. (2025). Analysing dialect (morpho)syntax in Austria: A non-aggregative dialectometric approach. *Journal of Linguistic Geography*, 12(2), 57–68. [CrossRef]
- von Schmid, J. C. (1831). *Schwäbisches Wörterbuch: Mit etymologischen und historischen Anmerkungen*. Schweizerbart'sche Verlagshandlung.
- von Sonnleithner, I. (1824). *Idioticon Austriacum, das ist: Mundart der Oesterreicher, oder Kern ächt österreichischer Phrasen und Redensarten* (2nd ed.). Wimmer.
- WA = Wenker, G. (1888–1923). *Sprachatlas des deutschen Reichs*. Research Centre Deutscher Sprachatlas.
- WBF = *Fränkisches Wörterbuch*. (2013–). Available online: <https://www.wbf.badw.de/wbf-digital/wörterbuch/> (accessed on 12 April 2026).
- Wieling, M. (2018). Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English. *Journal of Phonetics*, 70, 86–116. [CrossRef]
- Wieling, M., Nerbonne, J., & Baayen, H. (2011). Quantitative social dialectology: Explaining linguistic variation geographically and socially. *PLoS ONE*, 6(9), e23613. [CrossRef] [PubMed]
- Wiesinger, P. (1983). Die Einteilung der deutschen Dialekte. In W. Besch, U. Knoop, W. Putschke, & H. E. Wiegand (Eds.), *Dialektologie: Ein Handbuch zur deutschen und allgemeinen Dialektforschung* (=Handbücher zur Sprach- und Kommunikationswissenschaft 1.2) (pp. 807–900). De Gruyter.
- Wiesinger, P. (1988). Grundzüge der großräumigen bairischen Wortgeographie. In H. H. Munske, P. von Polenz, O. Reichmann, & R. Hildebrandt (Eds.), *Deutscher Wortschatz* (pp. 555–627). De Gruyter.
- Wirtz, M. A., Elspaß, S., Pickl, S., & Niehaus, K. (2025). Regional and social influences on pragmatic variation in colloquial German. *Zeitschrift für Sprachvariation und Soziolinguistik*, 1(1), 73–94. [CrossRef]
- Wood, S. (2006). *Generalized additive models: An introduction with R*. CRC Press.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.