

Including Swiss Standard German in GermaNet

Eva Huber, Erhard Hinrichs

University of Tübingen

eva.huber@student.uni-tuebingen.de

erhard.hinrichs@uni-tuebingen.de

Abstract

GermaNet (Henrich and Hinrichs, 2010; Hamp and Feldweg, 1997) is a comprehensive wordnet of Standard German spoken in the Federal Republic of Germany. The GermaNet team aims at modelling the basic vocabulary of the language. German is an official language or a minority language in many countries. It is an official language in Austria, Germany and Switzerland, each with its own codified standard variety (Auer, 2014, p. 21), and also in Belgium, Liechtenstein, and Luxemburg. German is recognized as a minority language in thirteen additional countries, including Brasil, Italy, Poland, and Russia. However, the different standard varieties of German are currently not represented in GermaNet. With this project, we make a start on changing this by including one variety, namely Swiss Standard German, into GermaNet. This shall give a more inclusive perspective on the German language. We will argue that Swiss Standard German words, Helvetisms, are best included into the already existing wordnet GermaNet, rather than creating them as a separate wordnet.

1 Introduction

GermaNet is a comprehensive wordnet of Standard German spoken in the Federal Republic of Germany. German is an official language or a minority language in many countries. It is an official language in Austria, Germany and Switzerland, each with its own codified standard variety (Auer, 2014, p. 21), and also in Belgium, Liechtenstein, and Luxemburg. German is recognized as a minority language in thirteen additional countries, including Brasil, Italy, Poland, and Russia.

However, the different standard varieties of German are currently not represented in GermaNet. More generally, among wordnets, there seems to be a lack of accounting for different standards of the same language. To the best of our knowledge, the Princeton WordNet (Fellbaum, 1998) is the only wordnet so far which accounts for standard varieties by marking specifically American or specifically British words. Moreover, a colloquial wordnet of English has recently been created (McCrae et al., 2017). Therefore, it seems worthwhile integrating other German varieties into GermaNet. The central question to this paper, therefore, is how we can successfully model standard varieties. The present study focuses on Swiss Standard German (Swiss StdG). Swiss StdG differs from German on all linguistic levels (Dürscheid and Sutter, 2014, p.37). An orthographic difference pertains to the Eszett ß (“sharp S”), which is in all cases replaced by ss in Swiss StdG (Dürscheid and Sutter, 2014). There are also remarkable phonological differences, such as the primary stress of the initial syllable in, for instance, *Büffet* (Clyne, 1984, p.16). Grammar differences are also found in word order, gender differences, and word derivation patterns. However, lexical differences are by far the most frequent (Dürscheid and Sutter, 2014). At a train station, Swiss people buy a *Billet* (German variant: *Fahrschein*; “ticket”) which they then show to a *Kondukteur* (*Schaffner*, “conductor”) in the *Erstklasswagen* (*Wagen der ersten Klasse*; “first class carriage”). Since wordnets consist of lexemes, we are concerned with the lexical differences. As is common in the literature, we will refer to words which are idiosyncratic for Swiss StdG as *Helvetisms* and to those idiosyncratic for German StdG as *Teutonisms*.

Our approach shall attain a broader representation of German in wordnets and offer a framework for other languages, of which different standard varieties exist, such as Portuguese, Swedish

or French. The paper is structured as follows. First, we will give an overview of GermaNet (Section 2). In Section 3, we will demonstrate how words of Swiss StdG can be collected from lexicographic sources (Section 3.1) and by corpus-based methods (Section 3.2). Section 3.3 presents characteristic examples of Swiss StdG words that have been harvested from lexicographic and corpus-based sources. Section 4 suggests a framework of how to integrate Swiss Standard German. We conclude by discussing possible future work with regard to German varieties (Section 5).

2 GermaNet

GermaNet is a lexical semantic network that is modelled after the Princeton WordNet for English. The resource has been under development for more than twenty years and is still being extended on a continuous basis. The GermaNet team aims at constructing a lexical resource in digital form that models the basic vocabulary of the language. GermaNet covers the most frequently used German adjectives, nouns, and verbs. The coverage of GermaNet is determined by frequency lists compiled from very large digital text corpora of contemporary German. The current data release 13.0 of GermaNet contains 128,100 synsets, 164,814 lexical units, and 148,929 literals. In addition to the inventory of lexical and conceptual relations used in the Princeton WordNet, GermaNet contains a set of lexical relations for nominal compounds. These relations indicate the semantic relations that hold between the constituent parts of a compound. Compounds are also morphologically decomposed into their constituent parts. Release 13.0 contains a total of 82,309 compounds that have been decomposed in this way (Hinrichs et al., 2013).

