Frequency Effects in Ongoing Analogical Change in German Imperatives: Converging Evidence from Corpus and Experimental Data

Inaugural-Dissertation

zur

Erlangung der Doktorwürde

der Philologischen Fakultät
der Albert-Ludwigs-Universität

Freiburg i. Br.

vorgelegt von

Anne Lerche (geb. Krause)

aus Karl-Marx-Stadt

WS 2016/17
Erstgutachter: Prof. Dr. Peter Auer
Zweitgutachter: Prof. Dr. Dr. h.c. Christian Mair
Drittgutachter: Dr. Göz Kaufmann

Vorsitzender des Promotionsausschusses
der Gemeinsamen Kommission
der Philologischen und
Philosophischen Fakultät: Prof. Dr. Joachim Grage

Datum der Disputation: 18.05.2017

Titel der eingereichten Dissertation:
Frequency Effects on Entrenchment: Converging Evidence from Ongoing Language Change
To my family
Acknowledgements

A number of people and institutions have had their share in the completion of the present dissertation, and I would like to take a minute to thank them for their professional and/or financial help or their mental support. First of all, I wish to express my sincere gratitude to my supervisors Peter Auer and Christian Mair for their constant support ever since I started working on the present dissertation project. I received very helpful comments from them and the third assessor of the dissertation, Göz Kaufmann, which greatly improved the present publication.

I am very thankful for the funding which I received from the German Research Foundation (DFG) during my time at the Research Training Group GRK DFG 1624 “Frequency Effects in Language” at the University of Freiburg. The primary corpus study could not have been conducted without the research assistant who compiled the corpus, and the success of the experimental study crucially depended on the participation of a sufficient number of participants. Thus, the financial support by the DFG practically enabled work on these studies. The study programme offered by the Research Training Group provided me with most of the statistical knowledge applied in this dissertation, and I would like to thank Lars Konieczny, Sascha Wolfer and Christoph Wolk personally for their tutoring and patience. Self-evidently, any flaws in the statistical analyses in the dissertation are to be blamed on me.

Of course, I am also indebted to the participants of the experiment for their time and unprompted anecdotes of their experience with the phenomenon under investigation. In this context, I expressly thank the participants of the pilot experiment who sacrificed a small share of their free time to provide me with indispensable data, without receiving an official reward.

I have profited greatly from participating in the colloquia of the Research Training Group, the German Linguistics Department and the Center for Cognitive Science at the University of Freiburg and at Doris Schönefeld’s chair of English Linguistics at Leipzig University, and I wish to thank the professors and colleagues for their helpful remarks, questions and recommendations, among them in particular my supervisors Peter Auer and Christian Mair, Doris Schönefeld, Adriana Hanulikova, Lars Konieczny and Stefan Pfänder. Work on this dissertation was also improved by the comments of participants, for example, at the Corpus Linguistics Conference 2015, the International Cognitive Linguistics Conference 2015, the UK Cognitive Linguistics Conference 2016 and the Annual Meeting of the German Cognitive Linguistics Association 2018.

Parts of the two major studies have been published in the journals Corpus Linguistics and Linguistic Theory (the Walkthrough Corpus study; Chapter 5) and Cognitive Linguistics (the experimental study; Chapter 6). I wish to use this opportunity to thank the editors and
reviewers of the manuscripts for their very helpful comments which improved both the articles and the present publication. I also wish to thank the editors and production teams of both journals for their excellent cooperation.

Last but not least, I am infinitely thankful to my family, friends and, in particular, my partner for their belief in me and the success of the present project, their great mental support, their continuous efforts at keeping me in a good mood and providing exactly the right type and amount of distraction during the past years.
Contents

List of tables ......................................................................................................................................... iv
List of figures ........................................................................................................................................... vi
1 Introduction ................................................................. 1
   1.1 Previous research on frequency effects in analogical change ................................................. 3
      1.1.1 Type frequency and productivity ..................................................................................... 3
      1.1.2 The conserving effect of high token frequency ................................................................. 4
      1.1.3 Frequency patterns as a reflection of mental entrenchment ............................................. 6
      1.1.4 Partial paradigmatic change: the conservation of individual forms ............................... 9
   1.2 German strong verbs with e/i-gradation .......................................................... 10
      1.2.1 Assumed change-in-progress in the imperative singular ................................................... 11
      1.2.2 Historical change in the first person singular ................................................................. 14
   1.3 What counts? Frequency measures reconsidered .............................................................. 17
      1.3.1 Verb lemma token frequency ............................................................................................. 18
      1.3.2 Base verb lemma token frequency .................................................................................... 18
      1.3.3 Type frequency .................................................................................................................. 20
      1.3.4 Intraparadigmatic frequencies ............................................................................................ 20
   1.4 Other factors with a potential influence on imperative formation ........................................ 23
      1.4.1 Language users as creatures of habit ............................................................................... 23
      1.4.2 Dialect variation .............................................................................................................. 27
      1.4.3 Gender patterns .............................................................................................................. 32
2 Methods ........................................................................ 33
   2.1 Corpus-linguistic studies ....................................................................................................... 33
   2.2 Experimental study ............................................................................................................... 34
   2.3 Statistical Analysis .......................................................... 35
3 Data ............................................................................ 40
   3.1 Primary corpora ..................................................................................................................... 40
      3.1.1 The DeReWo frequency list ............................................................................................... 40
      3.1.2 The Walkthrough Corpus ............................................................................................... 41
      3.1.2.1 In search of a corpus .................................................................................................. 41
      3.1.2.2 Compilation of the walkthrough corpus ....................................................................... 45
      3.1.3 Projekt Wortschatz Universität Leipzig ............................................................................ 47
      3.1.4 moca .................................................................................................................................. 47
      3.1.5 FOLK .............................................................................................................................. 48
   3.2 Reference corpora .................................................................................................................. 48
      3.2.1 Frequency dictionaries ................................................................................................... 48
      3.2.2 Duden Online ................................................................................................................ 49
   3.3 Experimental data .......................................................... 49
4 The direction of analogical levelling in strong verbs with e/i-gradation ................................. 50
   4.1 Introduction ........................................................................................................................... 50
      4.1.1 Productivity ..................................................................................................................... 50
      4.1.2 Hypotheses ..................................................................................................................... 51
   4.2 Data ................................................................................................................................. 52
      4.2.1 Previous accounts of frequency distributions .................................................................. 52
      4.2.2 The DeReWo list of lemma token frequency ................................................................ 53
   4.3 Method ............................................................................................................................... 54
      4.3.1 Verb extraction .............................................................................................................. 54
      4.3.2 Verb class annotation .................................................................................................. 54
   4.4 Results ............................................................................................................................... 55
      4.4.1 Schematicity ................................................................................................................ 55
      4.4.2 Type Frequency ............................................................................................................ 56
      4.4.3 Token Frequency ........................................................................................................... 57
   4.5 Discussion ......................................................................................................................... 58
   4.6 Summary ........................................................................................................................... 60
5 The trajectory of analogical levelling in the imperative singular ................................................. 61
  5.1 Hypotheses .............................................................................................................................. 61
    5.1.1 Variation or change-in-progress ....................................................................................... 61
    5.1.2 Verb token frequency ......................................................................................................... 62
    5.1.3 Intraparadigmatic type and token frequencies ................................................................. 62
    5.1.4 Persistence .......................................................................................................................... 63
    5.1.5 Sociolinguistic variables .................................................................................................... 64
    5.1.6 Traditional linguistic factors ............................................................................................ 65
  5.2 Data and method ....................................................................................................................... 66
    5.2.1 Corpus data ...................................................................................................................... 66
    5.2.2 Annotation of variables and distribution ........................................................................... 67
      5.2.2.1 Dependent variables .................................................................................................... 67
      5.2.2.2 Predictor variables ...................................................................................................... 67
      5.2.2.3 Random variables ....................................................................................................... 77
  5.3 Results .................................................................................................................................. 77
    5.3.1 Results of the multinomial regression analysis ................................................................. 78
    5.3.2 Results of the mixed-effects regression analyses ............................................................. 81
    5.3.3 Frequency effects ............................................................................................................. 85
    5.3.4 Persistence effects ............................................................................................................ 86
    5.3.5 Development over time - variation or change? ............................................................... 87
    5.3.6 Influence of Sociolinguistic Variables ........................................................................... 91
  5.4 Discussion of results .............................................................................................................. 93
  5.5 Summary ................................................................................................................................ 99

6 Experimental Study .................................................................................................................. 101
  6.1 Assumptions and hypotheses ............................................................................................... 101
    6.1.1 Measuring the entrenchment of competing forms in analogical levelling in terms of processing latencies ................................................................................................................... 102
    6.1.2 Introducing an off-line measure of entrenchment: recall accuracy .......................... 108
    6.1.3 Applying the apparent-time method in a study of language processing .................. 112
  6.2 Method .................................................................................................................................. 115
  6.3 Data ..................................................................................................................................... 117
    6.3.1 Stimuli ............................................................................................................................. 117
    6.3.2 Participants ....................................................................................................................... 124
    6.3.3 Annotation of variables .................................................................................................... 127
      6.3.3.1 Dependent variables .................................................................................................... 127
      6.3.3.2 Predictor variables ...................................................................................................... 130
      6.3.3.3 Random variables ....................................................................................................... 131
    6.3.4 Meta data from the background questionnaire ............................................................. 131
    6.3.5 Separation of data from participants who did or did not detect the aim of the study ... 132
  6.4 Results for participants who did not detect the aim of the study ........................................ 133
    6.4.1 Impact of the recall task on reading times ..................................................................... 133
    6.4.2 Reading times .................................................................................................................. 134
      6.4.2.1 Residualisation ............................................................................................................ 134
      6.4.2.2 Results of the regression analysis ............................................................................. 136
      6.4.2.3 Discussion of the reading time results ....................................................................... 141
    6.4.3 Recall accuracy ................................................................................................................ 144
      6.4.3.1 Results of the regression analysis ............................................................................. 144
      6.4.3.2 Discussion of the recall results .................................................................................. 147
  6.5 Results for participants who detected the aim of the study ............................................... 148
    6.5.1 Who detected the cover story? ....................................................................................... 149
    6.5.2 Reading times .................................................................................................................. 151
    6.5.3 Recall accuracy ................................................................................................................ 152
    6.5.4 Discussion of results for participants who were not deceived by the cover story ...... 155
  6.6 Discussion of the results of the experiment .......................................................................... 157
  6.7 Summary ................................................................................................................................. 160
7 Conservation of *i* in the second and third person singular .......................................................... 163
7.1 Previous accounts of partial change in morphological paradigms ..................................................... 163
  7.1.1 Natural morphology .................................................................................................................. 164
  7.1.2 Relevance theory .................................................................................................................. 167
  7.1.3 Prototypicality and the Paradigm Structure Hypothesis ........................................................... 169
7.2 Hypotheses .................................................................................................................................. 172
7.3 Data ........................................................................................................................................... 172
7.4 Method ...................................................................................................................................... 173
7.5 Results ...................................................................................................................................... 173
7.6 Discussion ................................................................................................................................. 174
7.7 Summary ................................................................................................................................... 176
8 Conclusions and outlook ..................................................................................................................... 177
References ................................................................................................................................... 181
  Corpora and frequency dictionaries ............................................................................................... 181
  Software ....................................................................................................................................... 182
  Bibliography ................................................................................................................................. 183
Appendices ....................................................................................................................................... 195
  A Eligible verbs for stimulus sentences in the experimental study ...................................................... 195
  B Implementation of regression models in R and presentation of results .............................................. 196
  C Extraction of data from the Walkthough Corpus ............................................................................ 200
  D Results of the MCMCglm regression model .................................................................................. 202
  E Materials for the main experiment ............................................................................................... 203
  F Materials for the pilot experiment ............................................................................................... 218
  G Corpus frequencies of variants of the imperative singular and the second and third person singular of strong verbs with e/i-gradation ................................................................. 224
Zusammenfassung ............................................................................................................................... 225
List of tables

Table 1: Conjugation table of the strong verbs with e/i-gradation geben ‘give’ and sterben ‘die’ ...... 10
Table 2: Present tense inflection of strong verbs with e/i-gradation in Old High German and Old Saxon .................................................................................................................................................... 15
Table 3: Stem vowel changes in the present indicative in seven German ablaut classes from Old High German to New High German.................................................................................................................................................... 16
Table 4: Relation between verb lemma and base verb lemma depending on morphological structure 19
Table 5: Intraparadigmatic type frequency of the stem vowel e in the verbs geben ‘give’ and sterben ‘die’ .................................................................................................................................................... 21
Table 6: Lists of strong verbs with e/i-gradation whose imperative singular is formed with stem vowel i or e, respectively, in Kaiserstuhl and surroundings (Noth 1993: 316-317)............................................................................................................. 31
Table 7: Sample output of the webcrawl for corpus texts ........................................................................ 45
Table 8: Sample output of the speech tagger ........................................................................................................ 46
Table 9: Sample output of the webcrawl for metadata (simulated data) ......................................................... 46
Table 10: Distribution of verb lemmata across verb classes in the DeReWo frequency list .................. 56
Table 11: Productivity constellations in the DeReWo data ................................................................................... 58
Table 12: Distribution of the dependent variables STEM VOWEL and SUFFIXATION in the dataset ............................................................................................................................................................... 67
Table 13: Intraparadigmatic token frequencies for forms of the verbs geben ‘give’ and sterben ‘die’ 71
Table 14: Distribution of five pretargets in the Walkthrough Corpus dataset ................................................ 72
Table 15: Interaction matrix between suffixation and verb type of a pretarget ........................................... 72
Table 16: Interaction matrix LAST.PRETARGET x SECOND.PRETARGET ............................................. 73
Table 17: Distribution of variables DIALECT and COUNTRY in the dataset .................................................. 75
Table 18: Results of the multinom regression analysis of imperative singular formation (Krause-Lerche 2019a: 20) .............................................................................................................................................. 79
Table 19: Conflation of factor levels of variable LAST.SECOND in variable LAST.SECOND.7 ...... 83
Table 20: Results of mixed-effects regression for dependent variable STEM VOWEL (Krause-Lerche 2019a: 18) .............................................................................................................................................. 84
Table 21: Results of mixed-effects regression for dependent variable SUFFIXATION (Krause-Lerche 2019a: 19) .............................................................................................................................................. 84
Table 22: Conditions of target sentences in the main experiment .............................................................. 119
Table 23: Realisation of stimulus sentence conditions in target sentence 11 ................................................ 119
Table 24: Conditions of target sentences in the pilot experiment .............................................................. 120
Table 25: Latin square for stimulus conditions (a-f) in the main experiment .............................................. 123
Table 26: Distribution of participants in group 1 (younger, from Baden-Württemberg) across participant lists ............................................................................................................................................................... 125
Table 27: Distribution of participants in group 2 (older, from Baden-Württemberg) across participant lists .................................................................................................................................................... 125
Table 28: Distribution of participants in group 3 (younger, from Saxony) across participant lists .... 126
Table 29: Distribution of participants in the control group across participant lists .................................. 127
Table 30: Distribution of factor levels of dependent variable REC.ACC ...................................................... 128
Table 31: Reading times in comprehension and repetition condition in Aaronson & Scarborough (1976) and Kuhn (2011) .............................................................................................................................................. 133
Table 32: Results of regression analysis of residual reading times (Krause-Lerche 2019b: 27) ........... 138
Table 33: Results of regression analysis of recall accuracies ................................................................. 145
Table 34: Results of regression analysis of residual reading times of participants who detected the aim of the study .................................................................................................................................................... 152
Table 35: Results of regression analysis of recall accuracies of participants who detected the aim of the study .................................................................................................................................................... 153
Table 36: Mean proportions of analogical variants of paradigm forms in three corpora of German. 163
Table 37: Sample token frequencies for paradigm forms in three corpora of German ................. 173
Table 38: Results of MCMCglmm model ................................................................................................. 202
Table 39: Token frequencies and proportions of established and analogical variants of three paradigm forms of strong verbs with e/i-gradation in the corpus Wortschatz Universität Leipzig......................... 224
Table 40: Token frequencies and proportions of established and analogical variants of three paradigm forms of strong verbs with e/i-gradation in the FOLK corpus ............................................................ 224
Table 41: Token frequencies and proportions of established and analogical variants of three paradigm forms of strong verbs with e/i-gradation in the moca database................................................................. 224
List of figures

Figure 1: Distribution of stem vowels i and e used for the formation of the imperative singular of strong verbs with e/i-gradation in dialects spoken on the territory of the Federal Republic of Germany .......................................................... 29
Figure 2: Main effects and interactive effect of stimulus word frequency and participant gender on reading times (simulated data) ........................................................................................................ 37
Figure 3: Caterpillar plot illustrating intercept adjustments for random factor participant in the analysis of reading times obtained in the experimental study ........................................................................ 39
Figure 4: Type and token frequency distribution of verb classes in the DeReWo list (particle verbs excluded) ............................................................................................................................ 57
Figure 5: Type and token frequency distribution of verb classes in the DeReWo list (particle verbs included) .......................................................................................................................... 58
Figure 6: Type and token frequency distribution of classes “weak, strong and mixed verbs” and “strong verbs with e/i-gradation” in the DeReWo list (particle verbs excluded) ......................... 59
Figure 7: α-persistence in imperative singular formation of strong verbs with e/i-gradation ................ 63
Figure 8: β-persistence in imperative singular formation strong verbs with e/i-gradation .................. 64
Figure 9: Distribution of instances in the dataset over the variable POST.YEAR ............................. 68
Figure 10: Distribution of variables PC1.VERB and PC2.VERB in the dataset ................................. 69
Figure 11: Distribution of variables PC1.BASE and PC2.BASE in the dataset .................................. 70
Figure 12: Effect of variable LAST.SECOND on suffixation of target imperative singular forms ......................... 82
Figure 13: Effect of base verb lemma token frequency (PC1.BASE) on stem vowel choice in imperative singular (Krause-Lerche 2019a: 22) ......................................................................................... 85
Figure 14: Interactive effect of last and second last pretarget (variable LAST.SECOND.7) on suffixation of target imperative singular (Krause-Lerche 2019a: 24) ......................................................... 86
Figure 15: Probability of occurrence of analogical imperative singular variants over time (POST.YEAR as a linear predictor) .................................................................................................. 88
Figure 16: Probability of stem vowel e in imperative singular over time (POST.YEAR as a categorical predictor) .................................................................................................................... 89
Figure 17: Probability of occurrence of analogical imperative singular variants over time (POST.YEAR) (Krause-Lerche 2019a: 26) .............................................................. 91
Figure 18: Intercept adjustment for random factor AUTHOR removes dialect-induced stem vowel preferences .......................................................................................................................... 92
Figure 19: Effect of DIALECT on dependent variable STEM VOWEL ........................................... 93
Figure 20: Sketch of processing latencies reflecting frequency and entrenchment of innovative and conservative variants of the same lemma competing during analogical leveling (Krause-Lerche 2019b: 6) ......................................................................................................................... 103
Figure 21: Frequency and probability of occurrence of imperative singular variants with the stem vowel i dependent on base verb lemma token frequency (Krause-Lerche 2019b: 9) .... 106
Figure 22: Frequency and probability of occurrence of imperative singular variants dependent on base verb lemma token frequency (Krause-Lerche 2019b: 10) ......................................................... 107
Figure 23: Expected processing patterns of imperative variants of strong verbs with e/i-gradation ........ 108
Figure 24: Sketch of replacement rates reflecting frequency and entrenchment of innovative and conservative variants of the same lemma competing during analogical leveling .................. 109
Figure 25: Expected results for dependent variable recall accuracy .............................................. 110
Figure 26: Interactive effect of base verb lemma token frequency and imperative form on residual reading times in the pilot experiment ................................................................. 111
Figure 27: Interactive effect of base verb lemma token frequency and imperative form on recall accuracy in the pilot experiment ......................................................................................... 112
Figure 28: Residuals of the regression of word length on (log-transformed) reading times ................. 135
Figure 29: Effect of word length on reading times in participant groups 1-3 ...................................... 135
Figure 30: Main and interactive effects of base verb lemma token frequency and imperative form on residual reading times obtained from three participant groups (adapted from Krause-Lerche 2019b: 20) ........................................................................................................................................................................ 139
Figure 31: Interactive effect of base verb lemma token frequency and imperative form on recall accuracies of participants from Saxony and effect of imperative form on recall accuracies of participants from Baden-Württemberg .......................................................... 146
Figure 32: Effects of participants’ AGE and GENDER on dependent variable DETECTED ........ 150
Figure 33: Effects of imperative form, participant gender and group on recall accuracies of participants who detected the aim of the study ................................................................................................... 154
Figure 34: Implicational hierarchy of New High German verb inflection .................................... 166
Figure 35: Prototypicality within the paradigm (Janda 2018: 311) ................................................. 170
Figure 36: Token frequencies of the imperative singular and the second and third person singular present indicative of strong verbs with e/i-gradation in three corpora of German .......................................................... 174
Figure 37: Outlier data points in analyses of recall accuracy from participants from Saxony........ 198
Figure 38: Outlier data points in analyses of recall accuracy from older participants from Baden-Württemberg ........................................................................................................................................ 198
Figure 39: Outlier data points in analyses of recall accuracy from younger participants from Baden-Württemberg ........................................................................................................................................ 199
1 Introduction

The present investigation takes as its starting point the observation that an ongoing analogical change seems to affect German verb morphology or, more precisely, the imperative singular form of a group of verbs known as starke Verben mit e/i-Wechsel (‘strong verbs with e/i-gradation’). This name hints at a peculiarity in the inflection paradigm of these verbs: their imperative singular and second and third person singular are traditionally formed with a stem vowel alternation from the infinitive e to i, e.g. geben ‘give’ > gib! ‘give!’, du gibst ‘you give’, er/sie/es gibt ‘he/she/it gives’. This stem vowel alternation appears to be increasingly replaced in the imperative singular by a regular formation with the stem vowel e, sometimes mockingly called the German “emperativ” (Wiktionary contributors 2012). However, the replacement process does not seem to affect the entire verb class to the same extent; while the traditional irregular imperative singular form of verbs like melken (milk!) and bergen (birg!) puzzle native speakers of German, the irregular variants of verbs like geben and nehmen are not questioned to the same extent.