The coverage of GermaNet is by and large restricted to Standard German. Regional variants and colloquial terms are included only to the extent that they occur frequently in large text corpora and are widely understood. The concept "bread roll" is expressed in Standard German by the lemma *Brötchen* and has many regional variants. One such variant is the term *Wecken*, which is included in GermaNet. *Wecken* belongs to Southern dialects of Germany, but its meaning is widely known, and it occurs with considerable frequency in German corpus data. Therefore, it is reasonable to include such a variant in GermaNet. Compared

to regional variants, colloquial words are included in GermaNet to a higher degree as long as their usage is stable over an extended period of time and as long as they are not offensive.

GermaNet is also linked to the Interlingual Index (ILI; Vossen 1998) that is used to link wordnets for different languages. The synsets for current release of the GermaNet records can be linked to the ILI via 28,566 ILI records. The lexical units in GermaNet can also be linked to a total of 29,550 Wiktionary sense descriptions (Henrich et al., 2014).

3 Detecting and Describing Helvetisms

Switzerland distinguishes itself from Austria and Germany in the sense that Swiss StdG is in a diglossic relationship with the Swiss dialects. While Swiss German dialects, so called *Mundarten*, are used in everyday communication, Swiss StdG occurs in written texts and in news media (Clyne, 1992, p. 119). The Swiss German dialects align themselves with canton boundaries and are acquired as children's first language. Swiss StdG is acquired only once children enter grade school. It is also worth noting that the German Alemannic dialects form a continuum that straddles the German and Swiss border. While it would be worthwhile to include regional varieties of both Germany and Switzerland, this project limits itself to the standard varieties only. In this section, we will discuss how relevant Swiss StdG words can be acquired by lexicographic resources and by data-driven methods.

3.1 Lexicographic Resources

The dictionary "Duden" is the common reference book for the German language, aiming at a full representation of the language ¹(Duden, 2017). The "Schweizerhochdeutsche Duden" (Swiss High German Duden), however, merely lists specific Swiss StdG terms (Bickel and Landolt, 2018). Additionally, the German Duden marks typically *schweizerisch* ("Swiss") or *österreichisch* ("Austrian") words, while Teutonisms, such as *Tesafilm* ("sellotape") are not marked. The German Duden allows for a detection of words which are present in Switzerland as well as in Southern Germany. For instance, the usage of *Nastuch* ("handkerchief") is entered as *süddeutsch*, *schweizerisch*. Furthermore, the

¹<https://www.duden.de/>

Swiss High German Duden specifies *mundart-nahe* words, i.e., words derived from Swiss dialects. Thus, both of these reference works make the gradual characteristics of Swiss StdG to the Mundarten and to German StdG, to a certain degree, explicit. Lexicographic resources offer a valuable data set of words to include in a wordnet. However, some words listed in lexicographic resources are no longer widely used or are used only in certain regions. We, therefore, also consult data-driven methods, which will be described in the following section.

3.2 Data-Driven Methods

Word lists were obtained from two different data sources: The German and the Swiss section of the Leipziger Wortschatz Corpus Collection and news crawls for German and Swiss online materials. The Leipziger Wortschatz Corpus was data-mined by Schneider (Schneider, 2018) using a document classification technique. This method yielded a word list of 21,788 lemmas of all parts-of-speech for which the corpus was tagged. Each lemma was accompanied by a score that indicated the degree to which a word belongs to one standard variety or the other, or whether the word is likely to occur in both varieties. Since the document classification technique does not control for frequency, we also used a frequency-based approach that was facilitated by the frequency lists for the Swiss and German section that are made available along with the Leipziger Wortschatz data. Both frequency lists were truncated to obey a frequency threshold of 50 occurrences. In order to obtain candidate lemmas for Helvetisms, all lemmas from the German frequency list were eliminated from the Swiss frequency list. The same frequency-based method was also applied to filter frequency lists for the news crawls for German and Swiss online domains.