The present dissertation offers an extensive investigation of the assumed analogical change in the imperative singular of the strong verbs with e/i-gradation. In three corpus studies, it tests whether and which frequency effects and other variables can explain the direction and the trajectory of the change and why only the imperative singular form of these verbs is affected by analogical levelling, while the irregular formation is preserved in the second and third person singular present indicative. The prevailing explanation of frequency effects in analogical change on the basis of the cognitive entrenchment of forms is put to the test in an experimental study, in which reading times of the competing traditional and analogical imperative singular variants in verbs of different frequency are measured. Thus, in the analysis of the ongoing change in the paradigm of German strong verbs with e/i-gradation, the present investigation draws on methods and findings from different fields of linguistics, among them in particular cognitive and psycholinguistics, historical linguistics and variationist sociolinguistics. Ultimately, it endeavours to explore as much of the explanatory potential and the limitations of the frequency-based approach with regard to phenomena of ongoing language change. To this end, it often contrasts several frequency measures in a search for the most appropriate variable in a given context.

Chapter 1 will provide the background for all the corpus and experimental studies conducted as part of the present investigation. It gives an overview of previous research on frequency effects in analogical change and explains the cognitive process of entrenchment.
which is assumed to be responsible for such frequency patterns. Moreover, it presents details about the assumed ongoing change in the imperative singular strong verbs with e/i-gradation and a potential historical antecedent. Finally, it also highlights which other factors, apart from frequency variables, are repeatedly found to influence linguistic variation and why they should be and have been taken into account in the present investigation of ongoing language change.

Chapter 2 summarises the corpus-linguistic and experimental methods which have been employed and the logic behind the mixed-effects regression analyses which were fitted on the data obtained from one of the corpus studies and the experiment. Chapter 3 introduces all primary and reference corpora used in the different studies and gives a summary of the data obtained during the experiment.

Chapter 4 is concerned with the first smaller corpus study which asks whether the direction of analogical levelling in the imperative singular of German strong verbs with e/i-gradation can be explained on the basis of the higher productivity of other verb inflection classes and whether their productivity is correlated with the type frequency of inflection classes and/or the token frequency of their members.

Chapter 5 raises two research questions, the first of which is whether analogical levelling in the imperative singular of strong verbs with e/i-gradation is to be considered a language change(-in-progress) or an example of linguistic variation. In addition, it attempts to capture which factors have an influence on the tendency of a strong verb with e/i-gradation to succumb to or resist analogical levelling in the imperative singular. The focus here is on finding out whether change/variation in German morphology is affected by the same or different frequency variables than in other languages.

Chapter 6 is devoted to the descriptions of the details and results of the experiment. This study focuses on the most prominent frequency effect in analogical change, the so-called conserving effect of high token frequency. This effect is typically explained as a consequence of the higher entrenchment of irregular forms of high frequency verbs. Several corpus-based studies have provided evidence for the validity of frequency patterns associated with the conserving effect. However, the present investigation is the first to produce converging evidence for the validity of the entrenchment explanation of this effect from an experiment in which traditional and analogical variants of a variable undergoing analogical levelling are contrasted.

Chapter 7 presents a second smaller corpus study which examines whether a frequency-based approach provides a more satisfactory explanation of the conservation of the
traditional stem vowel alternation in the second and third person singular present indicative of strong verbs with e/i-gradation than other accounts of partial levelling in paradigms.

The final Chapter 8 sums up the results of the present investigation, with a special focus on the explanatory potential and limitations of the frequency-based approach which were examined in the individual corpus studies and the experiment. It also answers the question in how far these findings are relevant to cognitive linguistic theory and gives a very brief outlook on the future of German strong verbs with e/i-gradation.

1.1 Previous research on frequency effects in analogical change

In a number of previous usage-based studies, analogical change like that in the formation of the imperative singular of strong verbs with e/i-gradation has been found to progress through paradigms in a gradual manner, i.e. it does not usually apply to all members of a paradigm to the same extent at the same time. Instead, it has often been observed that some members are levelled earlier and faster, whereas others are conserved in the irregular form for a long time; this conserving effect is typically correlated with the token frequency of the paradigm members. Apart from the trajectory of analogical levelling, frequency effects can also explain the direction of many such changes by pointing out productive patterns which work as a model for the analogy. Finally, a central question in the present investigation is why the imperative singular of strong verbs with e/i-gradation is changing while the second and third person singular present indicative forms with the same stem vowel alternation do not show any signs of being replaced. A frequency-based analysis may yield a more satisfactory answer to this question than previous theories of naturalness/markedness or relevance. The following section will present findings from previous research of frequency effects in analogical change. The methodology of these studies will be extended in the present investigation by introducing frequency measures which may be of special importance to an analysis of ongoing analogical change in the German language in general, or to the analysis of analogical levelling in the imperative singular of strong verbs with e/i-gradation in particular.

1.1.1 Type frequency and productivity

Joan Bybee, one of the forerunners of the frequency-based approach to language, was the first to spell out that type frequency can explain the direction of analogical change because it is a major determinant of productivity in language:

Productivity is the extent to which a pattern is likely to apply to new forms (e.g., borrowed items or novel formations). It appears that the productivity of a pattern, ex-
pressed in a schema, is largely, though not entirely, determined by its type frequency: the more items encompassed by a schema, the stronger it is, and the more available it is for application to new items. (Bybee 2001: 12-13)

Thus, whenever a word enters a language (neologisms, borrowings) and needs to be inflected, for example for plural in the case of nouns or tense in the case of verbs, the pattern which is selected for inflection is typically one that applies to a great number, if not the majority, of items in the same word class in the target language already. This explains why nouns borrowed into English and French usually receive an -s plural. In analogical change, “novel formations” are likewise generally modelled on productive patterns: therefore, in present-day English and German, it is more often the case that strong verbs acquire the productive past tense suffixes -ed (English) or -t (German) and thus become weak verbs than that weak verbs pass into the much less productive strong verb class with stem vowel change in the past tense (Bybee 1995; 2003; cf. Figure 4).

By saying in the passage above that productive patterns are “stronger”, Bybee refers to the relations between items in the mental lexicon of speakers. She explains that when the input contains a high number of items which share a certain pattern, then these items will be related in a speaker’s mental lexicon. The higher the number of items which the pattern applies to, the stronger will be its representation and the easier it is to access when new words need to be inflected (Bybee 1995: 434).

The number of German verbs which do not exhibit a stem vowel alternation in the imperative singular can be expected to be higher than that of strong verbs with e/i-gradation, the only verb class which display such a stem vowel alternation. The hypothesis that the productivity of the weak and other strong verb classes explains the direction of analogical levelling in the imperative singular of strong verbs with e/i-gradation will be tested in the present investigation.

1.1.2 The conserving effect of high token frequency

While type frequency can thus be expected to explain the direction of analogical levelling in the imperative singular of German strong verbs with e/i-gradation, in previous studies it has not been found to explain the trajectory of such changes. Instead, token frequency patterns in analogical change had been noted by linguists like Hermann Paul already in the late 19th century. They observed that throughout language history and cross-linguistically, analogical levelling of irregular formations is most strongly resisted by the “most common” or “most necessary”, i.e. the most frequent, items in a language (Paul 1877: 329; Kruszewski 1890: 139;
Wheeler 1887: 39). However, systematic analyses were not presented until the 1970s when Joan Bybee (Hooper) investigated a case of analogical levelling in English verb inflection. She found that of the Old English strong verbs, those which had become weak verbs by Modern English (past tense and past participle marked by inflectional suffixes) have a strikingly lower frequency than verbs which are still strong verbs in Modern English, i.e. whose past tense and past participle are formed with ablaut (Hooper 1976: 99). Bybee later terms this frequency pattern a **conserving effect**:

Using a morphological example again, high frequency forms with alternations resist analogical levelling ... while [low frequency forms] have a tendency to regularize ...

As a result, morphological irregularity is always centered in the high frequency items of a language” (Bybee & Thompson 1997: 380).

Similar examples of the conserving effect of high token frequency are found in other languages where irregularity is retained only in highly frequent members of a class. With regard to German, Augst explains that more than half of the Old High German strong verbs (222 out of 349) have become weak verbs by New High German (1975: 254-255); Chapter 4 will show that the strong verb class today predominantly contains high frequency verbs, whereas weak verbs generally have a mid or low token frequency (see Figure 4). Thus, Bybee’s observation is confirmed for German: while irregularity is removed in low frequency verbs, it is conserved in high frequency verbs which resist the pressure of analogical levelling. Further examples of this cross-linguistic correlation between frequency and irregularity in verbal and nominal inflection are presented by Haspelmath and Sims (2010: 274-276).

Interestingly, already in her seminal 1976 paper, Bybee hinted at the fact that such a conserving effect of high token frequency may also be observed in “modern leveling”, i.e. in ongoing analogical change:

One case I have investigated involves the six verbs **creep, keep, leap, leave, sleep,** and **weep,** all of which have a past form with a lax vowel ... Of these verbs, three, **creep, leap,** and **weep,** all may have, at least marginally, a past forms [sic] with a tense vowel, **creeped, leaped,** and **weeped.** The other three verbs are in no way threatened by leveling; past forms **kept, leaved, slept** are clearly out of the

---

1 The present investigation will only be concerned with the conserving effect of frequency; for a summary of the “reduction effect” or “reducing effect” of frequency in phonological change, see Bybee & Thompson (1997) and Bybee (2007).

---
question. ... Again the hypothesis that less frequent forms are leveled first is supported. (Hooper 1976: 99-100)

Apart from this example, however, and an examination of the present-day distribution of English and German verb inflection classes, the conserving effect of high token frequency in ongoing change has largely been investigated for morphosyntactic or syntactic phenomena. For example, papers in Bybee and Hopper’s edited volume (2001) present evidence of a conserving effect of frequency in auxiliary selection for the past perfect in English (Smith 2001) and in the use of the subjunctive in Canadian English (Poplack 2001). Bybee herself names further examples of conserving effects in ongoing morphosyntactic change concerning, among others, negation and morphosyntactic properties of English auxiliaries (Bybee 2007: 11) and pronouns (Bybee & Thompson 1997: 380-381). The only investigation of a truly morphological ongoing change concerns the Russian suffix shift (Janda et al. 2010; Nesson & Janda 2010): This shift affects a group of Russian verbs whose traditional -а suffixed forms in the present tense, imperative, present active participle and the gerund are in the process of being replaced by the more productive -ая suffixed forms (e.g. gerund mašа of the verb maxat’ ‘wave’ replaced by maxaja). In a mixed-effects logistic regression analysis of data from the Russian National Corpus, the authors found that the probability for the use of the more productive suffix decreases with verb frequency, i.e. verbs with a higher token frequency show a stronger tendency to resist analogical change (Janda et al. 2010: 40).

Analogical levelling in the formation of the imperative singular of strong verbs with e/i-gradation is a prime object for testing whether a conserving effect of high token frequency can be shown to affect ongoing analogical change in German morphology as well. If this ongoing change does indeed follow a frequency pattern in which low frequency verbs are affected by levelling much earlier and faster than high frequency verbs, this would explain why contemporary speakers of German seem to accept analogical imperative singular variants of low frequency strong verbs with e/i-gradation like melken and verbergen more readily than those of more frequent verbs like geben, nehmen, essen or brechen.

1.1.3 Frequency patterns as a reflection of mental entrenchment

Initially, Joan Bybee (Hooper) explained the conserving effect of high token frequency on the basis of “imperfect learning”; according to this hypothesis, “analogical change comes about as the result of over-generalization on the part of children acquiring their language” (Hooper 1976: 101; Hooper 1980). In their input, children encounter both regular and irregular verb forms; as soon as they start forming schemas from different regular verbs in their input, they
extend these regular schemas to irregular verbs. Bybee assumed that children “get away” more easily with such over-generalisations in low frequent verbs because are not noticed as much as they are in verbs of higher frequency. Therefore, higher frequency verbs are conserved in the irregular formation pattern even when they are transmitted to the next generation. In later studies (Bybee & Slobin 1982; Bybee 1995), however, she showed that the conserving effect is not restricted to acquisition: regularisation patterns as produced by adult speakers are not fundamentally different from the over-generalisations witnessed in children.

In what she called the “network model” (Bybee 1985; 1995), Bybee therefore proposed an alternative explanation based on a correlation between the frequency of a linguistic item and its strength of lexical representation in the mental lexicon:

If we metaphorically suppose that a word can be written into the lexicon, then each time a word in processing is mapped onto its lexical representation it is as though the representation was traced over again, etching it with deeper and darker lines each time. Each time a word is heard and produced it leaves a slight trace on the lexicon, it increases in lexical strength. (Bybee 1985: 117)

In this model, the lexical strength of a linguistic item is also connected to its accessibility (Bybee 1995: 428; 2007: 10): high frequency words are assumed to have such strong lexical representations in speakers’ mental lexicons that their forms are more easily retrieved from memory when needed than those of low frequency words which, by contrast, have weak mental representations. When a speaker wants to use the plural of an irregular high frequency noun, the plural form is quickly accessed in his or her mental lexicon. In contrast, the plural form of a low frequency irregular noun is not accessed as quickly, and the speaker may revert to a regular plural formation pattern instead. Thus, irregular formations in low frequency verbs are likely to be replaced, whereas high frequency words resist analogical levelling. This is how Bybee’s construct of lexical strength attempts to explain the conserving effect of high token frequency in analogical change.

A term which is perhaps more widely used than lexical strength was introduced by Ronald Langacker who describes the correlation between the frequency of linguistic items and their storage in the mental lexicon in terms of degrees of entrenchment: “Every use of a structure has a positive impact on its degree of entrenchment, whereas extended periods of disuse have a negative impact... units are variably entrenched depending on the frequency of their occurrence” (Langacker 1987: 59). The difference between these two models is chiefly
While in Bybee’s model, an increase in frequency leads to a strengthening of lexical representations, it leads to increased entrenchment in Langacker’s terms. In fact, Bybee’s formulation of the consequences of disuse is almost identical with Langacker’s: “[T]he lexicon I am describing here changes with use. Not only do words gain in [lexical] strength, but they can also decline in strength with disuse” (Bybee 1985: 118).

This entrenchment explanation of the conserving effect of high token frequency was echoed in many subsequent publications in the fields of morphology (Diesel 2007; Haspelmath & Sims 2010), morphosyntax (Poplack 2001; Smith 2001) and syntax (Bybee & Thompson 1997). Importantly, however, although entrenchment is a cognitive process, all these investigations are based exclusively on corpus data. As such, they would be considered examples of “cognitive corpus linguistics”, i.e. “research that formulates questions about human cognition in such a way that statistical analysis based on corpus data can yield answers to these questions” (Arppe et al. 2010: 2). Arguably, the question whether the conserving effect of high token frequency reflects mental entrenchment is not one which can be answered on the basis of corpus data alone. Precisely this kind of criticism is voiced by many researchers: On the topic of “corpus frequencies and psychological reality”, Gaëtanelle Gilquin remarks that “corpora are no shortcut to cognition ... [they] cannot stand in for experimental work” because their cognitive reality is at least doubtful (in Arppe et al. 2010: 8). On a related note, Alice Blumenthal-Dramé explains that it is “a long way from acknowledging [a correlation between corpus-extracted frequencies and language processing in the mind] to claiming that token frequencies in corpora should be seen as mirroring entrenchment in the mind” (2012: 44). Finally, Hans-Jörg Schmid who captured the assumed correlation between frequency and entrenchment in its perhaps strongest form in his so-called From-Corpus-to-Cognition-Principle - “Frequency in text instantiates entrenchment in the cognitive system” (2000: 39) - later acknowledges that “frequency as such is no more than an idealized and mechanical approximation of repeated use and exposure” (2016b: 18).

While a few years ago, Schmid even declared that “so far we have understood neither the nature of frequency itself nor its relation to entrenchment, let alone come up with a convincing way of capturing either one of them or the relation between them in quantitative terms” (2010: 125), the situation is not as hopeless as it seems. In the introduction to his recent edited volume Entrenchment and the psychology of language learning, Schmid gives an extensive overview of corpus-based and experimental studies testing the effect of frequency on entrenchment in terms of strength of representation and in terms of chunking and holistic units (2016a). More importantly, he lists studies which have attempted to find converging
evidence for the frequencies found in corpus data in behavioural data obtained from experimental tests. Among them are Alice Blumenthal-Dramé’s (2012) neuro- and psycholinguistic experiments testing the psychological reality of corpus findings on frequency effects on the entrenchment of chunks and Wiechmann’s (2008) comparison of measures of association strength between lexical items and syntactic patterns against data obtained from an eyetracking study.