The word lists obtained by the document classification method and by the frequency-based method need to be manually inspected in order to acquire reliable lexical material for Helvetisms relevant for inclusion in a wordnet. Amongst other things, this also means that the candidate lemmas need to be restricted to the three word classes of nouns, verbs and adjectives. Filtering out the other word classes, we obtained 3,712 lemmas of Helvetisms from the Leipziger Corpus and 3,139 from the crawl. The Duden includes approxi-

mately 3,500 lemmas. In order to estimate how many of the words are Helvetisms, we analysed samples including 10% of each data set. Based on the analysis of the samples, 57.14% of the Duden, 9.19% of the list of the Leipziger Corpus, and 5.48% of the crawl list are expected to be Helvetisms. Thus, our data set includes approximately 2,500 Helvetisms, without considering potential overlap between the data set. An analysis of the overlap between the samples of the Leipziger Corpus and the Duden and the crawl list and the Duden respectively shows that the overlap is relatively small. The overlap between the samples from the Leipziger Corpus and the samples from the Duden is 48.6% while the overlap between the samples from the crawl list and the samples from the Duden is merely 11.8%.

3.3 Swiss StdG Words

The Helvetisms that can be harvested from lexicographic resources or from digital corpora fall into different categories (see Lingg 2006; Clyne 1984): words that are derived from the Mundart, loanwords, particularly from French, and culture-specific words pertaining to domains such as politics or sports. The noun *Beiz* ("pub") is one example of a word that is derived from Mundart. It is used interchangeably with the word *Kneipe*, which belongs to the standard varieties spoken in Germany and Switzerland. French loanwords include lemmas such as *Jupe* ("skirt"), which corresponds to German StdG *Rock*. Additionally, Swiss StdG *Papeterie* ("stationary shop") is synonymous to the German StdG *Schreibwarengeschäft*. A further category includes words which are related to Switzerland's culture and tradition, administration and education, and government and political system. Switzerland has special sports, such as *Schwingen*, a kind of wrestling, and *Hornussen*, which obtains its name from a puck called *Hornuss*. Due to the different political systems in Germany and Switzerland, words related to politics are usually specific to its variety. The Swiss political system enables people to propose laws in the form of an *Initiative* ("popular initiative"). Furthermore, *Gegenvorschlag* ("counterproposal") is not as in the German variety merely a "counter proposal", but it is usually used to refer to a suggested alternative to a popular initiative. With regard to Switzerland's education system, we find words, such as *Sportferien* ("winter break") and

Maturitätsprüfung (“final exam”).

One phenomenon that cuts across the various categories of Helvetisms is the word formation process of compounding that is as productive in the Swiss StdG variety as it is in other German varieties. Compounds in Swiss StdG can either be composed of two words which are not associated with any particular variety, or they can include one or more Helvetisms. The constituent words of the nominal compounds *Süssgetränke* (“soda”), *Todesschein* (“death certificate”) and *Gratiseintritt* (“free admission”) are all words that are used in both Swiss and German StdG. Yet, all three compounds are characteristic of Swiss StdG, and have as their German StdG counterparts *Erfrischungsgetränke*, *Totenschein* and *freier Eintritt* respectively. Compounds of Swiss StdG also include loanwords from French, such as *Veloschloss* and *Retourbillet*. In *Veloschloss* the modifier is taken from French, whereas in *Retourbillet* both the head and the modifier are French loanwords.

4 Introducing Swiss StdG into the World of Wordnets

Representing Swiss StdG in a wordnet can be approached in two different ways. In this section, we discuss the two options and illustrate the approach we adopted by specific examples that show how to model Swiss StdG words in a wordnet.

4.1 Two Possible Approaches

The first option is to build a separate wordnet for Swiss StdG and map this new wordnet to the existing GermaNet via the Inter-Lingual-Index (ILI; Vossen 1998). This would generalise the approach taken in EuroWordnet, where several European languages are connected via the ILI. This approach provides a means for systematically linking synonymous and hyponymic words between the two varieties. However, please note that this approach treats Swiss and German StdG as separate languages in the same way as is done in EuroWordnet for, among others, French and German. Such a solution has the following major drawback: it disregards the fact that the vocabulary of Swiss and German StdG is largely overlapping, so that the construction of a separate Swiss wordnet would, to a considerable extent, be redundant with the existing GermaNet in both structure and lexical coverage. Recall that our current estimates for Helvetisms amount to approximately 2,500 lem-

mas (see 3.2), which is only around 10% of the words present in GermaNet.

The second option is to integrate Swiss StdG words directly into GermaNet. This approach follows the strategy adopted in the Princeton WordNet, where words particular to American and British varieties of English are explicitly marked by means of so-called domain region pointers. These pointers link the lexical units to geographical places. For instance, the word *boot*, which is the British expression for the American *trunk*, is marked with the domain region marker relating the word to the synset [United Kingdom, UK, U.K., Britain, United Kingdom of Great Britain and Northern Ireland, Great Britain]. The introduction of domain region pointers into GermaNet allows the modelling of Helvetisms and Teutonisms by linking them to the synsets of [Helvetien, Schweiz] and [BRD, Bundesrepublik Deutschland, Deutschland] respectively. In this approach, words that are used in both varieties are not linked to either of the two synsets. Note also that such an approach is easily generalisable to additional standard varieties of German, whose variety-specific vocabulary would have to be linked to the appropriate synset of the region in which it is spoken.

4.2 Specific Examples

The Swiss StdG words will be integrated into GermaNet so that they are consistent with the overall structure of GermaNet. The same relations will be used, and the only new addition will be the added regional marker to [Helvetien, Schweiz] or [BRD, Bundesrepublik Deutschland, Deutschland] in order to include the three word categories (nouns, verbs and adjectives)².

For the integration of Helvetisms into GermaNet, five different cases need to be observed, which are summarised in table 1. They involve lemmas that are different in both varieties for the same concept (case 1), lemmas that are particular to Swiss StdG or German StdG in addition to synonymous lemmas occurring in both varieties (case 2), and, lastly, lemmas for concepts only used in Swiss or German StdG (case 3). The three different cases are exemplified in tables 2 to 4, and involve in each case different parts-of-speech. The cases in which different lemmas are used for the

²As opposed to the Princeton WordNet, GermaNet does not contain adverbs

case	description
case 1	different lemmas for the same concept
case 2	additional lemma in Swiss StdG
	additional lemma in German StdG
case 3	lemma and concept used in Swiss StdG only
	lemma and concept used in German StdG only

Table 1: Case distinction for Swiss and German StdG words

same concept, e.g. "breakfast" (see table 2), are treated as co-hyponyms in GermaNet, and each lexical unit is tagged by the regional markers linking it to Switzerland, e.g. *Morgenessen*, and to Germany, e.g. *Frühstück*. The treatment of case 2 in GermaNet is also straightforward: words that are particular to Swiss StdG, e.g. *Estrich*, and to German StdG, e.g. *Kraftfahrzeug* (see table 3), are introduced as additional lexical units into the relevant synset, e.g. the synset for "car" or "attic", and are tagged by the appropriate regional domain pointer. The other members in the synset, which belong to both varieties, e.g. *Dachboden* and *Auto*, remain untagged. The lemmas that belong to case 3 denote concepts only used in Swiss or German StdG, e.g. *Sechseläuten* (a Swiss spring holiday) and *Mettwurst* (a German sausage) (see table 4). Thus, the synsets which include lemmas of case 3 contain (a) lexical unit(s) that are all tagged by a regional domain pointer.

If one merges the two standard varieties of German spoken in Switzerland and Germany in the way just outlined, which steps does a lexicographer have to follow to enter all words that appear in a list consisting of Swiss StdG words into GermaNet? Such a word list may have been compiled from a lexicographic resource, such as the Swiss StdG Duden, or from a corpus of Swiss StdG texts, such as the data from the Leipziger Corpus. Given the assumption that the new word should be incorporated into the existing structure of GermaNet, lexicographers need to follow a sequence of steps summarized as the flow chart in Figure 1. The first step is to ensure that the word is not already included in GermaNet. If this is the case, the lexicographer determines whether the word is a true

Helvetism or not. To make this decision, we rely on native speaker intuition, and also additional sources of information, such as Swiss High German corpora and German High German Corpora, are consulted. If the word, however, is not used in Swiss StdG only, the lexical unit is inserted as a new synset and tagged by the regional pointer to [BRD, Bundesrepublik Deutschland, Deutschland] if it is a Teutonism, else it is left unmarked. If the word is, indeed, a Helvetism, there are two possible next steps: either there is already a synset to which the Helvetism can be added (case 1 or case 2), or a new synset has to be created (case 3). In both cases, the lexical unit is marked with the regional domain pointer, linking it to [Helvetien, Schweiz]. If the Helvetism is inserted into an already existing synset, the other members of the synset have to be checked with respect to whether they are Teutonisms and have to be tagged by the regional domain pointer (case 1), or whether they are used in both varieties and are thus left unmarked (case 2).