As concerns the relation between frequency and entrenchment in analogical change, there is abundant evidence from behavioural experiments that frequent words are processed and produced faster (Whaley 1978; Hauk & Pulvermüller 2004; Oldfield & Wingfield 1965; Mousikou & Rastle 2015). In analogical levelling, however, where two forms or variants of the same word, a traditional or conservative irregular one and an incoming or innovative analogical one, are in competition, the question imposes itself whether these variants are also entrenched to a variable extent. In this regard, Alice Blumenthal-Dramé is right in deploring an “essential missing link in the literature [in the form of] empirical studies on the relationship between usage frequency and entrenchment” (2012: 28). The present investigation aims to fill this gap. Thus, it does not only test in a corpus-based study whether a frequency pattern can be found in the ongoing change in imperative singular formation of strong verbs with e/i-gradation (cf. Chapter 5) - and thereby further “increase[s] the size of data sets of the same type” (Arppe et al. 2010: 6). A related experimental study will try to supply converging evidence for the explanation of the conserving effect of high token frequency on the basis of mental entrenchment in terms of reading times and recall of the competing imperative singular variants of strong verbs with e/i-gradation (cf. Chapter 6).

1.1.4 Partial paradigmatic change: the conservation of individual forms

Frequency effects are not usually held responsible for the fact that some forms in a paradigm are subject to analogical levelling, while others are preserved in their traditional form. A notable exception is the study of the Russian suffix shift by Nesset and Janda (2010; Janda et al. 2010). They suggest that morphological paradigms have an inherent structure, and that analogical change proceeds through an irregular paradigm on a hierarchical scale, affecting one form after the other. As outlined in 1.1.2, in their investigation of the Russian suffix shift, Janda and Nesset observed a conserving effect of verb lemma token frequency insofar that verbs with a higher token frequency tend to resist the suffix shift (Janda et al. 2010: 40). In addition, the token frequency of forms in the paradigms of the Russian verbs is highly correlated with their propensity to change: for example, the third person singular present indicative form has the highest intraparadigmatic token frequency, and it is also the form which shows
the highest proportion of (traditional) a-suffixes (Nesset & Janda 2010: 720). These findings thus suggest a conserving effect of high token frequency within paradigms.

The present investigation will test whether token frequencies of forms within the paradigms of strong verbs with e/i-gradation can explain why the present tense stem vowel alternation (e > i) is replaced by analogical variants with the stem vowel e only in the imperative singular but preserved in the second and third person singular indicative. In this regard, the explanatory power of the frequency-based approach will be compared to that of the Paradigm Structure Hypothesis (Janda et al. 2010; Nesset & Janda 2010) and other explanations of partial paradigmatic change, such as natural morphology (Bittner 1996; Dressler 2003) and relevance theory (Bybee 1985).

1.2 German strong verbs with e/i-gradation

There is a class of German strong verbs which, in addition to vowel gradation according to tense, exhibit a stem vowel gradation within the present tense. As Table 1 illustrates, the first person singular and all persons of the plural of the indicative, all forms of the subjunctive, and the second person plural of the imperative are formed with the same stem vowel as the infinitive. In contrast, the second and third person singular indicative and the imperative singular are formed with the stem vowel i, e.g. du gibst ‘you give’, er gibt ‘he gives’, gib! ‘give!’.

Table 1: Conjugation table of the strong verbs with e/i-gradation geben ‘give’ and sterben ‘die’

<table>
<thead>
<tr>
<th>Present</th>
<th>Indicative</th>
<th>Subjunctive</th>
<th>Imperative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singular 1</td>
<td>gebe sterbe</td>
<td>gebe sterbe</td>
<td>gib stirb</td>
</tr>
<tr>
<td>2</td>
<td>gibst stirbst</td>
<td>gebest sterbest</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>gibt stirbt</td>
<td>gebe sterbe</td>
<td></td>
</tr>
<tr>
<td>Plural 1</td>
<td>geben sterben</td>
<td>geben sterben</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>gebt sterbt</td>
<td>gebet sterbet</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>geben sterben</td>
<td>geben sterben</td>
<td></td>
</tr>
<tr>
<td>Past</td>
<td>Indicative</td>
<td>Subjunctive</td>
<td>Present Participle</td>
</tr>
<tr>
<td>Singular 1</td>
<td>gab starb</td>
<td>gäbe stürbe</td>
<td>gebend sterbend</td>
</tr>
<tr>
<td>2</td>
<td>gabst starbst</td>
<td>gäbst stürbest</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>gab starb</td>
<td>gäbe stürbe</td>
<td></td>
</tr>
<tr>
<td>Plural 1</td>
<td>gaben starben</td>
<td>gäben stürben</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>gabt starbt</td>
<td>gäbt stürbet</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>gaben starben</td>
<td>gäben stürben</td>
<td></td>
</tr>
</tbody>
</table>

The imperative singular of these verbs is the focus of the present investigation because it seems to be in a state of change from the traditional formation with the stem vowel i, which can be found in dictionaries and grammars of present-day standard German, to a formation with the stem vowel e. The following sections will explain why this change can be considered a case of analogical levelling (1.2.1), and why it may be part of an overarching process of...
change in the paradigm of strong verbs with e/i-gradation which started out in the formation of the first person singular present indicative several centuries ago (1.2.2).

1.2.1 Assumed change-in-progress in the imperative singular

Examples of imperative singular variants of strong verbs with e/i-gradation formed with the stem vowel $e$ are by no means restricted to present-day German. Diachronic grammars observe that they have been occasionally used in dialects and colloquial German at least since the 16th century (Dammers et al. 1988: 430-431; Lasch 1914: 227; Lübben 1882: 90; Schmidt 2007: 319). Hermann Paul finds individual examples even in literary works composed between the 16th and 19th century (1958: 218-232). According to Tschirch, the spread of these imperative singular variants is a “natural development” which was only halted through the efforts of prescriptive norms:

Daß diese in Mundart und niederer Umganssprache heute allgemein verbreiteten

Imp. ($helfe$; $gebe$; $esse$!) sich die Hochsprache jedoch nicht zu erobern vermocht ha-

ben, das haben die rote Tinte des Lehrers und die Wachsamkeit der Setzer verhindert

und so durch bewußten Eingriff eine natürliche Entwicklung gewaltsam unterbun-

den. (Tschirch 1969: 193-194)

‘That these imperatives which are nowadays commonly used in the dialect and lower colloquial language ($helfe$; $gebe$; $esse$!) have not been able to conquer the standard [original high] language, that has been inhibited by the red ink of the teacher and the attention of the typesetters and thus through conscious intervention [they] violently prevented a natural development’

It may be true that to the present day, the traditional irregular imperative singular of strong verbs with e/i-gradation with the stem vowel $i$ (cf. Table 1) predominates in standard German. However, as Tschirch admits, this form is increasingly replaced by variants with the infinitive stem vowel $e$ in non-standard modern German and different domains of everyday language use; this shows that prescriptive efforts have not been successful in impeding these variants from spreading altogether. Unlike the traditional irregular form, the imperative singular variants formed with the stem vowel $e$ can be either unsuffixed or suffixed. A small set of examples below contains occurrences of the incoming variants in (1) the lyrics of the song “Voo-
doo” by the Hamburg band Deichkind, (2) the children’s page of an advertising booklet, (3) a reader’s poem published in a Thuringian newspaper, and on (4) a meal evaluation computer in a university canteen in Freiburg im Breisgau.
(1) Trete hinein, du bist herzlich willkommen.
Schwebe davon!
Ein Schlick vom Flakon?
Vom leuchtenden Saft mit der feurigen Kraft,
der heute Nacht deine Träume entfacht? (Deichkind 2006)

(2) 

(3) Der Pilzkenner

(4) 

Although the modern and historical examples seem to imply that the imperative singular of strong verbs with e/i-gradation is a variable of stable linguistic variation (Labov 1990:

(Müller Holding Ltd. & Co. KG 2012: 9)

(Spill 2012)
and that an apparent increase in their use is no more than a “frequency illusion” (Zwicky 2005), it can be argued that the paradigm of the strong verbs with e/i-gradation is in fact subject to a process of ongoing change, known as “analogue levelling”: the traditional irregular imperative singular form with the stem vowel $i$ is replaced by regularly inflected imperative singular forms without stem vowel change. Evidence for this assumption can be found in the fact that contemporary speakers react with some surprise when being told that the traditional irregular imperative singular forms of verbs such as *melken* ‘to milk’, *flechten* ‘to braid/ to plait’ and *verbergen* ‘to conceal’ are *milk!*, *flcht!* and *verbIr!* (cf. Nübling et al. 2013: 255). Moreover, in recent popular scientific articles the regularised imperative singular formation is commented as one of the increasingly used ‘quick forms’ which are hardly ever corrected: “Und wer dem Volk genau aufs Maul schaut, kann immer öfter Schnellformen erhaschen wie ... ‘helf mal!’ ... - die zwar grausam klingen, aber oft kaum noch korrigiert werden” (Hinrichs 2016: para. 4).

A genuine change-in-progress in the formation of the imperative singular of strong verbs with e/i-gradation thus seems to be a relatively recent phenomenon, but it may have evolved at least since the 1970s. In the third edition of the perhaps most widely used Duden grammar of German, the description of the formation of the imperative singular of strong verbs with e/i-gradation was annotated with a footnote:

> Wenn Klassiker wie Goethe oder Herder die dem Infinitiv angeglichenen Formen mit “e” (*trete, verspreche, schelte, nehme* usw.) gebrauchen, so ist dies aus dem noch nicht fest gewordenen Gebrauch zu erklären. ... Auch die heutige Umgangssprache bevorzugt sie, sie gelten aber als nicht hochsprachlich. (Grebe 1973: 100)

> ‘When classics like Goethe or Herder use the forms with “e” levelled on the infinitive (*trete* ‘kick/step’, *verspreche* ‘promise’, *schelte* ‘scold’, *nehme* ‘take’ etc.), then this is to be explained from an as yet unfixed use. ... Likewise, present-day colloquial language prefers them, but they are regarded as non-standard.’

In the following fourth edition, the second part of this formulation is slightly modified: “Heute gelten sie mit wenigen Ausnahmen (vgl. *melke!* und nicht mehr: *milk!*) als nicht standardsprachlich.” ‘With few exceptions (cf. *melke!* and no longer: *milk!* ‘milk!’), they [the levelled forms with *e*] are today regarded as non-standard’ (Drosdowski 1984: 174). Moreover, this edition contains another footnote in a general paragraph on strong verbs with e/i-gradation:
In der Gegenwartssprache ist ein Zug zum Systemausgleich insofern festzustellen, als der Imperativ Sing. auch ohne eli-Wechsel gebildet wird. So kann man in der gesprochenen (Umgangs)sprache eß! esse! statt standardspr. iß!, werf!/werfe! statt standardspr. wirf!, brech!/breche! statt standardsprachlich brich! hören. (Drosdowski 1984: 128)

‘In present-day language, a system levelling can be observed insofar as the imperative singular can also be formed without e/i-gradation. Thus, in spoken (colloquial) language one can hear eß!/esse! instead of standard iß! ‘eat!’ , werf!/werfe! instead of standard wirf! ‘throw!’ , brech!/breche! instead of standard brich! ‘break!’ .

This indicates that while analogical imperative forms of other verbs are still regarded as non-standard, at least the analogical form of the verb melken has spread sufficiently by this time to be recorded as standard. Both footnotes are retained in two subsequent editions of the Duden grammar (Drosdowski & Eisenberg 1995; Eisenberg 1998). However, when the grammar is published in a completely redesigned and extended seventh edition (Eisenberg & Kunkel-Razum 2005), both footnotes have disappeared. This is most probably a decision of the new editorial board: while they abound in all previous editions, the newly designed edition contains hardly any footnotes at all. The fact that no changes with regard to the explanation of imperative singular formation of strong verbs with e/i-gradation are introduced in the revised reprint of the seventh and in the subsequent eight edition of the Duden grammar (Eisenberg & Kunkel-Razum 2006; 2009) does not necessarily mean that analogical forms of strong verbs with e/i-gradation have declined in use. Although linguistic change, in particular if it is ongoing, may always revert (Mair 2008: 1110), the examples given above and contemporary speaker’s reactions to the traditional irregular imperative singular forms of verbs like melken, flechten and verbergen draw a different picture. In any case, this rather confusing state of affairs calls for a systematic analysis of the imperative singular of strong verbs with e/i-gradation, hereby providing a more conclusive answer to the question whether it is indeed affected by analogical levelling.

1.2.2 Historical change in the first person singular

The assumed change-in-progress in the formation of the imperative singular of strong verbs with e/i-gradation may be an extension of a levelling process which affected the forms of the first person singular present indicative in the same verb class. In Old High German (500-1050
AD) and Old Saxon/ Old Low German (800-1150/1200)\textsuperscript{2,3}, inflection in the present tense was divided: the (infinitive) stem vowel $\ddot{e}$/$e$ was used for all plural forms and the stem vowel $i$ for all singular forms (indicative and imperative; see Table 2).

Table 2: Present tense inflection of strong verbs with e/i-gradation in Old High German and Old Saxon

<table>
<thead>
<tr>
<th></th>
<th>Old High German</th>
<th>Old Saxon/Old Low German</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Indicative</td>
<td>Imperative</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Singular</td>
<td>hilfu (-o)</td>
<td>quithu (-o)</td>
</tr>
<tr>
<td></td>
<td>hilfis (-ist, -est)</td>
<td>hilfis (-ist, -est)</td>
</tr>
<tr>
<td></td>
<td>hilfit (-et)</td>
<td>quithid (-id, -it)</td>
</tr>
<tr>
<td>Plural</td>
<td>hëlffumês (-amês,</td>
<td>quetha (-and, -ant)</td>
</tr>
<tr>
<td></td>
<td>-emês, -ém, -ên)</td>
<td>hëlffat (-et)</td>
</tr>
<tr>
<td></td>
<td>hëlffat (-et)</td>
<td>hëlffat (-et)</td>
</tr>
<tr>
<td></td>
<td>hëlffant (-ent)</td>
<td>quetha (-and, -ant)</td>
</tr>
</tbody>
</table>

(Schmidt 2007: 240; Galleé 1993: 245)

Note: Parentheses give alternative inflectional suffixes.

Since the beginning of the Middle Low German (1150/1200-1600/1650 AD) and Early New High German period (1350-1650 AD), the first person singular present indicative was increasingly used with the plural stem vowel $\ddot{e}$/\textit{e}. According to Lasch (1914: 224) and Lübben (1882: 71-75), this replacement process was complete by the beginning of Middle Low German, and the first person singular was used exclusively with the stem vowel $e$. Dammers et al. (1988: 443-456) present evidence that the change progressed gradually southwards in High German: in Middle German areas (areas 6-10 in Figure 1 and the Greater Berlin area), the first person singular occurs only with the stem vowel $\dot{e}$ by the second half of the 15\textsuperscript{th} century, in Nuremberg and the Swabian and Alsatian language areas by the late 17\textsuperscript{th} century, and in Bavarian and High Alemannic first person singular forms with the stem vowel $i$ are not completely replaced by the end of the Early New High German period but linger on until the middle of the 18\textsuperscript{th} century.

Schmidt explains that this change in the first person singular present indicative form of strong verbs with e/i-gradation is a case of analogical levelling: among the seven ablaut verb classes in German, the sixth and seventh already had umlaut in the second and third person singular present indicative but not in the first person. The Early New High German change in the paradigm of strong verbs with e/i-gradation (ablaution classes 3b, 4, and 5) and a similar change in ablaut class 2 are thus part of a “levelling movement affecting the first, second and

\textsuperscript{2} In historical linguistics, German is traditionally divided into High German (Upper and Middle German) and Low German along the so-called “Benrath line”, which corresponds largely to the lower border of areas 3-5 in Figure 1 (Middle German includes the Greater Berlin area), on the basis of the High German consonant shift. This shift affected Upper German (in the South of Germany) and Middle German (areas 6-10 in Figure 1 and the Greater Berlin area), but the pronunciation of Low German consonants (in the North) remained unchanged (Schmidt 2007: 73-74).

\textsuperscript{3} cf. table “Periodisation of the German language” in Schmidt (2007: 22)
third persons singular present indicative in most ablaut classes” (Schmidt 2007: 393-394; cf. Ernst 2012: 145-146).

Table 3 illustrates that Old and Middle High German indeed had three different systems for present tense inflection of strong verbs. Ablaut classes I, IIIa, and VIIc showed the same stem vowel throughout the present tense. In classes VI and VIIa and later also VIIb and VIId-f, one stem vowel was used for the first person singular and all persons of the plural, and another stem vowel was used to form the second and third person singular. In the remaining classes, II, IIIb, IV and V, present tense inflection was divided according to number: the infinitive stem vowel was used for the plural and a different stem vowel for the singular.

<table>
<thead>
<tr>
<th>Ablaut Class</th>
<th>Old High German</th>
<th>Middle High German</th>
<th>Early New High German</th>
<th>New High German</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>sg/pl</td>
<td>î</td>
<td>î</td>
<td>ie</td>
</tr>
<tr>
<td></td>
<td>2/3sg</td>
<td>u</td>
<td>iu</td>
<td>iu</td>
</tr>
<tr>
<td></td>
<td>pl</td>
<td>io/e</td>
<td>ie</td>
<td>ie</td>
</tr>
<tr>
<td>IIIa</td>
<td>sg/pl</td>
<td>i</td>
<td>i</td>
<td>i</td>
</tr>
<tr>
<td>IIIb</td>
<td>1sg</td>
<td>i</td>
<td>i</td>
<td>i</td>
</tr>
<tr>
<td></td>
<td>2/3sg</td>
<td>i</td>
<td>i</td>
<td>i</td>
</tr>
<tr>
<td></td>
<td>pl</td>
<td>ê</td>
<td>ê</td>
<td>e</td>
</tr>
<tr>
<td>IV</td>
<td>1sg</td>
<td>i</td>
<td>i</td>
<td>i</td>
</tr>
<tr>
<td></td>
<td>2/3sg</td>
<td>ê</td>
<td>ê</td>
<td>e</td>
</tr>
<tr>
<td></td>
<td>pl</td>
<td>ê</td>
<td>ê</td>
<td>e</td>
</tr>
<tr>
<td>V</td>
<td>1sg</td>
<td>i</td>
<td>i</td>
<td>i</td>
</tr>
<tr>
<td></td>
<td>2/3sg</td>
<td>ê</td>
<td>ê</td>
<td>e</td>
</tr>
<tr>
<td></td>
<td>pl</td>
<td>ê</td>
<td>ê</td>
<td>e</td>
</tr>
<tr>
<td>VI</td>
<td>1sg</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>2/3sg</td>
<td>ê</td>
<td>ê</td>
<td>e</td>
</tr>
<tr>
<td></td>
<td>pl</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>VIIa</td>
<td>1sg</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>2/3sg</td>
<td>e</td>
<td>e</td>
<td>e</td>
</tr>
<tr>
<td></td>
<td>pl</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>VIIb,d,e</td>
<td>1sg</td>
<td>â, ô, ou</td>
<td>â, ô, ou</td>
<td>a, o, au</td>
</tr>
<tr>
<td></td>
<td>2/3sg</td>
<td>æ, Æ, öu</td>
<td>æ, Æ, öu</td>
<td>â, Æ, öu</td>
</tr>
<tr>
<td></td>
<td>pl</td>
<td>â, ô, ou</td>
<td>â, ô, au</td>
<td>a, o, au</td>
</tr>
<tr>
<td>VIIc</td>
<td>sg/pl</td>
<td>ei</td>
<td>ei</td>
<td>ei</td>
</tr>
<tr>
<td>VIIf</td>
<td>1sg</td>
<td>uo</td>
<td>ü</td>
<td>u</td>
</tr>
<tr>
<td></td>
<td>2/3sg</td>
<td>üe</td>
<td>ü</td>
<td>ü</td>
</tr>
<tr>
<td></td>
<td>pl</td>
<td>uo</td>
<td>ü</td>
<td>ü</td>
</tr>
</tbody>
</table>

(Schmid 2013: 118-129; Schmidt 2007: 242-247, 303-305, 390-398)

Note: Grey shaded cells denote strong verbs with e/i-gradation.
When in Early New High German the first person singular form of the latter classes (II, IIIb, IV and V) was levelled on the stem vowel used for the plural, this seems at first to complicate matters - however, only when the respective ablaut classes are considered in isolation. When New High German inflection of strong verbs is considered as a whole, analogical levelling in the first person singular of strong verbs with e/i-gradation and in the singular of the second ablaut class creates a dichotomous system in which either i) all present indicative forms have the same stem vowel (classes I, II, IIIa, VIIc, VIII) or ii) only the second and third person singular present indicative have a different stem vowel from the infinitive (IIIb, IV, V, VI, VIIa,b,d,e).

When the imperative is taken into account, the inflection of strong verbs with e/i-gradation constitutes an exception from the New High German dichotomous system insofar that it is the only verb class which retains a stem vowel change which does not concern the second and third person present indicative. Thus, analogical levelling of the imperative singular form of these verbs on the stem vowel used in the present indicative for the first person singular and the plural (which is identical with that of the infinitive) may be seen as an extension of the levelling process in the first person singular. Upon completion, analogical levelling in the imperative singular of strong verbs with e/i-gradation would eventually install a perfectly dichotomous present tense inflection system for the indicative and imperative mood. Chapter 7 will return to the idea of a dichotomy in the inflection of German strong verbs in relation to the conservation of the e/i-gradation in the second and third person singular present indicative of strong verbs with e/i-gradation.

It would be interesting to examine as well whether the historical change in the first person singular of strong verbs with e/i-gradation not only progressed gradually throughout Germany but also across the verbs in the paradigm. Examples given by Dammers et al. (1988: 443) suggest a frequency pattern which is assumed for the change in imperative singular formation as well: the forms of low frequency verbs succumb to analogical levelling more readily than those of higher frequency verbs (see 1.1.2). However, such an analysis would go beyond the scope of the present investigation.

1.3 What counts? Frequency measures reconsidered

From previous research, it seems that the definition, annotation and influence of some frequency variables can be considered relatively unproblematic, while others are not taken into account in many studies although they might have an influence on the formation of the imperative singular of strong verbs with e/i-gradation. In fact, even after several studies (e.g.
Poplack 2001; Smith 2001) had confirmed the existence of frequency effects in language change which Bybee suggested from her earliest work on (Hooper 1976; Bybee 1995; Bybee & Hopper 2001), she explicitly addressed the topic of which frequency measures are important or “useful” and says that the question of “what to count” depends on the investigated phenomenon (Bybee 2007: 17). The present investigation aims to explore the potential, but also the limits, of the frequency-based account in explaining ongoing morphological change; therefore, the following studies will take into account frequency measures whose effect is by now widely accepted, but it will extend the approach by testing the impact of frequency measures which have so far been largely neglected. The German strong verb with ei-gradation *geben* ‘to give’ (and occasionally also *sterben* ‘to die’) will be used in order to illustrate the different frequency measures.

1.3.1 Verb lemma token frequency

As explained in 1.1.2, the relationship between high token frequency and irregularity is widely accepted in the linguistic community. This is interesting because Joan Bybee (Hooper) is the only one of the authors quoted above who explicitly says what kind of “frequency” they counted: “To obtain these [verb] frequency counts, I totalled all entries that could be verb forms, for example, *drive, drives, driven, drove,* and *driving*” (Hooper 1976: 99). This frequency measure, commonly referred to as verb lemma token frequency, thus represents a count of all inflected forms of a verb in all tenses, moods/aspects, numbers (singular and plural) and persons (e.g. first and third person); it will be used in the present analysis as well. For the verb *geben*, this means counting all tokens of all forms in the paradigm, e.g. *gebe, geben, gibt, gab, gäbe, gegeben* etc. (cf. Table 1).

1.3.2 Base verb lemma token frequency

This frequency measure is one that is not usually distinguished in the literature, but it may be very important in the present investigation: The syntax of German imperative sentences differs depending on the morphological make-up of the verb. Simplex verbs (5) and complex verbs which consist of a simplex verb and a prefix (6) behave very much like their English equivalents:

(5) *Gib mir das Buch!*

‘Give me the book!’

(6) *Vergib mir, Vater, denn ich habe gesündigt.*

‘Forgive me, father, for I have sinned.’
Particle verbs, i.e. complex verbs which consist of a simplex verb base and a particle, in contrast, show a very different pattern in the imperative. While in English, objects/complements and adverbials follow or precede the verb and the particle, in German both objects/complements and adverbials are inserted between the verb and its particle.

(7) **Gib die Hoffnung niemals auf!**

Give the hope never up

‘Never give up hope!’

Such structures with stranded verb particles exist in English as well; however, as the translation of an example sentence (modified from Bybee 1995: 437) shows, the number of constituents which can occur between the base verb and its particle is limited in English, but in German a theoretically infinite number of constituents can be inserted between them:

(8) **Write the number down (a thousand times) (on different pages) (tomorrow).**

Schreib die Zahl auf (ein tausend Mal) (auf unterschiedlichen Seiten) (morgen).

‘Schreib die Zahl (morgen) (ein tausend Mal) (auf unterschiedlichen Seiten) auf.’

This makes the connection between the particle and its verb base appear to be so weak that in the case of the analogical change in the imperative singular of strong verbs with e/i-gradation, it might not be the token frequency of the verb lemma (including prefixes and particles) which has a conserving effect. Therefore, in addition to verb lemma token frequency, a second frequency measure will be considered in the present investigation, viz. the lemma token frequency of the base verb. This term is borrowed from Lüdeling (2001; cf. Bybee 1995: 437; Müller 2002) to refer to a verb, including inseparable prefixes (cf. 6), but excluding separable particles (cf. 7 and 8). For simplex verbs like *geben* ‘give’ and prefixed verbs like *vergeben* ‘forgive’ (cf. 6), the base verb lemma token frequency is identical with the verb lemma token frequency, i.e. all tokens of all forms in the paradigm. For particle verbs like *zurück-gaben* ‘give back’, however, the two measures yield different numbers: while for the measure verb lemma token frequency all tokens of all forms in the paradigm of *zurückgeben* are counted, for base verb lemma token frequency all tokens of all forms in the paradigm of the base verb *geben* are counted.

<table>
<thead>
<tr>
<th>Table 4: Relation between verb lemma and base verb lemma depending on morphological structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>morphological structure</td>
</tr>
<tr>
<td>-------------------------</td>
</tr>
<tr>
<td>simplex verb</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>prefixed verb</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Thus, as Table 4 shows, no difference is made between the base verb lemma token frequency of, for example, the lexemes *geben* and *zurückgeben* or *wiedergeben* because it is assumed that the conserving effect of token frequency does not depend on whether a speaker says *Gib mir das Buch!* ‘Give me the book!’ or *Gib mir das Buch ... zurück/ wieder!* ‘Give me the book ... back!’.

### 1.3.3 Type frequency

Type frequency again seems to be a rather unproblematic concept; it denotes the number of items which a pattern applies to (Bybee 1985; 2007), i.e. in the present investigation the number of verbs which a particular inflection pattern applies to. However, also with regard to this frequency measure, attention must be paid to how the group of German particle verbs is counted. For example, Joan Bybee criticised a paper by Clahsen and Rothweiler (1992) on the productivity of German past participle formation patterns because the authors counted particle verbs as distinct types instead of subsuming them under their simplex bases; she claims that this method “artificially inflates the type frequency of the irregular verbs” (Bybee 1995: 435). In order to avoid a similar bias in the present investigation, type frequencies will be presented separately for all lexical verbs, i.e. counting particle verbs as distinct types, and for all base verbs, excluding particle verbs.

### 1.3.4 Intraparadigmatic frequencies

The last group of frequency measures is not entirely new, but they have not been considered in many studies, most probably because they are only relevant in languages like Russian or other Slavic languages with “extensive paradigms” (Corbett et al. 2001: 220). German similarly has paradigms with a relatively high number of distinct forms (see Table 1); therefore, additional frequency measures should be annotated in order to capture which frequency exactly exerts a conserving effect. Corbett and his colleagues have studied irregularity in plural inflection in Russian nouns on a scale from full suppletion to irregularities in stress and identified three levels of frequency:

> There are three morphological levels which might be relevant for frequency effects.

> The first is the level of the lexeme as a whole; the second is the level of the subparadigm of the lexeme; and the third is the level of the individual cell. (Corbett et al. 2001: 220)
Thus, in addition to lemma token frequencies on the first level, intraparadigmatic frequencies may capture the propensity of paradigm members to become subject to analogical levelling more accurately in languages with rich verb inflection paradigms like German. Evidence from psycholinguistic experiments suggests that speakers keep track of how often they encounter and use individual forms in the paradigm of a lexeme, such that, for example, the singular forms of “singular-dominant” nouns - which, on average, occur more frequently in the singular than in the plural - are recognised faster than the plural forms of the same nouns (e.g. Baayen et al. 2003; Baayen & Schreuder 2003; New et al. 2004; Sereno & Jongman 1997).

On this “subparadigm” level, Corbett et al. counted whether plural forms in different noun paradigms occurred more frequently than in the corpus as a whole; this measure had a significant effect on irregularity in that “nouns which have an irregularity involving a split between singular and plural will tend to be nouns which occur frequently in the plural” (2001: 219-220). On the third level, the authors examined whether a high relative intraparadigmatic frequency of a particular cell, e.g. the genitive plural, is correlated with the irregularity ranking of the nouns which exhibit this “cell anomaly”. Corbett et al. found no significant effect for this third level; however, in the present investigation, where analogical levelling affects only one form in the paradigm of strong verbs with e/i-gradation, an unusually high proportion of this cell may be related to the respective verb’s tendency to resist levelling (another possible conserving effect).

<table>
<thead>
<tr>
<th>Present</th>
<th>Indicative</th>
<th>Subjunctive</th>
<th>Imperative</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Singular</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>gebe</td>
<td>sterbe</td>
<td>gebe</td>
</tr>
<tr>
<td>2</td>
<td>gibst</td>
<td>stirbst</td>
<td>gebest</td>
</tr>
<tr>
<td>3</td>
<td>gibt</td>
<td>stirbt</td>
<td>gebe</td>
</tr>
<tr>
<td><strong>Plural</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>geben</td>
<td>sterben</td>
<td>geben</td>
</tr>
<tr>
<td>2</td>
<td>gebt</td>
<td>sterbt</td>
<td>gebet</td>
</tr>
<tr>
<td>3</td>
<td>geben</td>
<td>sterben</td>
<td>geben</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Past</th>
<th>Indicative</th>
<th>Subjunctive</th>
<th>Present Participle</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Singular</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>gab</td>
<td>starb</td>
<td>gäbe</td>
</tr>
<tr>
<td>2</td>
<td>gabst</td>
<td>starbst</td>
<td>gäbst</td>
</tr>
<tr>
<td>3</td>
<td>gab</td>
<td>starb</td>
<td>gäbe</td>
</tr>
<tr>
<td><strong>Plural</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>gaben</td>
<td>starben</td>
<td>gäben</td>
</tr>
<tr>
<td>2</td>
<td>gibt</td>
<td>starbt</td>
<td>gäbt</td>
</tr>
<tr>
<td>3</td>
<td>gaben</td>
<td>starben</td>
<td>gäben</td>
</tr>
</tbody>
</table>

Note: Grey shaded cells denote paradigm forms with the stem vowel e.

Four intraparadigmatic frequency measures will be used in the present study; the first of these is the INTRAPARADIGMATIC TYPE FREQUENCY of the stem vowel e. Table 5
illustrates that there are 12 forms in the paradigm of *geben* which have this stem vowel, whereas there are only 11 forms in the paradigm of *sterben* with this stem vowel because this verb forms the past participle with the stem vowel *ö*. A higher (absolute) intraparadigmatic type frequency of forms with the stem vowel *e*, as observed for the verb *geben*, may reinforce a verb’s tendency to succumb to levelling towards this stem vowel in the imperative singular, hereby counteracting the conserving effect of high lemma token frequency.

In addition, the potential influence of the RELATIVE INTRAPARADIGMATIC TOKEN FREQUENCIES OF THE STEM VOWELS *I* AND *E* will be assessed. For example, the third person singular present indicative form of *geben* has a very high token frequency because it is part of the German existential construction which, in contrast to the English equivalent, is only used in the singular: *es gibt X* ‘there is/are X’. Conversely, the verb *sterben* is very often used in the preterite and past participle forms *starb(en)* and *gestorben* as in “someone (has) died”. The token frequencies for individual forms in the paradigms of strong verbs with *e/i*-gradation with the stem vowel *e* or *i* will be extracted from a corpus and entered into a table (see Table 13 and Section 5.2.2 for further details). Within this table, the relative token frequencies for all forms in the paradigm of one verb with the stem vowel *e* and *i*, respectively, are added up. This yields relative intraparadigmatic token frequencies for the verb *geben* of 49.15 % for the stem vowel *i* and 30.72 % for the stem vowel *e*. In contrast, the frequencies for the verb *sterben* are 10.10 % for the stem vowel *i* and 20.63 % for the stem vowel *e*. Assuming that high intraparadigmatic token frequency can exert a conserving effect on a verb’s inflection, comparable to that of high lemma token frequency, the probability for the conservation of the traditional imperative singular formation with the stem vowel *i* can be expected to be higher in verbs like *geben* with a higher intraparadigmatic token frequency of *i*-forms than in verbs like *sterben* with a higher intraparadigmatic token frequency of *e*-forms.

The table also yields the values for the last frequency variable considered in the present investigation, viz. the RELATIVE INTRAPARADIGMATIC TOKEN FREQUENCY OF THE IMPERATIVE SINGULAR. As explained above, verbs whose imperative singular form is used very infrequently in comparison to other paradigm forms may succumb to analogical levelling more readily than verbs which are used in the imperative singular very often. The verbs *geben* and *sterben* both have a very low relative intraparadigmatic token frequency of the imperative, at 0.09 % and 0.55 %, respectively (see Table 13). By contrast, the verb

---

4 Some verbs have past tense forms with *o* (indicative) or *ö* (subjunctive). The sole important type frequency here is that of the stem vowel *e*.
hersehen has a relatively high value of 4.37%; the form *sieh her!* can therefore be expected to resist analogical levelling more strongly than *gib!* and *stirb!*.

1.4 Other factors with a potential influence on imperative formation

Studies of ongoing language change present an inherent difficulty insofar as they need to tease apart evidence of diachronic change from patterns of synchronic variation (cf. Mair 2008: 1110). Apart from the frequency variables explained above, other factors may have an impact on variation in the formation of the imperative singular of strong verbs with e/i-gradation, among them most notably persistence effects, dialectal differences and other socio-linguistic variables.

1.4.1 Language users as creatures of habit

In trying to account for repercussions of synchronic variation in the imperative singular formation of strong verbs with e/i-gradation, the present study will take into account the potential influence of a context-related variable which is here referred to as *persistence* (Szmreczanyi 2005; 2006). This term denotes the tendency that speakers’ choice between two variants of a variable, for example between the traditional irregular and analogical variants of the imperative singular of strong verbs with e/i-gradation, could be (partly) explained on the basis of their previous choices concerning the same variable shortly before. In order to avoid confusion with related concepts of repetition in language, such as *priming* and *recency*, these terms are explained in a separate section. The study does not aim to answer the question whether recency has an influence on variation in imperative formation, but the potentially disturbing effect of priming is of importance for the experimental study (Chapter 6).

**PERSISTENCE**

In his dissertation project, Benedikt Szmreczanyi (2005; 2006) examined five well-known cases of synchronic (morphosyntactic) variation in spoken English and tried to explain speakers’ use of a particular variant of a variable on the basis of the presence or absence of identical or related words or patterns in the previous context. He distinguishes between two types of *persistence* and illustrates them by the following two examples:

(9) Matt’ll find this out, and, I mean, we’ll get involved in it. (Corpus of Spoken American English, text “Bank Products”)
(10) You go look, and every horse’s hoof is shaped different. It doesn’t matter. Every horse is gonna have a little different shape. (Corpus of Spoken American English, text “Actual blacksmithing”)

(examples (2) and (3) in Szmreczanyi (2005: 117))

Szmreczanyi refers to (9) as an instance of $\alpha$-persistence: if a speaker has the choice between two or more semantically equivalent variants of a variable, the fact that he/she has used one variant before will make it more likely that this same variant is reused. Sentence (10) is an example of $\beta$-persistence in which a speaker likewise faces the choice between two or more semantically equivalent variants of a variable and his/her previous use of a “linguistic pattern which is not necessarily variable itself, but which shares one or more syntactic, morphological, or lexical properties with one of the variable’s variants” increases the probability for this variant to be used (Szmreczanyi 2005: 116-117; 2006: 2). It seems that in both cases, the speakers can freely choose between two variants of the variable “future marker” in English, be going to and will: the sentences could also have been formulated as “we’re going to get involved in it” and “every horse will have a little different shape”, respectively. Szmreczanyi argues that this choice is not entirely free but that it partly depends on the previous use of the variant will in sentence (9) and of a linguistic pattern containing the auxiliary verb go in sentence (10). Szmreczanyi is by no means the first researcher to investigate these patterns in corpus data. Earlier studies, in which variation has been partly attributed to persistence (review in Szmreczanyi 2006: 28-38), have been concerned with morphosyntactic variables, such as active-passive alternation (Weiner & Labov 1983; Estival 1985), dative alternation and particle placement in English (Gries 2005), or pronoun choice in Montreal French (Sankoff & Laberge 1978a). Studies of morphological variation include investigations of variable plural-marking in Puerto Rican Spanish (Poplack 1980) and Brazilian Portuguese (Scherre & Naro 1991) and past-tense marking in Nigerian Pidgin English (Poplack & Tagliamonte 1993; 1996).

Analogical levelling in the formation of the imperative singular of German strong verbs with e/i-gradation will be argued to have moved past a stage of pure synchronic variation; however, persistence may still affect the choice of imperative singular variants to a non-negligible degree. In fact, a close reading of some of the corpus texts used in the present investigation suggested that $\alpha$-persistence might play a role in situations when speakers use the traditional or analogical imperative singular variant of a strong verb with e/i-gradation presumably because they have used the same variant shortly before in the same verb or in an-
other strong verb with e/i-radation. In the following example, the speaker may simply reuse the form *nehme* in the second sentence:


‘Threaten the banker and take the bag. You are inside? Then take the shovel with Clever and dig up the hatchet.’

Similarly, the author of example (12) may be triggered to use the analogical imperative singular variant of the verb *nehmen* ‘take’ on the basis of his or her previous use of the analogical variant of *aufbrechen* ‘break open’.


‘When you are there, break open the untidy desk with the skeleton key and take the key for the harbour office (tool) from the drawer.’

Furthermore, β-persistence in Szmreczanyi’s terms may explain the use of the traditional variant of *werfen* with the stem vowel *i* in (13) and the analogical variant of the same verb with the stem vowel *e* in (14).


‘Pick up the nut and turn right. There you see a switch. Throw the nut towards the switch; if you hit (it), the machine turns off.’


‘When the rider rides left, run to Sandyman’s mill and hide in the mill yard. Throw stones towards the village and when the rider rides there, you run away across the bridge.’

In (13), the speaker uses unsuffixed variants of the imperative singular of verbs that are not strong verbs with e/i-gradation before the target imperative; in (14) suffixed variants of such verbs precede the target form. In both cases, the target imperative is preceded by a variant of

---

5 All examples in this section are quoted from the Walkthrough Corpus (see 3.1.2). In order to protect personal data of the authors of corpus texts, only the corpus-internal abbreviation of the respective text is given.
“a linguistic pattern which is not necessarily variable itself, but which shares one or more syntactic, morphological, or lexical properties with one of the variable’s variants” (Szmreczanyi 2005: 116-117; 2006: 2), the shared property in this case being suffixation. Thus, like α-persistence, β-persistence is expected to influence the suffixation of the imperative singular of German strong verbs with e/i-gradation in the present study. In addition, β-persistence can be assumed to indirectly affect stem vowel choice for imperative singular forms of these verbs. A suffixed version of the traditional irregular imperative variant with the stem vowel i does not exist (e.g. *gibe, *wirfe); thus, if speakers are triggered to use a suffixed imperative singular variant of a strong verb with e/i-gradation by their previous use of suffixed imperative variants of other verbs, as in example (14), the only suffixed imperative variant they can choose is the one with analogical stem vowel, for example wɐɾfe.