Already existing words in GermaNet must be re-examined as to whether they are Helvetisms, Teutonisms or used in both varieties. This does not only concern words on the Swiss word list which are already included in GermaNet, but it applies to all words present in GermaNet.

5 Discussion and future work

In this paper, we have shown how to include Swiss StdG into GermaNet by following the approach taken in the Princeton WordNet for linking words from different standard varieties to regional domain pointers. We have emphasised the need for distinguishing between Swiss Mundarten and Swiss StdG and have limited our modelling to the latter. As data sources, we have consulted both lexicographic sources and corpus material and have shown the relative merits of these two sources. It would be worthwhile to broaden the empirical base for identifying Helvetisms by using other data sources, such as informant studies, a traditional method for collecting data on language varieties, and crowdsourcing, which has already been applied to collect colloquial words in a wordnet context by McCrae et al. (2017).

Once the integration of Helvetisms into GermaNet has reached a stable state, the additional data will be released with the yearly updates of the GermaNet resource. GermaNet can be licensed

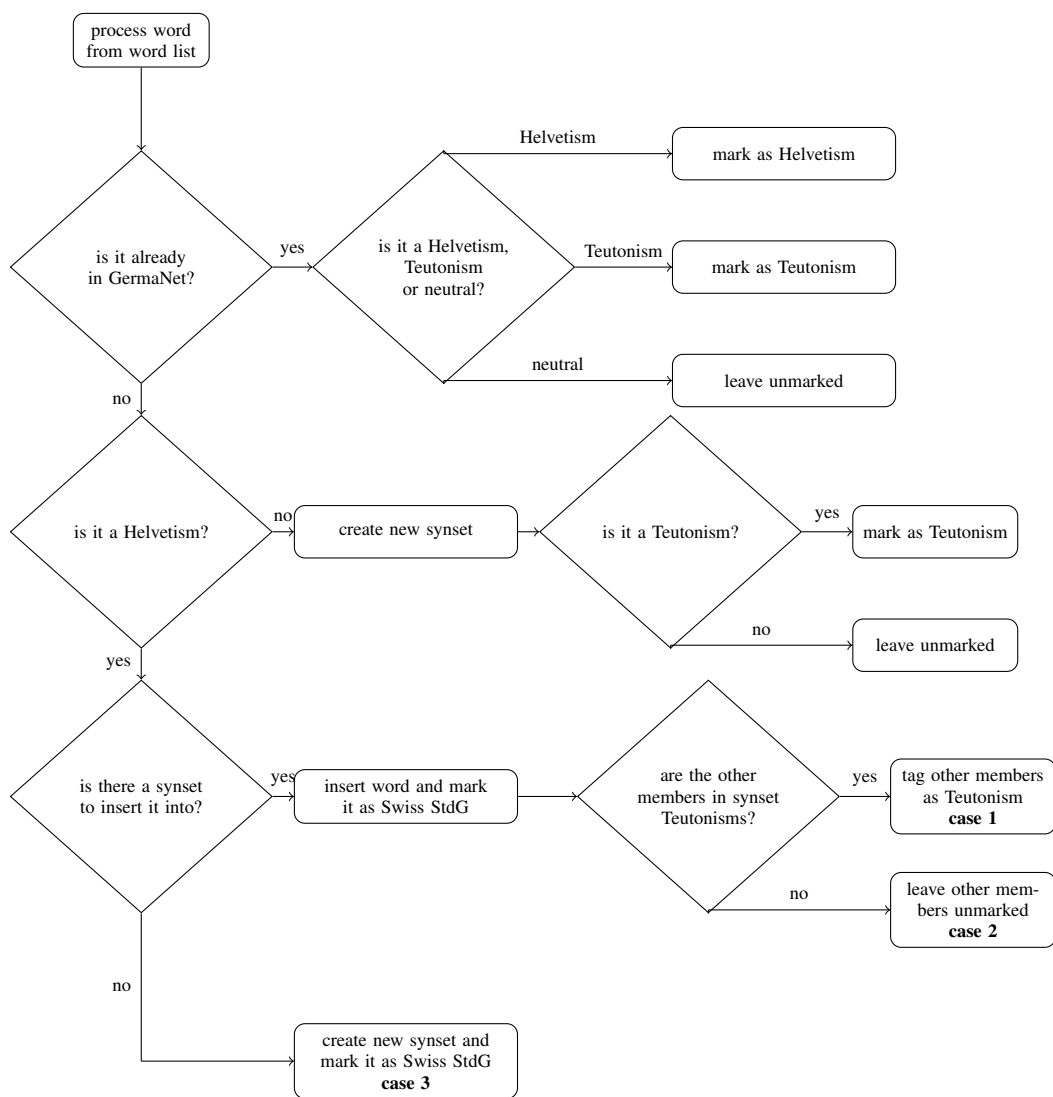


Figure 1: Workflow for lexicographers to include lexemes from the Swiss word list

	example	variety	meaning	part-of-speech
1.	Morgenessen Frühstück	Swiss StdG German StdG	breakfast	noun
2.	parkieren parken	Swiss StdG German StdG	park	verb
3.	Abdankung Trauerfeier	Swiss StdG German StdG	funeral service	noun
4. rule	Aktion Sonderangebot	Swiss StdG German StdG	bargain offer	noun

Table 2: Different lemmas in Swiss and German StdG for the same concept (case 1)

	example	variety	meaning	part-of-speech
1.	Beiz Kneipe	Swiss StdG Swiss StdG and German StdG	breakfast	noun
2.	Estrich Dachboden	Swiss StdG Swiss StdG and German StdG	attic	noun
3.	gehäuselt kariert	Swiss StdG Swiss StdG and German StdG	chequered	adjective
4.	übrissen übertrieben	Swiss StdG Swiss StdG and German StdG	excessive	adjective
5.	Kraftfahrzeug Auto	German StdG Swiss StdG and German StdG	car	noun
6.	artig brav	German StdG Swiss StdG and German StdG	well-behaved	adjective
7.	lauschen hinhören	German StdG Swiss StdG and German StdG	eavesdrop	verb
8.	schmuck dekorativ	German StdG Swiss StdG and German StdG	decorative	adjective

Table 3: Additional lemma in Swiss StdG (1-4) and German StdG (5-8) (case 2)

by academic institutions for research purposes free of charge. Non-academic institutions can license GermaNet for the purpose of internal research and development or for the development of commercial products or services.

A natural next step would be to extend the current approach to other standard varieties, such as the standard varieties spoken in Lichtenstein and Austria. These two countries are of particular interest since both border with Switzerland, and Austria also borders with Germany. Another variety of German worthwhile studying is the German spoken in Luxembourg, a country with Letzeburgisch, German and French as the three of-

ficial languages. Letzeburgisch has been officially recognised as an independent language, but historically has been influenced by Dutch, French and German.

Another issue that we have only touched upon briefly in this paper is the modelling of regional varieties, such as the Swiss Mundarten or regional varieties spoken in Germany. It would be interesting to explore to what extent the approach taken in the Princeton WordNet and also in this paper to the treatment of standard varieties could be generalised to the treatment of regional varieties as well. Here, we can only give some examples from different regional varieties of Switzerland in or-

	example	variety	meaning	part-of-speech
1.	Ausgang -	Swiss StdG German StdG	nightlife	noun
2.	Gegenvorschlag -	Swiss StdG German StdG	counterproposal (in the context of a referendum)	noun
3.	strahlen -	Swiss StdG German StdG	to look for mountain crystals	verb
4.	Sechseläuten -	Swiss StdG German StdG	traditional spring holiday	noun
5.	- Mettwurst	Swiss StdG German StdG	German sausage	noun
6.	- Autohaus	Swiss StdG German StdG	car dealer	noun
7.	- Jahresurlaub	Swiss StdG German StdG	annual holiday	noun
8.	- dufte	Swiss StdG German StdG	smashing	adjective