PRIMING AND RECENCY

Before Szmrecsanyi introduced the term persistence⁶, repetition patterns in corpus data such as the ones outlined in the previous section had been referred to as “priming” or “recency”. These terms describe patterns and processes which are related to but different from persistence; they are briefly explained in the following in order to underline their unsuitability for the present investigation.

Psycholinguistic studies have repeatedly shown that the reactions of participants in comprehension and production experiments can be explained on the basis of priming. In lexical decision and naming tasks, for example, semantic priming is held responsible for the observation that the word nurse is generally recognised faster and more accurately when it follows the presentation of the word doctor from the same semantic field than when it follows a semantically unrelated word like butter (Meyer et al. 1975; cf. Hutchison et al. 2013). Similar priming effects have been found for stimuli which are orthographically (Evett & Humphreys 1981; Ziegler et al. 2000), phonologically (Humphreys et al. 1982), orthographically and phonologically (Tanenhaus et al. 1980), morphologically (Feldmann 1994; Rueckl et al. 1997; Sonnenstuhl et al. 1999) or syntactically related (Ledoux et al. 2007; Fine & Jaeger 2013). The term production priming refers to the finding that previous exposure to a prime not only facilitates processing of a semantically, lexically, morphologically, or (morpho-) syntactically related target stimulus. When participants have been exposed to a related stimulus previously in an experiment, i.e. when they have “primed”, they are also more likely to produce a par-

---

⁶ More precisely, he reintroduced the term which was (and is still occasionally) used in psycholinguistics for instances of syntactic production priming (e.g. Bock 1986; Segaert & Hagoort (to appear)), as opposed to syntactic priming in language processing, referred to as “priming”.

26
ticular word or phrase (e.g. Hartsuiker & Westenberg 2000; Smith & Wheeldon 2001). Priming effects have been attributed, amongst others, to spreading activation: The idea is that linguistic items are represented by nodes in memory which are connected to related nodes by “associative pathways”. When two linguistic items are connected in this way, then the activation of one of them triggers preactivation of the second, so that when the latter is presented, the threshold for its activation is lower (Cortese & Balota 2012). Even though priming may be the cause for persistence, i.e. an observed “surface parallelism in corpus data” (Szmrecsanyi 2006: 3), this cannot be proven on the basis of the corpus data alone; therefore, this term is avoided in the corpus studies presented in this paper (however, see Chapter 6 about the relevance of priming for the experimental study).

Recency is originally described in psycholinguistic studies of working memory; in contrast to investigations of priming effects in language production or processing, these studies examine whether temporal distance between exposure to and recall of stimuli affects recall accuracy. Baddeley summarises studies conducted since the 1960s which discovered such a recency effect: “If a subject is presented with a list of words for immediate free recall, there will typically be extremely good recall of the last few items presented, the so-called recency effect. If recall is delayed by a few seconds of interpolated counting, then the recency effect disappears” (1994: 9; emphasis original). Moreover, Szmreczanyi himself uses “recency” in a corpus-linguistic sense to refer to the textual distance between a previous and current instance of a linguistic variable. Its effect interacts with that of persistence, i.e. the presence of identical or related material in the previous context (2005: 119-120). Thus, in order to refer to the observation that speakers in the present study might reuse imperative singular forms and patterns without implying short distance (temporal or textual) between them, “recency” was similarly abandoned in favour of Szmreczanyi’s “relatively neutral term persistence” (2006: 3).

1.4.2 Dialect variation
As explained above, the analogical levelling process affecting imperative singular formation of German strong verbs with e/i-gradation is assumed to be a case of ongoing language change rather than an instance of linguistic variation. Nevertheless, it may be affected (accelerated or inhibited) by dialectal variation in the formation of the imperative singular of these verbs.

In the literature on German dialect morphology, information about the formation of the imperative singular of strong verbs with e/i-gradation is either not provided at all (e.g. Hünecke 2012; Niebaum 1973; Stickel 1997) or on various scales. While Žirmunski’s
Deutsche Mundartkunde (2010) and Russ’ Dialects of Modern German (2013a) offer comprehensive descriptions of the phonology and morphology of all German dialects spoken in the territory of the modern Federal Republic of Germany (FRG), other authors examine the features of more or less locally restricted dialects and vernaculars in the same territory. A coherent picture of the wealth of information presented in all these publications emerges most clearly when the distribution of patterns of imperative singular formation of strong verbs with e/i-gradation is illustrated in a map of the modern FRG:

The division of dialect areas in Figure 1 largely follows the widely-used maps in Mitzka (1943), Wiesinger (1983) and the Neue Duden Lexikon (Adam 1989), but it was slightly modified on the basis of information from chapters in Russ (2013a). The boundary lines are traditionally based on phonological isoglosses; therefore, the Berlin area is usually included in the Upper Saxon dialect region. The distribution of patterns of imperative singular formation of strong verbs with e/i-gradation, however, makes it more reasonable to group this area within the East Low German dialect (cf. Schönfeld 2013: 95).

The legend indicates that for some regions in Germany, several publications agree on which of the stem vowels i or e is used for the formation of the imperative singular of strong verbs with e/i-gradation in the respective dialects or vernaculars spoken there, i.e. whether speakers say, for example, gib mir das Buch! or geb(e) mir das Buch. According to Žirmunski (2010: 563-564) and chapters in Russ (Russ 2013b: 355; Rowley 2013: 431; Wiesinger 2013: 493), speakers of Swabian and Bavarian (North and Central) use the stem vowel i; this is supported by Gebhardt (1968: 292), Merkle (1975: 48) and Steininger (1994: 88-94) for Bavarian. Žirmunski further explains that vowel gradation between e and i is completely eliminated only in a group of dialects in the middle West (2010: 563), i.e. the stem vowel e is used in all singular forms, including the singular imperative, in Central Franconian, Hessian and the Palatinate. Out of the chapters in Russ (2013a) which cover these dialect areas, imperative singular formation is only addressed explicitly in the chapter on Hessian, likewise giving e as the stem vowel used for the imperative singular of strong verbs with e/i-gradation (Durrell & Davies 2013: 235). However, an impressive number of other publications further confirm Žirmunski’s observation: Borchers (2009) cites the imperative singular with the stem vowel e for urban dialects in Central Franconia, Hessa and the Palatinate, Münch (1970: 169) and Tiling-Herrwegen (2002: 81) for the (wider) Cologne area in Central Franconian, Bräutigam (1934: 104-105) for Mannheim and Mottausch (2002: 5) for the Lorsch area in the Palatinate.

7 Descriptions are also given for areas outside the territory, for example Higher Alemannic in Switzerland and Southern Bavarian in Austria, but these dialects are less relevant for the present investigation.
8 The map in Figure 1 was created using the programme Inkscape (version 0.91; https://inkscape.org/en/).
Figure 1: Distribution of stem vowels $i$ and $e$ used for the formation of the imperative singular of strong verbs with $\text{e/i}$-gradation in dialects spoken on the territory of the Federal Republic of Germany.
Both Lindow (1984; 1998) and Stellmacher (1983: 269) ascertain that speakers of all Low German dialects (areas 1-5 in Figure 1) likewise generally make use of the stem vowel e for the imperative singular of strong verbs with e/i-gradation. Their observations match those by Walker (2013: 19) for Frisian, Goltz and Walker (2013: 48) and Sass and Thies (2010: 53) for North Saxon, Freise (2009: 34-39) for Hildesheim in Eastphalia, Langner (1977: 76) for Wittenberg (in the East Low German region), and Schönfeld for the “semi-dialect” of Berlin (2013: 124) and the East Low German dialect in general (2013: 113-114). Finally, Spangenberg observes in two publications (1998: 177; 2013: 282) that the stem vowel e is used for imperative singular formation of strong verbs with e/i-gradation in the Thuringian dialects.

For the remaining three dialect areas, Upper Saxon, East Franconian and Low Alemanic, the literature mainly presents negative evidence. In the chapters in Russ’ edited volume, “dialects are described in contrast to New High German, or standard German” (2013a: xx). Thus, rather than indicating that the authors did not have sufficient data or were not interested in the matter, the lack of an explicit description of imperative singular formation of strong verbs with e/i-gradation can be taken to mean that the same stem vowel is used in the dialects as in Standard German, namely the vowel i. For East Franconian dialects, Wagner similarly supplies negative evidence by observing that differences from the norms of the standard language only occur in the formation of the first person singular of strong verbs with e/i-gradation, i.e. producing ich nimm instead of ich nehme (1987: 70-71). If the formation of the imperative singular of these verbs deviated from standard usage, he would have mentioned it as well; thus, speakers of East Franconian dialects most probably use the stem vowel i for forming the imperative singular of strong verbs with e/i-gradation. In the case of Upper Saxon, the dictionary part of Albrecht’s (1881) description of the urban dialect of Leipzig lists strong verbs with e/i-gradation with the stem vowel i in the imperative singular. Admittedly, Albrecht’s publication is old, but Bergmann explained recently that Saxon today is “characterized by the loss of almost all the most localizable dialect features and the retention of the features which are more widespread” (2013: 309). The retained features he mentions are exclusively phonological and lexical; if the Upper Saxon dialects are moving closer to the German Standard in other regards, imperative singular formation of strong verbs with e/i-gradation will employ the stem vowel i. Interestingly, one publication by Harald Noth which explicitly addresses the imperative singular in a Low Alemannic dialect observes that both stem vowels of strong verbs with e/i-gradation are used for its formation. Noth (1993) has studied the dialects of the Kaiserstuhl region and its surroundings, approximately 20 kms North-West of Freiburg im Breisgau. He explains that the imperative singular is generally formed like the
first person singular and provides two lists of verbs whose first person singular is formed with the stem vowel \( i \) or \( e \), respectively:

<table>
<thead>
<tr>
<th>stem vowel ( i )</th>
<th>stem vowel ( e )</th>
</tr>
</thead>
</table>

Although Noth does not state this, as a matter of fact, his division hints at a frequency pattern: the verbs he lists with an imperative singular formed with the stem vowel \( e \) are exclusively low frequency verbs of German (see Appendix A), whereas the ones listed with the stem vowel \( i \) are highly frequent verbs (except for melken and verderben). Thus, Noth’s dialectological study can be taken as first evidence that change in the formation of the imperative singular of strong verbs with \( e/i \)-gradation is underway in Low Alemannic and that it is characterised by the conserving effect of high token frequency outlined in Bybee and Thompson (1997): low frequency verbs occur with an analogical imperative singular, while higher frequency verbs are used only with the traditional (standard) stem vowel \( i \).

Thus, change in the formation of the imperative singular of strong verbs with \( e/i \)-gradation may well be influenced by dialect morphology in the German dialects. Whereas it may be initiated and spread more easily in the dark grey areas in North and Central Germany in Figure 1, where speakers use the analogical stem vowel \( e \) in the dialect, it can be expected to be inhibited in the white areas in Southern Germany and Upper Saxon and East Franconian, where speakers generally use the traditional stem vowel \( i \), both in Standard German and in the dialect. Finally, Noth’s (1993) account of Low Alemannic morphology hints at the possibility that the ongoing change in imperative formation is simply not documented in many dialect grammars, yet.

An exhaustive study of the formation patterns of the imperative singular of strong verbs with \( e/i \)-gradation in all German dialect areas cannot be the aim of the present investigation. Rather than that, it will examine the assumed change-in-progress on the basis of a corpus study, hereby testing whether, apart from the influence of frequency and other factors, the change is affected by dialectal differences in imperative formation, as suggested by this overview. In search for further evidence of whether change occurs throughout Germany or only in parts of the country, the reactions of speakers from two different regions to imperative singular forms with both stem vowels will be compared in the experimental study.
1.4.3 Gender patterns

It has been shown in a number of studies, first by William Labov and later by other researchers, that female and male speakers often behave differently towards forms involved in language variation or change. According to Labov’s “principles of sexual differentiation”, women lead in using innovative variants of a variable in change “below the level of social consciousness”, i.e. when speakers are not aware of change (principle II in Labov 1990: 215; renamed principle 4 in Labov 2001: 292). Among other examples quoted by Labov, women were in advance of men in vowel raising and/or fronting in New York City and Detroit (Labov 1990: 216). On the other hand, according to principles I and Ia (principles 2 and 3 in Labov 2001), “women lead in ... the elimination of stigmatized forms” (Labov 1990: 212) in stable linguistic variation or change “from above the level of consciousness”, i.e. when speakers are aware of the change. Sociolinguistic studies of stable linguistic variation quoted by Labov (1990; 2001) include investigations of the realisation of unstressed -ing as [ɪŋ] or [ɪŋ] and the realisation of interdental [θ] and [ð] in English (Labov 1990: 211); change from above is found, for example, by Milroy and Milroy (1985) in the raising of /e/ from [a] to [e] in Belfast. These conflicting patterns, also termed the Gender Paradox, make it difficult to set up hypotheses for the phenomenon under investigation in the present study: if analogical leveling in the imperative singular of strong verbs with e/i-gradation is indeed a case of ongoing language change of which speakers are not aware, then women would be expected to use a higher rate of analogical imperative variants than men. If, on the contrary, the imperative singular is a variable of stable linguistic variation or if it is changing and speakers are aware of this change, men would be expected to use more of the analogical imperative variants. Arguably, none of these patterns may apply to the present case of analogical levelling: Labov mentions very few examples of morphological variation and change from previous research in which such gender differences apply; his main body of evidence comes from studies of phonological or phonetic variables. In part due to the apparent discrepancy between Labov’s principles, many sociolinguists have adopted a different methodology: instead of splitting a community up into predefined, perhaps superficial, groupings, they start out from the notion that “gender is a social identity, [which] emerges, like all social identities, in the ways we interact with others” (Meyerhoff 2011: 232, see pages 231-246 for example studies). The present study cannot delve into an elaborate discussion of the influence of social identities on imperative singular formation in strong verbs with e/i-gradation to the extent that previous studies have done. Even if the Labovian method may be simplistic, in the present investigation speaker gender (or participant gender in the experimental study) will be taken into ac-
count as a predictor in order to uncover whether men and women indeed use different variants of the imperative singular or react differently towards the competing variants.

2 Methods

Along the lines of earlier usage-based studies, the assumed change-in-progress in the formation of the imperative singular of strong verbs with e/i-gradation will be analysed by means of corpus-linguistic studies, outlined in Section 2.1. An experimental study will try to supply converging evidence for the explanation of frequency patterns in corpus data on the basis of the cognitive processes of entrenchment (Section 2.2). Datasets obtained in the major corpus study and from the experimental study will be analysed by means of mixed-effects regression models. The general procedure of such statistical methods will be explained in Section 2.3; details of the software implementation are outlined in Appendix B.

2.1 Corpus-linguistic studies

In three separate corpus studies, the present investigation will try to answer the questions whether the higher productivity of other verb inflection classes is responsible for analogical levelling in the imperative singular of German strong verbs with e/i-gradation, whether this case of analogical levelling is an example of language variation or change(-in-progress) and which factors have an influence on the change, and why the stem vowel alternation in the paradigm of these verbs is levelled in the imperative singular but conserved in the forms of the second and third person singular present indicative.

CHAPTER 4: THE DIRECTION OF CHANGE

Analogical levelling has been shown to usually proceed in the direction of a productive formation pattern, and the type frequency of patterns is claimed to be a major determinant of productivity (see 1.1.1). The first corpus study will address the question whether analogical levelling in the imperative singular formation of strong verbs with e/i-gradation can likewise be explained on the basis of the type frequency of verb inflection classes in which the imperative singular is formed without a stem vowel alternation. Instead of relying on existing (arguably outdated) figures for the type frequency of German verb inflection classes, this study will present distributions from recent corpus data.
CHAPTER 5: THE TRAJECTORY OF CHANGE

It was mentioned before that analogical levelling does not seem to affect all strong verbs with e/i-gradation to the same extent: whereas speakers are astonished when encountering the analogical imperative variants of verbs like *geben* ‘give’ or *nehmen* ‘take’, they are equally surprised that verbs like *verbergen* ‘conceal’ belong to this verb class and thus traditionally have an imperative singular form with the stem vowel *i*: *verbirg*. The first corpus study will test which factors have an influence on a verb’s propensity to change. To this end, observations of the use of the imperative singular of strong verbs with e/i-gradation will be extracted from a corpus (see Section 3.1.2) and annotated for a variety of factors such as the above-outlined frequency, persistence and sociolinguistic variables. Statistical analyses will show which of these factors have a significant influence on stem vowel choice and suffixation of the imperative singular of the strong verbs with e/i-gradation. Unfortunately, due to a lack of meta information, the impact of the sociolinguistic factors could not be exhaustively assessed in the corpus study; they will therefore be incorporated in the design of the experimental study.

CHAPTER 7: RESISTANCE TO CHANGE

This third corpus-linguistic study attempts to explain why analogical levelling does not affect all forms in the paradigm of strong verbs with e/i-gradation to the same extent: the second and third person singular forms of these verbs do not seem to change although they exhibit the same stem vowel alternation which is replaced by analogical formations in the imperative singular forms in the same verb paradigms. This study will test whether the token frequency of the second and third person singular present indicative forms is higher than that of the imperative singular form, so that their conservation can be explained on the basis of their higher entrenchment in speakers’ mental lexicons. In order to explore the explanatory potential of the frequency-based approach against alternative accounts of partial paradigmatic change, token frequencies for the three paradigm forms will be extracted from three very recent corpora of spoken and written German.

2.2 Experimental study

Although *cognitive corpus linguistics* (Arppe et al. 2010) has by now become an established methodology, cognitive explanations of patterns discovered in corpus data are often contested. Criticism is typically voiced in the form of a plea for “methodological pluralism” or “converging evidence”, as in Doris Schönefeld’s introduction to the eponymous volume:
Corpus data can suggest assumptions ... [but] the corpus does not give first-hand information on what was actually going on in the formulation process of the written and spoken text respectively. In order to tap the procedures going on in language production and comprehension, experimentation should follow in the research agenda to test whether the hypotheses formulated on the basis of corpus findings can be corroborated. (Schönefeld 2011: 24)

This recommendation has been accepted in the present investigation: in pursuit of converging evidence for the prevalent explanation of the conserving effect of high token frequency in analogical change on the basis of mental entrenchment, an experimental study has been conducted. In this study, the entrenchment of the competing imperative singular variants of the German strong verbs with e/i-gradation is measured in terms of reactions by native speakers of German. Participants not only processed the imperatives which were embedded in stimulus sentences, they were also asked to repeat the sentences. Reactions are thus measured in terms of response times and recall accuracies of the presented stimulus imperatives. For a better understanding, the assumptions of the experiment design, details about method and data and the results of the experimental study are presented collectively in Chapter 6.

2.3 Statistical Analysis

The data obtained in the two major studies, i.e. the second corpus study and the experimental study, will be analysed using multifactorial statistical methods. In contrast to monofactorial methods which test whether there is a relation between maximally two variables, multifactorial regression models allow for testing whether several predictor variables have an effect on a dependent variable at the same time. Their results are, therefore, more accurate. For example, in an investigation of particle placement in English, Diessel and Tomasello (2005) compared the results of a monofactorial and a multifactorial analysis of the same corpus dataset. The monofactorial method suggested that four predictor variables had a significant effect on particle placement; however, the multifactorial regression analysis revealed that the influence of two of these predictors was overestimated (cf. Gries 2003; Bresnan et al. 2007).