Table 4: Lemma and concept used in Swiss StdG only (1-4) or in German StdG only (5-8) (case 3)

der to sketch what such an extension would look like. In Swiss Mundarten, the German and Swiss StdG verb *weinen* ("to cry") has the two variants *briäggä* and *brüele* in the dialect spoken in the canton of Zurich and *grännä* is the variant used in the canton of Berne. Similarly, the noun *Brötchen* ("bread roll") has the Mundarten variants *Weggli* used in the canton of Zurich, *Mütschli* in the canton of Berne and *Schwööbli* in the canton of Basel. Modelling such variants in GermaNet would mean to include the variants, e.g. *Weggli*, *Mütschli* and *Schwööbli* or *grännä*, *briäggä* and *brüele*, in one synset that also contains the lexical unit *Brötchen* used in Standard German. The regional variants are then linked to the appropriate domain pointers for the Swiss cantons, while the lexeme *Brötchen* remains unmarked.

Acknowledgments

We thank Reinhild Barkey, Çağrı Çöltekin and Christiane Fellbaum for providing insight and expertise from which this project has greatly benefited. Furthermore, we gratefully acknowledge the financial support of our research by the German Ministry for Education and Research (BMBF) as part of the CLARIN-D research infrastructure grant given to the University of Tübingen.

References

- Peter Auer. 2014. Enregistering pluricentric German. In Augusto Soares da Silva, editor, *Pluricentricity: Language Variation and Sociocognitive Dimensions*, chapter 1, pages 19–48. de Gruyter, Berlin/Boston.
- Hans Bickel and Christoph Landolt. 2018. *Schweizerhochdeutsch*. Duden, Schweizerischer Verein für die deutsche Sprache.
- Michael Clyne. 1992. German as a pluricentric language. In Michael Clyne, editor, *Pluricentric Languages: Differing Norms in Different Nations*, chapter 2, pages 117–148. Walter de Gruyter.
- Michael G Clyne. 1984. *Language and society in the German-speaking countries*. Cambridge University Press Cambridge.
- Duden. 2017. *Duden - Die deutsche Rechtschreibung*, 27th edition. Dudenredaktion.
- Christa Dürscheid and Patrizia Sutter. 2014. Grammatische Helvetismen im Wörterbuch. *Zeitschrift für angewandte Linguistik*, 60(1):37–65.
- Christiane Fellbaum, editor. 1998. *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press.

- Birgit Hamp and Helmut Feldweg. 1997. Germanet - a lexical-semantic net for German. In *Proceedings of the ACL workshop Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications*, pages 9–15, Madrid. Association for Computational Linguistics.
- Verena Henrich and Erhard Hinrichs. 2010. Gernedit - the Germanet editing tool. In *Proceedings of the ACL 2010 System Demonstrations*, pages 19–24, Uppsala, Sweden. Association for Computational Linguistics.
- Verena Henrich, Erhard Hinrichs, and Tatiana Vodolazova. 2014. Aligning Germanet senses with wiktionary sense definitions. In *Human Language Technology: Challenges for Computer Science and Linguistics*, volume 8387 of *Lecture Notes in Computer Science*, pages 329–342. Springer.
- Erhard Hinrichs, Verena Henrich, and Reinhild Barkey. 2013. Using part-whole relations for automatic deduction of compound-internal relations in germanet. In *Language Resources and Evaluation, special issue on "Wordnets and Relations"*, volume 47:4. Springer.
- Anna-Julia Lingg. 2006. Kriterien zur Unterscheidung von Austriazismen, Helvetismen und Teutonismen. In Christa Dürscheid and Martin Businger, editors, *Schweizer Standarddeutsch: Beiträge zur Varietätenlinguistik*, chapter 1, pages 23–49. Gunter Narr Verlag.
- John P McCrae, Ian Wood, and Amanda Hicks. 2017. The colloquial Wordnet: Extending Princeton Wordnet with Neologisms. In *International Conference on Language, Data and Knowledge*, pages 194–202. Springer.
- Gerold Schneider. 2018. Differences between Swiss High German and German German via data-driven methods. In *SwissText 2018: 3rd Swiss Text Analytics Conference, Winterthur, 12 Juni 2018*.
- Piek Vossen. 1998. Introduction to EuroWordnet. In *EuroWordNet: A multilingual database with lexical semantic networks*, pages 1–17. Springer.