More precisely, mixed-effects regression models were used in these analyses in order to generalise over the idiosyncratic variation patterns of individual speakers or linguistic items. The following sections will briefly outline and illustrate the working principle of linear and logistic regression and of the mixed-effects variants of these regression models and why
they were used in the present study. Details of how regression models were implemented in the statistics programme R and how results are reported in regression tables and graphs are presented in Appendix B.

LINEAR AND LOGISTIC REGRESSION
Regression analyses are multifactorial statistical methods by which the influence of several predictor variables (ratio-scaled, ordinal or nominal) on a dependent variable can be assessed; thereby, they yield answers to a variety of questions:

i) which of the predictor variables entered into a regression model has an influence on the dependent variable,

ii) the direction in which the predictors influence the dependent variable,

iii) in addition to their individual influence, do any of the predictors have an (significant) interactive effect on the dependent variable,

iv) is this influence significant, i.e. can it be generalised beyond the data sample (thereby determining whether some variables/interactions may be better or stronger predictors than others),

v) how much of the variation in the dependent variable is explained by a) each of the predictor variables and b) all of the predictors considered together,

The terms linear and logistic regression derive from the quality of the dependent variable in the particular analysis: Linear regression models are used for ratio-scaled dependent variables (cf. Gries 2013: 261-293); in the present paper, (the mixed-effect variant of) this regression type was used for the analyses of reading times obtained in the experimental study (Chapters 6.4.2 and 6.5.2). Logistic regressions are used when the dependent variable is a nominal or ordinal variable with two or more levels (cf. Gries 2013: 293-316); (mixed-effects) binary logistic regressions were fitted on the corpus data of stem vowel and suffixation choice of imperative singular forms of strong verbs with e/i-gradation (Chapter 5.3.2) and on the recall accuracy data obtained in the experimental study (Chapters 6.4.3 and 6.5.3). The chief difference between the two procedures is the way in which the dependent variable is handled by the statistical model: in linear regressions, the original values of the dependent variable are retained (e.g. reading times in msec); in binary logistic regression, one of the levels of the dependent variable is treated as the reference level and the other as the response level (e.g. incorrect recall and correct recall). The linear model tests whether a change in one

---

9 Binary logistic regressions are used for dependent variables with exactly two levels (“dichotomous” variables), as opposed to ordinal regressions for dependent variables with two or more ordered levels and multinomial regression models for dependent variables with more than two levels (cf. Gries 2013: 316-324)
of the predictor variables is associated with a change in the value of the dependent variable, for example, whether reading times increase with stimulus word frequency or whether female participants read faster than male participants. In contrast, the logistic regression tests whether a change in the predictor is associated with the chances of one of the levels of the dependent variable to occur, for example, whether chances for the correct recall of a stimulus word increase with its word frequency or whether chances for correct recall are higher among female than among male participants (see questions i and ii above).

The individual effects of two or more predictors on a dependent variable may add up, for example, when reading times increase with decreasing stimulus word frequency or when they are generally higher in male than in female participants. If the predictors do not have an interactive effect, this would mean that the reading times for male participants reading stimuli with a low word frequency can be calculated by adding the reading time coefficients for male participants and that of low frequency stimuli; similarly, the reading times of female participants reading low frequency stimulus words corresponds to adding the reading time coefficient for female participants and that of low frequency stimulus words (see left panel of Figure 2).

![Figure 2: Main effects and interactive effect of stimulus word frequency and participant gender on reading times (simulated data)](image)

By contrast, if the two predictors have an interactive effect on the dependent variable, this interactive effect cannot be predicted from the individual effects of the single predictors (see question iii): the increase in reading times with decreasing stimulus word frequency may be stronger for male than for female participants. In this case, the sum of the reading time coefficients of female participants and low frequency stimuli would be higher than observed, and
the sum of the coefficients for male participants and low frequency stimuli would be lower than observed in the data, as illustrated in the right panel of Figure 2.

Both regression types are methods of inferential statistics: a researcher practically never has the chance of analysing an entire population, for example obtaining reading times from all speakers in Germany. Instead, they typically obtain reading times from a carefully selected sample of speakers. Regression models compute significance values for effects which allow the researcher to say whether the statistical results of an analysis are true only for the analysed data sample or whether they can be generalised to the population from which the sample was drawn (see question iv).

The process of selecting the final regression model for an analysis which captures as much of the variation in the dependent variable as possible, i.e. which includes all and only the predictors and interaction terms that are necessary, can work “backward” or “forward” (Gries 2013: 260). The backward selection procedure starts out from a full regression model with all possible predictors and interactions between them (a “maximal” model in Gries’ terms); predictors (first interactions, then single) are removed from the model one after another until none of the predictors can be removed anymore without a concurrent decline in model quality. In the forward selection process used in the present analyses, a baseline regression model is specified which contains only the model mean (and random variables; see next section). In successive steps, single predictors and interaction terms are added if their addition improves model quality, i.e. if the resulting regression model explains significantly more of the variation in the dependent variable than the baseline regression or a regression without the respective predictor or interaction term. By adding predictors to or subtracting them from a regression model and comparing the resulting regression models, the researcher is also able to answer question (v) as to how much of the variation in the dependent variable is accounted for by a) any and b) all of the predictors in the model.

MIXED-EFFECTS REGRESSION

The so-called mixed-effects regression models are a derivative of conventional regression models whose name derives from the fact that the influences of fixed factors (hitherto predictors) and random factors on a dependent variable are determined by the same regression model. The chief difference between fixed and random factors lies in the question whether they are “repeatable”: “A factor is repeatable if the set of possible levels for that factor is fixed, and if, moreover, each of these levels can be repeated” (Baayen 2008: 260; emphasis added). For illustration, Baayen uses the example of an auditory lexical decision experiment
in which white noise may or may not be added to the stimulus. Once data have been obtained for these two levels of the factor and analysed in a regression model, it is still possible to add new data points to the same regression because the new observations would fall into either of the existing levels of the factor, “white noise” or “no white noise”. Random factors, such as experiment participants or stimulus items, are not repeatable: “items and subjects are sampled randomly from populations of words and participants, and replicating the experiment would involve selecting other words and other participants” (Baayen 2008: 260; emphasis added). If an existing dataset of observations from Baayen’s simulated experiment had been analysed in a regression model which accommodates for individual differences in response latency between the participants in the sample, observations for a new set of participants could not be analysed in this same regression model because the levels of the factor “participant” would not correspond to any of the existing levels of this factor in the previous regression analysis.

In comparison to conventional regression models, in practice, mixed-effects regressions involve an additional prestage in which the individual effects of the random factors are assessed. In the analysis of reading times obtained in the experimental study, for example, the mixed-effects regression first determined whether individual participants are slow or fast readers. A value of zero on the x-axis in Figure 3 indicates that the participant (on the y-axis) does not read faster or slower than the average in the participant sample; values greater than zero indicate that the average reading times of a participant are above the average of the participant sample (“slow readers”), and values smaller than zero are assigned to participants whose average reading times are below the average of the participant sample (“fast readers”).

![Caterpillar plot illustrating intercept adjustments for random factor participant in the analysis of reading times obtained in the experimental study](image)

**Figure 3:** Caterpillar plot illustrating intercept adjustments for random factor participant in the analysis of reading times obtained in the experimental study
It is only after such individual intercepts (and slopes) are adjusted that the mixed-effects regression model starts to work like conventional regression analyses, i.e. it starts to determine the (interactive) effects of the fixed factors on the dependent variable. Which random factors need to be taken into account depends in part on the type of investigation, e.g. corpus linguistic or experimental study, and in part on the details of the specific investigation, for example, whether the nested structure of corpus data needs to be adjusted for (speakers in files in registers in modes; cf. Gries 2015). The particular random variables included in the present investigation are explained in Chapter 5 for the corpus analysis and Chapter 6 for the analysis of the experimental data.

3 Data

Different corpora were used as the primary source of data for the three corpus studies outlined above. The choice of a suitable corpus was based on the research questions and hypotheses which were the focus of the respective study. Reference corpora were consulted in the second corpus study as a source for values of a number of frequency variables whose influence on imperative singular formation in strong verbs with e/i-gradation was tested. Some corpora, for example the Projekt Wortschatz Universität Leipzig, served several functions in more than one of the corpus studies.

As the data basis in the experimental study and the data obtained from it are very different from the corpus-linguistic studies, Section 3.3 below only presents a brief summary of the data which were obtained during the experimental study. Details will be provided in Chapter 6.

3.1 Primary corpora

3.1.1 The DeReWo frequency list
The DeReKo (Deutsches Referenzkorpus, ‘German Reference Corpus’) is the “largest linguistically motivated collection of contemporary German texts”, provided by the Institute of German Language (IDS) “as an empirical basis for linguistic research” (Institut für deutsche Sprache 2012a). The corpus contains over 25 billion words of German (as of 15/09/2014), compiled from electronic sources and comprising literary, scientific, popular scientific and newspaper texts as well as several other text types. On the basis of this corpus, the IDS has compiled the DeReWo (Korpusbasierte Wortgrundformenliste, ‘Corpus-based lemma list’), a token frequency list of all the recorded lemmata; as of 2013, this list comprises 326,946 en-
tries (Institut für deutsche Sprache 2012b). Token frequency is measured in logarithmic classes on an increasing scale, with the determiner der/die/das ‘the’ as the reference level “0”. Thus, the higher the value in the list, the lower is the token frequency of the respective lemma in the corpus.

The DeReWo list was used as the primary source of data in the first corpus study; it served a second purpose as one of the reference corpora in the second corpus study. Finally, the compilation of stimuli for the experimental study was partly based on the frequency information in the DeReWo list.

3.1.2 The Walkthrough Corpus

The primary corpus for the second corpus study was compiled specifically for the present investigation. It contains walkthroughs, i.e. game guides for video games; this text type was chosen because, in contrast to most corpora of German, it contains a high number of imperative singular forms. In addition, the great diversity of video games provides for variation in the frequency of the verbs contained in them. The texts were extracted from the website spieletipps.de (STRÖER media brands AG 1998-2019), one of the most popular gaming websites in Germany. The corpus contains approximately 7 million word forms, tagged for parts of speech and aligned with meta information about authors of texts (where available). The website was crawled in July 2013; thus, the corpus covers a time span from 2000 (when the website was launched in its present form) to July 2013. The following sections will provide a more detailed description of the text type selection and the compilation procedure of the final Walkthrough Corpus.

Unfortunately, this data base has two major disadvantages. First, because of the type of texts contained in the walkthrough corpus, results from the analysis of the texts may not be representative of how the investigated change progresses in the German language at large (cf. Leech 2007), even during the short time span covered by the data. Second, collections of computer-mediated communication can be affected by changes in the underlying technology. In fact, Section 5.3.5 will illustrate how the diagnosis of the current state of analogical leveling in the imperative singular of strong verbs with e/i-gradation is impaired by the introduction of spell-checking in German web browsers. The experimental study, however, will corroborate the classification of the phenomenon under investigation as a change-in-progress.

3.1.2.1 In search of a corpus

This chapter could as well be called “in search of imperatives” as the most important task in the present project was to find a corpus containing a sufficient number of tokens of the im-
The imperative singular of strong verbs with e/i-gradation. In contrast to languages such as English, German as a more highly inflected language differentiates between three forms for the imperative, one in the singular and two in the plural: compare German singular *gib!*, plural *gebt!* and formal plural *geben Sie!* to English *give!* in all forms. Unfortunately, the use of the imperative singular is restricted to conversations in which speakers are very familiar with each other (Eisenberg & Kunkel-Razum 2009: 548-49) and, perhaps most importantly, contexts which invite a request or command. These conditions are not met by many corpora because they are interview- or narration-based rather than action-oriented. Especially the singular imperative form is thus used much less frequently than the plural forms. In addition, speakers of German have a variety of other constructions at their disposal to express commands or requests, such as indicative, infinitive or modal constructions (Eisenberg & Kunkel-Razum 2009: 549-550). In consequence, the problem of finding a suitable corpus for this study did not promise to be easily solved.

CORPORA OF SPOKEN GERMAN
The data base of the present study should ideally have been located in spoken discourse for at least two reasons: (i) As explained above, the use of the imperative singular in German presupposes familiarity between conversation partners, and speakers can be expected to display more familiarity in spoken than in written interaction. (ii) It has been shown previously for language change of different kinds that incoming forms are first instantiated in spoken discourse before spreading to written language (Mair 2008: 117-118). However, all pre-compiled available corpora of spoken German (as of February 2013; among others moca, Deutsch heute, DGD) proved to contain an insufficient number of tokens of the imperative singular of strong verbs with e/i-gradation. A database of spoken language could thus only have been used if imperative singular forms had been elicited in (potentially artificial) interviews.

COMPUTER-MEDIATED COMMUNICATION
In order to base the corpus study on data as close to spoken language as possible, computer-mediated communication (CMC) was considered an acceptable compromise, in case it yielded sufficient data. The term *computer-mediated communication* originally referred to communication via a computer, i.e. email and chat communication, conversation in forums and instant messaging. Its focus has widened in the last years to include new platforms like Facebook and Twitter and communication via other electronic devices, such as short messaging service (SMS) and competing services like What’s app. Except for professional emails, these forms of communication fall into what Koch and Oesterreicher (1985) refer to as “conceptionally spo-
ken language”: in addition to the medium in which language is produced, they take into account the conception underlying the communicative action. Thus, instead of a strict dichotomy of spoken and written language, they assume a continuum between these two poles on which different forms of communication are located, depending on the extent to which they are characterised by features of spoken/intimate language (“Sprache der Nähe”) or written/distant language (“Sprache der “Distanz”) (Koch & Oesterreicher 1985: 17; 1994; cf. Biber 1991: 121-129). For example, a scientific talk is characterised as “conceptionally written” because it is typically pre-formulated in a written manuscript, but the presentation medium is oral (Koch & Oesterreicher 1994: 587). In contrast, a printed interview or CMC are contained in a written medium, but their conception is oral. As such, CMC is close to spoken language and in fact exhibits features of spoken discourse: for example, in e-forum postings, Biber and Conrad observe a high number of ellipses, particularly of subject pronouns, the expression of emotions and attitudes via punctuation in place of intonation, pitch and speed, and a high frequency of verbs and adverbs (2009: 194-195).

Unfortunately, there are to date no pre-compiled freely available corpora containing a large amount of German language data from, for example, Facebook, Twitter or What’s app. Moreover, the yield rate of a corpus compiled from social network data, i.e. the proportion of imperative singular forms contained in the data base, must be expected to be rather low: the use of an imperative to some extent presupposes the addressee to be able to perform the action denoted in the imperative; thus, the range of executable imperatives in social networks is likely to be very restricted. Although other large web-based corpora like deWaC with 1.7 billion word forms (Baroni et al. 2009; WaCKy 2013) would probably contain a sufficient number of observations, the corpus was not used for the present study because it has a number of disadvantages, particularly concerning metadata of hits. Apart from the source URL, no information is provided about the author of the text (e.g. age or gender) and, more importantly for an investigation of language change, no date of writing is provided. In order to be able to assess the influence of time and social factors, these metadata should be contained in the corpus. Therefore, instead of drawing on existing corpora, a smaller web-based corpus was compiled which promised to contain a high number of observations of the imperative singular of strong verbs with e/i-gradation.

WALKTHROUGHS
Any person who has ever been stuck in a video game and tries to find help on the internet will soon be directed to one of many websites containing so-called “walkthroughs”, “game
guides” or, in German, “Komplettlösungen” ‘complete solutions’. These manuals are usually written by gamers for other gamers and consist in a chronological, concise explanation of which actions a gamer’s avatar (his/ her graphical alter ego in the game) needs to perform in order to progress in the game and finally successfully complete it. The “situational characteristics” of walkthroughs are comparable to e-forum postings, described by Biber and Conrad (2009: 191-193). For at least two reasons, a large proportion of these texts are written employing the imperative singular of verbs: they consist almost exclusively of directions, and the authors and (intended) audience of the walkthrough texts constitute a “virtual community of practice” (Campagna et al. 2012: 10) by their shared interest in video games, creating familiarity between them. These features and the time pressure which authors experience by wanting to write down the instructions for an action in a similar pace as the “real” action takes place in the game can be expected to result in very naturalistic language use, undriven by concerns for correctness or appropriateness. The great diversity of video games in terms of genres, topics, etc., is reflected in the diversity of verbs contained in them; thus, instances of singular imperatives of strong verbs with ei-gradation can be expected to stem from different points on the token frequency continuum.

A further advantage of the walkthrough genre for the present investigation are the characteristics of the authors of the texts. They are predominantly young and thus represent an age group commonly perceived as one picking up language change early. A comparable database containing older speakers’ language was expected to be found in the websites frag-mutti.de and frag-vati.de ‘ask-mom/ask-dad’; however, instructions on these sites were mainly given in the imperative plural, the infinitive or by using modal constructions. The resulting lack of data for older age groups can largely be accounted for by the facts that i) no spoken language corpora are available with a sufficient yield rate and representing the language of different age groups, and ii) middle-aged and especially old speakers do not participate in such forms of communication as walkthroughs (or even Facebook, Twitter or chat with a lower yield rate) to the same extent as young speakers do. This data bias in terms of the factor speaker age is not considered problematic, however, as the phenomenon under investigation is assumed to be a change-in-progress affecting predominantly the language of younger speakers.

---

10 Both platforms have been founded by Bernhard Finkbeiner and Hans-Jörg Brekle. Frag-mutti.de (established 2003) serves as a forum for (predominantly young) users concerning questions of housekeeping (cooking, laundry, shopping etc.); frag-vati.de represents an equivalent for questions concerning home improvement/Do-It-Yourself-tasks but extending to technical and financial problems (Wikipedia 2014).
In order to select an appropriate source website for the research corpus, a random sample of 20 texts yielded by the Google search for the phrase “komplettsung X” ‘walk-through X’ was examined (where X stands for different imperative singular forms of strong verbs with e/i-gradation with the analogical stem vowel e). This procedure proved that walk-throughs really were as rich a data source of imperative singular observations as they appeared. 12 of the 20 texts (60 %) stemmed from a variety of different websites which did not yield sufficient data. Three texts (15 %) came from scharesoft.de (now elderscrollsportal.de); however, this website presents the same problem as the corpus deWaC: neither the date of writing nor metainformation about the authors of texts are provided on this website. Finally, five texts (25 %) came from the website spieletipps.de which offers the necessary metadata for the present study: the posting date of the walkthrough and the author of a text are specified, and authors can provide personal information in their member profile (e.g. first name, last name, place and country of residence, age), which is linked to their posts on the website and can thus be included in the corpus; therefore, the corpus was compiled from this website.

3.1.2.2 Compilation of the walkthrough corpus

A software package consisting of a webcrawler, a speech tagger and a webserver was developed in the programming language Java. The webcrawler is tailored to the specific structure of the website spieletipps.de to ensure that all postings (hence “texts”) and nothing else (e.g. links, commercials etc.) are downloaded. The webcrawler saves all texts from the website in one plain text file (CSV format); each text is stored in a single row of that file, along with the URL of the source webpage and the URL of the author’s member profile (in order to connect texts to metadata, see Table 9):

<table>
<thead>
<tr>
<th>text URL</th>
<th>profile URL</th>
<th>text</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://www.spieletipps.de/pc/elder-scrolls-4-oblivion/tipps/21314/2/#2">http://www.spieletipps.de/pc/elder-scrolls-4-oblivion/tipps/21314/2/#2</a></td>
<td><a href="http://www.spieletipps.de/profil/s">http://www.spieletipps.de/profil/s</a> paske/</td>
<td>... Nehmt dazu am besten, schwache Waffen beim Torwächter da man so öfter einschlagen bzw schießen kann. ...</td>
</tr>
</tbody>
</table>

The output of the webcrawler is sent on to the speech tagger programme, applying the open source TreeTagger (Schmid 1995) to the contents of the plain text CSV file. The result of the speechtagger programme is a new CSV file with an additional column containing the

---

11 Originally, the search yielded 48 texts. Texts with imperative singular forms of only one verb were excluded and texts representing parts of one and the same walkthrough were conflated, resulting in a total of 20 texts.
12 The corpus was compiled by a research assistant from the Computer Science department at the University of Freiburg. I thank Daniel Alcón for his help, especially in acting as an intermediary while giving instructions for the corpus compilation.
POS tags and lemmas of the word forms in the plain text, for example the word form *nehmt* is tagged as as finite form (VVFIN) of the verb *nehmen*:

<table>
<thead>
<tr>
<th>Table 8: Sample output of the speech tagger</th>
</tr>
</thead>
<tbody>
<tr>
<td>text URL</td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td><a href="http://www.spieltipps.de/pc/elder-scrolls-tipp/4/2/2131">http://www.spieltipps.de/pc/elder-scrolls-tipp/4/2/2131</a></td>
</tr>
</tbody>
</table>

In order to extract metadata such as the release date of the respective video game and the metainformation provided on the member’s profile page, the webcrawler was adjusted to create CSV files like the ones containing the corpus texts themselves, one for the release date of the game and one with the author’s metadata (Table 9). Note that at the time of compilation, the posting dates of walkthroughs were not displayed along with the text. Therefore, the game release date and the author’s registration date were originally used for extrapolating the posting date of texts for which the original date was not available from an author’s profile page (for details, see Section 5.2.2).

<table>
<thead>
<tr>
<th>Table 9: Sample output of the webcrawl for metadata (simulated data)</th>
</tr>
</thead>
<tbody>
<tr>
<td>profile URL</td>
</tr>
<tr>
<td>-------------</td>
</tr>
<tr>
<td><a href="http://www.spieltipps.de/profil/nutzer/">http://www.spieltipps.de/profil/nutzer/</a></td>
</tr>
</tbody>
</table>

Finally, the web server programme merges the CSV files created by the speech tagger Table 8) and those containing the metadata about authors (Table 9) and release dates and creates a so-called “inverted index” to enable a search via a web interface as known from web search engines such as Google. When the user enters a simple query like “gib” or a POS query like “VVIMP2geben”, a preview of texts is displayed containing the respective word or POS tag and a link via which a CSV file containing all query hits can be downloaded.

13. The number “2” in the POS tag column in Table 8 functions as a separator between the POS tag and the lemma and has no further meaning. The POS tagset used for this corpus can be found in Schiller et al. (1999).
3.1.3 Projekt Wortschatz Universität Leipzig

The corpus *Projekt Wortschatz Universität Leipzig* is part of the Leipzig Corpora Collection and it contains more than 400 million words from the last 30 years (Quasthoff et al. 2013). Frequencies extracted on the basis of word-level searches in the web-based version of this corpus (Leipzig University 1998-2019) were used as reference counts in the second corpus study (Chapter 5); the corpus was also used as one of the primary corpora in the third corpus study (Chapter 7).

It needs to be noted that the number of displayed query results in Wortschatz Universität Leipzig online is limited to 256 hits. When the actual number of hits exceeds this maximum, a multiple is displayed, e.g. one fifth in the case of 1,250 hits, amounting to 250 displayed results. Distributions in these samples, for example the proportion of uses of the particle verb *gibt* ... *aus* among the search hits for *gibt* and the proportion of uses of the imperative singular and third person singular present indicative among search hits for the ambiguous word form *tritt*, were set in relation to the actual number of hits. In order to test the validity of the sample distributions, the latter were tested for their correlations between verbs and in relation to the frequency values from the other reference corpora. The lemma token frequency values from Projekt Wortschatz Universität Leipzig (i.e. the aggregated intraparadigmatic token counts) are highly correlated with the four scales from the remaining reference corpora, described in Section 3.2: \( r > 0.88 \) (p < 0.001).\(^{14}\) Correlations within the sample sets are slightly lower but significant (\( r > 0.63, \) p < 0.01).

3.1.4 moca

Another corpus used in the third corpus study is *moca* (short for “multimodal oral corpora administration”), “an online system for the administration of spoken language corpora” maintained by the University of Freiburg (2019). It contains searchable transcripts of different corpora along with time-aligned audio/video recordings, three of which were used in the present investigation. The *BigBrother* corpus contains transcripts (233,458 words) and aligned video recordings of the first season of the television show Big Brother, recorded and screened in Germany in 2000. The *CallHome* corpus consists of telephone conversations between native speakers of German in North America and their families and friends at home in Germany. It was collected in 1997 and comprises 205,067 words. The *ReDi* corpus, short for “Regionaldialekte” ‘regional dialects’, was collected by members of the University of Freiburg in

\(^{14}\) Pearson-coefficient (product-moment-correlation); the absolute token frequency scales from Wortschatz Universität Leipzig and the two frequency dictionaries by Ruoff (1990) and Jones & Tschirner (2006) were log-transformed before calculating the correlations in order to ensure comparability with the logarithmic scales provided in the DeReWo list and Duden Online.
2007. It contains interviews (1,591,563 words) with speakers of different ages and from different dialect regions in the Alemannic tri-border area.

3.1.5 FOLK

For decades, the Institute of the German Language (IDS) has collected corpora of predominantly written German language material, many of which have been freely available online for a longer time via the COSMAS portal. In 2012, the Institute added to its collection a further set of searchable corpora, representing spoken language material. The 23 corpora included in the resulting ‘Database for Spoken German’ (*Datenbank für gesprochenes Deutsch*, Institut für deutsche Sprache 2012-2019) cover a variety of communicative situations, dialects, and speaker groups and were compiled at different points in time between 1955 and 2012.

Transcripts of 14 of the 23 corpora in the database are searchable on the database website. Only one of these corpora could be selected for the present investigation because its compilation date roughly matches the time spanned by the Walkthrough Corpus and thus mirrors contemporary language use. **FOLK**, the “Forschungs- und Lehrkorpus Gesprochenes Deutsch” (‘research and didactic/instructional corpus of spoken German’) was collected between 2003 and 2015 and contains approximately 1.6 m word forms. It is designed as a stratified corpus of spoken language and as such comprises “language data from different areas of social life”, such as small talk as well as institutional and public communication (Institut für deutsche Sprache 2018).

3.2 Reference corpora

As mentioned above, the DeReWo list and Wortschatz Universität Leipzig served a double purpose as the primary corpus for the first or third corpus study and as a reference corpus for the Walkthrough Corpus study. For the latter, frequencies were extracted from three additional sources.

3.2.1 Frequency dictionaries

Two of the available frequency dictionaries for German\(^\text{15}\) were consulted as reference dictionaries for the Walkthrough Corpus study. The first of these has been compiled by Rouff (1990)

\(^{15}\) At least one further frequency dictionary is available for German (Quasthoff et. al 2011). However, instead of lemma token frequency counts, it only provides a list of absolute frequencies on the basis of word level searches in the underlying corpus Wortschatz Universität Leipzig which is not manually corrected and thus less useful for the investigation than a separate search in the same corpus.
and is based on exclusively spoken language. It provides absolute token frequency counts of all lemmata contained in the underlying corpus (appr. 2.5 m word forms, Ruoff 1990: 9).

The frequency dictionary by Jones and Tschirner (2006) is based on the Herder/BYU-Corpus which contains 1 million words each from literary, newspaper and academic texts as well as spoken language material, and 200,000 words from instructional texts (appr. 4.2 m word forms in total). It aims at “core vocabulary for learners [of German]” and therefore lists only the 4,037 most frequently used words in the corpus (Jones & Tschirner 2006: 1). It offers token frequency counts (per 1 million words) as well as ranked frequencies; in order to ensure comparability with the other sources, the frequency counts were used in Chapter 5.

3.2.2 Duden Online
The Duden is one of the most widely used dictionaries of German. As is the case for its English equivalent, the Oxford English Dictionary, an online version of the dictionary has been available for some years. This online version provides lemma token frequency information for all entries on the basis of occurrences in the Dudenkorpus. The corpus contains more than 3 billion words from fictional and non-fictional texts, newspaper and journal articles and other sources from the last fifteen years (Bibliographisches Institut GmbH 2011-2019). Token frequency is measured on a logarithmic scale ranging from “5” (top 100 entries in the Duden Korpus) to “1” (below the 100,000 most frequent entries).

3.3 Experimental data
As mentioned before, the competing analogical and traditional forms of the strong verbs with e/i-gradation were presented to native speakers of German in an experimental study. These participants were asked to read the sentences in self-paced reading and repeat each sentence orally. Stimulus sentences were compiled in such a way that they mirror the frequency distributions of the competing imperative variants found in the Walkthrough Corpus study, by including analogical and traditional imperative forms of verbs with a different lemma token frequency. Data were obtained from three participant groups, to test for differences between the reaction patterns of two generations of speakers and those of speakers from two regions in Germany, which can be interpreted as signs of a change-in-progress in the formation of the imperative singular of strong verbs with e/i-gradation. All further details of how data were obtained in the experiment and analysed statistically are provided in Chapter 6.
4 The direction of analogical levelling in strong verbs with e/i-gradation

4.1 Introduction

Analogical levelling in the imperative singular of German strong verbs with e/i-gradation entails the replacement of traditional imperative variants exhibiting stem vowel gradation to \( i \) by variants with the stem vowel \( e \) of the infinitive. The present chapter raises the question whether the direction of this levelling process can be explained on the basis of frequency effects, and whether type or token frequency, or both, are at work in this regard.

4.1.1 Productivity

Processes of analogical levelling have typically been shown to work in the direction of productive patterns. As several scholars (e.g. Bybee 1995; 2010; Barðdal 2008) have pointed out, productivity hinges to a great amount on the type frequency of a pattern, i.e. in the present case the number of different verbs which an inflection pattern already applies to. In cognitive terms, it is assumed that the more often a pattern is instantiated in different lemmata, the more strongly the respective pattern is entrenched in speakers’ mental grammars.

Previous research also agrees on the importance of at least one more determining factor of productivity: The terms/formulations “schematicity” (Bybee 2010: 67), “simplicity of the formation” (‘die Einfachheit der Bildungsweise’, Nübling et al. 2013: 61) or “the number of possible words that can be formed according to a certain pattern” (Haspelmath & Sims 2010: 130) all refer to the fact that low (phonological) constraints in the formation according to a pattern enhance its applicability to new items and thus its productivity. Haspelmath and Sims stress the importance of the interaction between the two factors: “There are simply too many cases of (more or less) unproductive rules that do not seem to be restricted in any general way.” (2010: 130) To say it in Bybee’s terms, it is only “when high schematicity is combined with high type frequency, [that] a maximally productive construction results” (2010: 67). The present chapter thus aims to demonstrate whether the direction of analogical levelling in the imperative singular of strong verbs with e/i-gradation in German is likewise determined by the productivity of verb inflection classes, and if so, whether this productivity is triggered by type frequency or schematicity\(^{16}\), or an interaction of both.

\(^{16}\) For convenience, Bybee’s term “schematicity” is chosen to represent the idea of the degree to which a pattern can be applied with little or no phonological restrictions.
With regard to one further potentially conditioning frequency factor, Bybee has found in her own and cited work that “productivity is governed by type frequency, and that high token frequency can actually detract from productivity” (Bybee 1995: 434, emphasis added; cf. Bybee 1985; Bybee & Hopper 2001 and works cited therein). She argues that linguistic items with a high token frequency acquire such high autonomy that they are no longer recognised as members of a morphological pattern. Due to this autonomy, their high token frequency does not contribute to the productivity of the respective pattern, be it of high or low type frequency. Owing to the fact that the present paper investigates both the potential and limitations of frequency effects in explaining a morphological change-in-progress, the concept of token frequency as a driving force in analogical levelling is taken up as well in this chapter, if only to confirm that it does not influence productivity in the present investigation.

4.1.2 Hypotheses

As concerns the importance of type frequency as a determining factor of productivity of German verb inflection classes, an indication of the prevalence of the weak (or regular) verb paradigm is found in two figures cited in most grammars and research studies: the weak verb paradigm is said to apply to approx. 4000 simplex verbs, as compared to approx. 170 strong simplex verbs (Eisenberg 2013; Nübling et al. 2013; see Chapter 4.2.1). Although these numbers will be replaced by data from a different source in the present study, it can be expected that the class of weak verbs nevertheless has the highest type frequency in German. Along the lines of previous studies (Bybee 1995; 2010), it is assumed that owing to its high type frequency, this inflection pattern is strongly entrenched in the mental lexicon of speakers and may therefore be extended to the inflection of strong verbs in case of doubt. More precisely, it is extended to the formation of the imperative singular of strong verbs with e/i-gradation.

With regard to the role of the schematicity of patterns, it is hypothesised that the productivity of the weak verb paradigm is reinforced by the fact that its formation of the imperative singular can be extended to verbs of any phonological shape. It remains to be seen, however, whether this lack of constraints and the resulting high degree of schematicity are restricted to the weak verb paradigm.

Finally, it has been shown in several studies (Bybee 1985; 1995) that instead of making their respective morphological class or pattern productive, high token frequency members of a particular pattern reduce productivity due to the high degree of autonomy induced by their frequency of usage. Thus, it can be assumed that if the present study should identify patterns with members of high token frequency, these patterns will not be the productive ones.
which serve as a model for the analogical levelling in the imperative singular of the strong verbs with e/i-gradation.

4.2 Data

4.2.1 Previous accounts of frequency distributions

As mentioned above, two figures are provided in many grammars of German with regard to the frequency distribution of patterns of verbal inflection. Bittner (1996), Eisenberg (2013) and Nübling et al. (2013) refer to approx. 4000 verbs in the weak and 169 verbs in the strong inflection class; however, only the latter cites the precise source for these numbers (Nübling et al. 2013: 257): In 1975, Gerhard Augst counted the number of weak, strong and fully irregular verbs among the entries for the initial “letters” K and Sch (1975: 234) in the rank frequency list provided by Helmut Meier in his Deutsche Sprachstatistik (1967). On the basis of these counts, he extrapolated that, overall, approx. 3811 simplex verbs in German are weak verbs and 169 simplex verbs are strong verbs (Augst 1975: 235). Even though these extrapolations could give an indication of the productivity which might be used for the present investigation, the (extrapolated) type frequency of the strong verbs with e/i-gradation is not listed separately so that it would not be possible to compare this inflection class directly with the weak verbs.

More importantly, however, although Meier’s Sprachstatistik was published in 1967, it relies on frequency counts obtained by Friedrich Wilhelm Kaeding between 1891 and 1897 (Meier 1967, vol.1: 7). By citing Meier’s or Augst’s frequency figures for “contemporary” (Augst 1975: 235) or “today’s” German (Nübling et al. 2013: 257), the authors seem to neglect Meier’s note of caution as to the source and nature of the frequency data.

In these lines from the preface to the “rank book of German word forms and words”, the author stresses two important points for the present investigation: (i) the frequency data are taken from (edited) German prose written around the turn of the century (1900); (ii) subsequent cultural and political changes have repercussions in the lexicon which should be reflected in new frequency counts. That the frequency counts from Kaeding’s work were used in contemporary grammars of German can be understood in light of the facts that (i) more recent precompiled lemma frequency lists have not been available until a few years ago (2007, see following section) and (ii) these somewhat outdated figures do not cast doubt on the content of the authors’ work. However, bearing in mind that at least one of the verbs under investigation in the current study, *melken* ‘milk’\(^\text{17}\), belongs to a group which according to Nübling et al. (2013: 253) is decreasing in frequency on the basis of a ‘change in circumstances’ (“Wandel der Lebensumstände”), it seems more than reasonable to rely on a more recent available source for frequency data in the present analysis.

4.2.2 The DeReWo list of lemma token frequency

Figures from the DeReWo lemma token frequency list offered by the Institute for the German Language from 2012 (Institut für deutsche Sprache 2012b; see Chapter 3.1.1) are assumed to reflect the linguistic reality of the present topic of investigation better than the data from Meier’s language statistics (1967) because they mirror contemporary German at the time that the change-in-progress in the imperative singular of strong verbs with e/i-gradation is examined.

As mentioned before, there is a high number of particles verbs in German, such as *aufgeben* ‘give up’, which may be problematic for the following analysis. An earlier study of the productivity of verb inflection classes in German (Clahsen & Rothweiler 1992) was criticised because the authors had chosen to count particle verbs as distinct types, i.e. counting *geben*, *aufgeben*, and *ausgeben* as different lemmas, which according to Bybee “artificially inflates the type frequency of the irregular verbs” (1995: 435). In order to avoid a similar distortion of the results of this study, type and token frequencies will be presented for a subset of non-particle verbs and for the whole dataset extracted from the DeReWo list\(^\text{18}\).

\(^{17}\) Other strong verbs with e/i-gradation such as *dreschen* ‘thresh/thresh’ and *fechten* ‘fence’ are not listed; they can be assumed to be similarly affected by cultural changes.

\(^{18}\) As concerns token frequency, a different possible method would be to add up the frequency of the simplex verb and those of the particle verbs derived from it. However, as the DeReWo list provides frequency classes instead of summable absolute frequency counts, this is not possible in the present case.
4.3 Method

4.3.1 Verb extraction
While the corpus DeReKo itself is part-of-speech tagged and morphosyntactically annotated, the DeReWo list only contains part-of-speech information in ambiguous cases, distinguishing, for example, the pronoun sein (‘his/its’) from the verb sein (‘to be’). Thus, it was not possible to simply extract all verb lemmata from the list and analyse the result. In order to arrive at a list of all verbs, therefore, structural properties of German verb morphology were exploited. As a partial automation of the search procedure, with the help of R-Studio all entries were extracted which fulfilled two criteria: i) they should start in a lower case letter, to remove all (proper) nouns, and ii) they should end in the letter “n” because all verb lemmata in the list are recorded in the infinitive form, carrying this inflectional suffix. Due to the fact that these two criteria also apply to some adjectives, pronouns, prepositions and participles, a number of false hits remained after the automatic search which were sorted out manually during the annotation process.

Apart from the false hits, all entries in the resulting list were annotated for the verb inflection paradigm which they belong to. As explained in the following section, five paradigms were distinguished: weak verbs, strong verbs, strong verbs with e/i-gradation, mixed verbs, and modal verbs. In cases of doubt, the online version of the Duden dictionary (Duden Online) was consulted.

4.3.2 Verb class annotation
There is no general consensus even in the prevalent and most recent grammars of German with regard to the number of basic verb inflection classes. The only common distinction is between weak (or regular) and strong (or irregular) verbs. In addition, some grammars describe a class of modal verbs (Eisenberg 2013; Sahel & Vogel 2013; Sommerfeldt & Starke 1998), others add a class of “mixed verbs” (Dürscheid et al. 1994; Konopka 2007; Römer 2006; Wellmann 2008). For the purpose of verb class annotation in the DeReWo list, the present study follows the classification in the most recent edition of the Duden grammar (Eisenberg & Kunkel-Razum 2009) which lists all four of these inflection classes. As the investigation focuses on the paradigm of strong verbs with e/i-gradation, this group is taken up as an additional inflection class. The chief differences between the resulting five classes (a-e) shall be briefly explained; in line with the common practice in the grammars mentioned above, four distinguishing forms from the paradigm of one example verb each are listed as an illustration: 3rd sg.pres. = third person singular present indicative, 3rd pl.pres. = third person
plural present indicative, 3rd sg.pret. = third person singular preterite indicative, imp.sg. = imperative singular.

a. **Weak verbs** are regularly inflected verbs, i.e. all forms are derived from the same verb stem and inflectional morphemes distinguish number, person, tense, and mood.

   infinitive translation 3rd sg.pres. 3rd pl.pres. 3rd sg.pret. imp.sg.

   sagen 'to say' sag-t sag-en sag-te sag(-e)!

b. In contrast, **strong verbs** show a stem vowel alternation according to tense. Within the tenses, number, person and mood contrasts are expressed by inflectional morphemes.

   infinitive translation 3rd sg.pres. 3rd pl.pres. 3rd sg.pret. imp.sg.

   fliehen 'to flee' flieh-t flieh-en flieh flieh(-e)!

c. **Strong verbs with e/i-gradation**, the object of the present study, exhibit an additional stem vowel alternation according to number and mood within the present tense.

   infinitive translation 3rd sg.pres. 3rd pl.pres. 3rd sg.pret. imp.sg.

   helfen 'to help' hilf-t helf-en half helfen!

d. So-called ‘**mixed verbs**’ (“gemischte Verben”) owe their name to the fact that they exhibit both a stem vowel alternation (like the strong verbs) and suffixation (like the weak verbs) in the past tense.

   infinitive translation 3rd sg.pres. 3rd pl.pres. 3rd sg.pret. imp.sg.

   rennen 'to run': renn-t renn-en renn-te renn(-e)!

e. In German grammars, the class of **modal verbs** also goes by the name of “Präterito-präsentia” ‘preterite presents’, reminiscent of the fact that they are originally past forms of strong verbs which were used in the present tense to express obligation, volition etc.

   infinitive translation 3rd sg.pres. 3rd pl.pres. 3rd sg.pret. imp.sg.

   wollen 'to want' will woll-en woll-te ---

   Today, they do not form an imperative at all because they describe processes or events which are not under the control of the subject agent (Eisenberg & Kunkel-Razum 2009: 549). As their imperative can thus not function as a model for analogy, they will be disregarded in the remaining sections.

### 4.4 Results

#### 4.4.1 Schematicity

As can be seen from the inflection of the example verbs in Section 4.3.2, the imperative singular of several of the German verb classes is formed according to the same pattern. For weak verbs, strong verbs and mixed verbs, the pattern is [[PRESENT TENSE STEM][(-e)] or, since
the present tense stem for the respective verbs is identical with the infinitive stem, \([\text{INFINITIVE STEM}](-e)\). This pattern applies easily to all verbs with a single present tense stem, i.e. the vast majority of verbs; this minor phonological constraint thus makes the pattern highly schematic.

The imperative singular of strong verbs with e/i-gradation is similarly formed from the present tense stem; nevertheless, the formation is not as straightforward as for the groups mentioned above. Strong verbs with e/i-gradation have two present tense stem vowels, \(e\) for the first person singular and all persons of the plural and \(i\) for the second and third person singular. The latter stem vowel is also used for the formation of the imperative singular, which yields the pattern \([\text{PRESENT TENSE STEM WITH VOWEL GRADATION}]\)-ø]. This pattern can only apply to verbs with two present tense stems (one identical with the infinitive stem, one with a different stem vowel); due to this constraint, it has a low degree of schematicity.

4.4.2 Type Frequency

All in all, the DeReWo list contains 7,568 verb lemmata (16,999 including particle verbs). Their distribution across the four inflection classes is summarised in Table 10:

<table>
<thead>
<tr>
<th>verb class</th>
<th>type frequency in DeReWo list excluding particle verbs</th>
<th>type frequency in DeReWo list including particle verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td>weak verbs</td>
<td>6,954 91.89</td>
<td>13,342 78.49</td>
</tr>
<tr>
<td>strong verbs</td>
<td>475 6.28</td>
<td>2,876 16.92</td>
</tr>
<tr>
<td>strong verbs with e/i-gradation</td>
<td>114 1.50</td>
<td>607 3.57</td>
</tr>
<tr>
<td>mixed verbs</td>
<td>25 0.33</td>
<td>174 1.02</td>
</tr>
<tr>
<td>total</td>
<td>7,568 100.00</td>
<td>16,999 100.00</td>
</tr>
</tbody>
</table>

A Chi-squared test shows that the observed distribution of verb lemmata across inflection classes is significantly different from the expected distribution (\(p < 0.001\)). Regardless of whether particle verbs are included in the distribution or not, the weak verb class clearly contains the majority of verbs listed in the DeReWo, while all other classes taken together cover only 8.11% or 21.51%, respectively. The second class of verbs which is found to have a higher type frequency than strong verbs with e/i-gradation is “other” strong verbs, i.e. strong verbs without vowel gradation within the present tense.

Productivity in word-formation, as displayed by the probability of members from a verb class to be combined with a particle to form a separable verb, is very different from productivity in inflection and will not be discussed in the present study. The observation that strong and mixed verbs form more particle verbs than the members of the weak verb class can be understood in light of the fact that the former constitute the ‘basic vocabulary’ of German
(“Grundwortschatz”, Nübling et al. 2013: 257) which speakers modify more often with the use of particles than the concepts expressed by weak verbs.

### 4.4.3 Token Frequency

The question remains whether the inflection classes also apply to verb lemmata of different token frequency. To this end, entries in the DeReWo list were cross-tabulated according to the variables verb class and frequency class. A Chi-squared test shows that the observed distributions are significantly different from the expected distributions ($\chi^2(3) = 18117.71$, $p < 0.001$ when excluding particle verbs; $\chi^2(3) = 26928.07$, $p < 0.001$ when including particle verbs).

Thus, as the association plots in Figures 4 and 5 below illustrate, some combinations are over-, others underrepresented. Grey bars (below the dotted lines) indicate that the observed values are lower than expected, black bars (above the dotted lines) indicate that they were higher than expected. The y-axis represents the original frequency classes in the DeReWo list, with the highest frequency classes on the upper and the lowest frequency classes on the lower end.

![Diagram](image)

**Figure 4:** Type and token frequency distribution of verb classes in the DeReWo list (particle verbs excluded)

Note: SVe/iG = strong verbs with e/i-gradation

The weak verb inflection class thus applies to fewer high token-frequency verbs and more lower token-frequency verbs than would be expected, while the opposite is true for strong verbs, strong verbs with e/i-gradation and mixed verbs.
4.5 Discussion

The analysis of the DeReWo list has revealed three constellations in the German verb inflection classes with regard to the factors type frequency, schematicity and the token frequency of members:

<table>
<thead>
<tr>
<th>constellation number</th>
<th>instantiated by</th>
<th>type frequency</th>
<th>degree of schematicity</th>
<th>token frequency of members</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>weak verbs</td>
<td>high</td>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td>2</td>
<td>strong verbs, mixed verbs</td>
<td>low</td>
<td>high</td>
<td>high</td>
</tr>
<tr>
<td>3</td>
<td>strong verbs with e/i-gradation</td>
<td>low</td>
<td>low</td>
<td>high</td>
</tr>
</tbody>
</table>

There are striking similarities and differences between the verb classes which can potentially work as a model for the analogical levelling in the imperative singular of strong verbs with e/i-gradation. Owing to the fact that they make use of the same inflection pattern for the formation of the imperative singular, both the weak verb class as well as the strong and mixed verb classes exhibit the same high degree of schematicity. On the basis of this feature, they could all qualify as productive patterns and thus models for analogy. However, while the weak verb class has a high type frequency and low token frequency members, the opposite is true for the strong and mixed verbs with a low type frequency and high token frequency members. Therefore, on the basis of the constellations in Table 11, it is not possible to say whether type or token frequency is the major determinant of productivity and whether it is the
weak or the strong/mixed verb class which dictates the direction for the analogical change in the strong verbs with e/i-gradation.

In order to find out whether type and/or token frequency are the determinants of productivity in the present case, a different division of verb classes may be useful. Since all the weak, strong and mixed verbs follow the same pattern of [[INFINITIVE STEM](-e)] for the formation of the imperative singular which has been shown to be highly schematic, it might be reasonable to assume one overarching verb class. The relative type frequency of this “new” verb class is very high; it amounts to almost 100% of all verbs. The remaining verb class, that of strong verbs with e/i-gradation, with the lowly schematic imperative singular pattern [[PRESENT TENSE STEM WITH VOWEL GRADATION]-ø] has a relative type frequency of only 1.5% or 3.57%, depending on whether particle verbs are in- or excluded. (cf. Table 10).

![Figure 6: Type and token frequency distribution of classes “weak, strong and mixed verbs” and “strong verbs with e/i-gradation” in the DeReWo list (particle verbs excluded)](image)

Figure 6: Type and token frequency distribution of classes “weak, strong and mixed verbs” and “strong verbs with e/i-gradation” in the DeReWo list (particle verbs excluded)

The association plots of the distribution of these two verb classes in Figure 6 clearly shows the dominance of the class comprising weak, strong and mixed verbs in terms of type frequency; more importantly, however, it also reveals that high token frequency members are underrepresented in this class. Thus, the productivity of the imperative pattern [[INFINITIVE STEM](-e)] is once again shown to be determined primarily by its type frequency and degree of schematicity. This conclusion is supported by the fact that the composite verb class consists mainly of the weak verb class with a high type frequency and only few strong and mixed verbs with a high token frequency, which does not render either of these two verb classes more productive.
4.6 Summary

An examination of the distribution of German verb inflection classes proves that besides schematicity, frequency is one of the main explaining factors for the direction of the ongoing analogical levelling in the imperative singular of strong verbs with e/i-gradation. The regular imperative singular formation pattern [[INFinitive STEM](-e)] is highly schematic: it applies to all verbs whose infinitive stem is used for all present tense forms, including the imperative singular. Analogical imperative singular variants of strong verbs with e/i-gradation with the infinitive stem vowel e emerge as a result of the extension of this regular pattern. This analogical levelling is reinforced by the type frequency of the regular pattern; it applies to 96% of the German verbs. This pattern can thus be assumed to be highly entrenched in the mental lexicon of speakers and all the more readily applied to strong verbs with e/i-gradation when their traditional irregular imperative singular form with the stem vowel i is not directly accessed. Finally, as has been found in previous studies (Bybee 1985; 1995), the token frequency of verbs does not have an impact on the productivity of a verb class: in the most productive class of German weak verbs high token frequency verbs are underrepresented, while the least productive class of strong verbs with e/i-gradation has a high number of high token frequency verbs.

In contrast to the relation between the conserving effect of high token frequency and the mental entrenchment of competing traditional and analogical forms which has never been tested in an experimental study (cf. Section 1.1.3), evidence for the entrenchment effect of the type frequency of patterns has been provided from nonce-probe experiments conducted by Moder (1992) and Wang and Derwing (1994) for English. This evidence can be expected to apply to the closely related language German as well; therefore, a separate test for German, which would go beyond the scope of the present investigation, was deemed unnecessary.
5 The trajectory of analogical levelling in the imperative singular\textsuperscript{19}

The previous chapter has shown that frequency effects can explain the direction of analogical levelling in the imperative singular of German strong verbs with e/i-gradation. The present chapter raises the question whether this levelling process is properly described as a change-in-progress or simply represents a case of linguistic variation. It also investigates whether it affects all strong verbs with e/i-gradation to the same extent, and if, on the contrary, some verbs occur in an analogical formation more often than others, in how far this trajectory can be explained on the basis of frequency and persistence effects and the influence of other factors. Finally, it explores the impact of sociolinguistic variables, such as speakers’ age or dialectal background on imperative singular formation, to the extent that this is possible.

The first section introduces the hypotheses derived from usage-based and sociolinguistic research on language variation and change which will be tested in this corpus study (5.1). Section 5.2 gives a summary of how observations of the imperative singular of strong verbs with e/i-gradation were extracted from the Walkthrough Corpus and how they were annotated for the variables whose influence on imperative singular formation was tested in the present study, including why and which random variables were entered in the mixed-effects regression analysis of the corpus dataset. Section 5.3 presents the results of the analysis of the corpus data, which confirm the classification of the analogical levelling process in the imperative singular of strong verbs with e/i-gradation as a “change-in-progress”, the existence of a conserving effect of high token frequency in analogical change-in-progress and the importance of persistence effects for such levelling phenomena. The relevance of the results for the present project and the usage-based approach to language are discussed in Section 5.4. The final Section 5.5 provides a summary of the findings of the study and an outlook into how questions that could not be exhaustively answered in the current corpus study were tackled in the experimental study (Chapter 6).

5.1 Hypotheses

5.1.1 Variation or change-in-progress

Although imperative singular formation in strong verbs with e/i-gradation has varied at least since the 16\textsuperscript{th} century (see 1.2.1), it can be argued that it has by now moved beyond the stage

\textsuperscript{19} Parts of the present chapter have been published in Corpus Linguistics and Linguistic Theory (Krause-Lerche 2019a).
of pure variation to an ongoing change by which the traditional imperative singular variants with the stem vowel \(i\) are gradually replaced by analogical variants with the stem vowel \(e\). The present study should thus reveal an increase in the use of analogical variants of the imperative singular with the stem vowel \(e\) over the time span covered by the corpus data.

### 5.1.2 Verb token frequency

This ongoing change does not seem to progress through the verb class uniformly, i.e. it does not affect all strong verbs with \(e/i\)-gradation to the same extent. On the basis of previous research, the imperative singular forms of low frequency verbs, such as *bestehlen* and *zertreten*, are expected to occur in an analogical variant with the stem vowel \(e\) more often than those of high frequency verbs, e.g. *nehmen* and *sehen*. In this regard, the present study will answer the question whether in a more highly inflected language like German this conserving effect is exerted by verb lemma token frequency or base verb lemma token frequency (cf. 1.3 and 5.2.2).

### 5.1.3 Intraparadigmatic type and token frequencies

In addition to lemma token frequency, intraparadigmatic type and token frequencies can be assumed to affect imperative singular formation in strong verbs with \(e/i\)-gradation. The past participle of some of the verbs is formed with the stem vowel \(e\) (e.g. *gegeben*); these verbs have a higher intraparadigmatic type frequency of the stem vowel \(e\) than verbs with a past participle with the stem vowel \(o\), e.g. *gestorben*. On the basis of the higher type frequency of the stem vowel \(e\) in the paradigm of *geben*, this verb can be expected to show a stronger tendency to be used with the analogical imperative singular variant *geb(e)*, whereas the conservation of the traditional irregular imperative singular variant of the verb *sterben* (stirb) may be reinforced by the lower intraparadigmatic type frequency of the stem vowel \(e\).

The token frequency of forms is likewise very different across paradigms of strong verbs with \(e/i\)-gradation. As mentioned in Section 1.3, paradigm forms of *geben* with the stem vowel \(i\) have a higher token frequency than forms with the stem vowel \(e\); the opposite is true for the verb *sterben*. Conservation of the traditional imperative formation of strong verbs with \(e/i\)-gradation with the stem vowel \(i\) can therefore be expected to concentrate in verbs with a high intraparadigmatic token frequency of forms with this stem vowel. A similar conserving effect of high relative intraparadigmatic token frequency is expected for the imperative singular as such: The higher the token frequency of the imperative singular in the paradigm of a strong verb with \(e/i\)-gradation, the lower will be this verb’s propensity to occur with an analogical imperative singular variant.
5.1.4 Persistence

As outlined in Chapter 1.4, although analogical levelling in the imperative singular of strong verbs with e/i-gradation is assumed to represent a case of ongoing change, the choice of a stem vowel and suffixation in the imperative may still be affected by variables which have been shown to influence linguistic variation. In this regard, the verb class and suffixation of previously used imperative singular forms may have an influence on the suffixation of target instances of the imperative singular of strong verbs with e/i-gradation, and thereby indirectly stem vowel choice in these forms. α-persistence (Szmrecsanyi 2005; 2006) is expected to explain cases in which an imperative singular variant of one strong verb with e/i-gradation is repeated in the same verb (path α1 in Figure 7) or triggers the use of the same variant of another strong verb with e/i-gradation (path α2), for example, gebe followed by another instance of gebe or by the form nehme (1.) or gib followed by gib or nimm (3.).

Figure 7: α-persistence in imperative singular formation of strong verbs with e/i-gradation

<table>
<thead>
<tr>
<th>previous imperative singular of strong verb with e/i-gradation</th>
<th>α1</th>
<th>target imperative singular of same strong verb with e/i-gradation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. [e]-e&lt;sub&gt;a&lt;/sub&gt;</td>
<td>1. [e]-e&lt;sub&gt;a&lt;/sub&gt;</td>
<td></td>
</tr>
<tr>
<td>2. [e]-φ&lt;sub&gt;a&lt;/sub&gt;</td>
<td>2. [e]-φ&lt;sub&gt;a&lt;/sub&gt;</td>
<td></td>
</tr>
<tr>
<td>3. [i]-φ&lt;sub&gt;a&lt;/sub&gt;</td>
<td>3. [i]-φ&lt;sub&gt;a&lt;/sub&gt;</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>previous imperative singular of strong verb with e/i-gradation</th>
<th>α2</th>
<th>target imperative singular of other strong verb with e/i-gradation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. [e]-e&lt;sub&gt;b&lt;/sub&gt;</td>
<td>1. [e]-e&lt;sub&gt;b&lt;/sub&gt;</td>
<td></td>
</tr>
<tr>
<td>2. [e]-φ&lt;sub&gt;b&lt;/sub&gt;</td>
<td>2. [e]-φ&lt;sub&gt;b&lt;/sub&gt;</td>
<td></td>
</tr>
<tr>
<td>3. [i]-φ&lt;sub&gt;b&lt;/sub&gt;</td>
<td>3. [i]-φ&lt;sub&gt;b&lt;/sub&gt;</td>
<td></td>
</tr>
</tbody>
</table>

β-persistence (Szmrecsanyi 2005; 2006) concerns the previous use of imperative singular forms of verbs other than the strong verbs with e/i-gradation. When unsuffixed imperative singular forms of such verbs like lauf-ø ‘run’ or verlass-ø ‘leave’ precede the target imperative singular form of a strong verb with e/i-gradation, the latter is expected to be unsuffixed as well (e.g. gib-ø or geb-ø; path β1 in Figure 8). When preceding imperative singular forms carry a suffix (e.g. lauf<sub>e</sub>, verlass<sub>e</sub>), the target imperative singular form will most likely be suffixed as well. A suffixed version of the traditional imperative singular variant of strong verbs with e/i-gradation with the stem vowel <i>i</i> does not exist (e.g. *gibe); thus, when suffixation of preceding imperative singular forms triggers the target imperative singular form to be suffixed, the stem vowel must be <i>e</i> (e.g. gebe; path β2)
In his studies of persistence patterns in five cases of variation in English, Szmrecsanyi uses the term CURRENT to refer to the target instance whose form is influenced by a PREVIOUS (2005; 2006, capitalisation original). In the present study, the target imperative singular instance of a strong verb with e/i-gradation will be referred to as the target imperative singular forms preceding this target, whose form is assumed to have an influence on the target imperative, are referred to as pretargets (cf. Czoska 2016).

5.1.5 Sociolinguistic variables

The survey of dialectological literature in Chapter 1.4.2 has shown that in some regional varieties of German, speakers use the traditional variant of the imperative singular of strong verbs with e/i-gradation with the stem vowel i (the Standard German variant), while an analogical formation with the stem vowel e is used by speakers of other regional varieties. It can thus be assumed that the choice of a stem vowel for the imperative singular forms recorded in the corpus data is - at least partly - influenced by speakers’ dialectal background. More precisely, speakers from Northern Germany, the Middle West and Thuringia can be expected to use imperative singular variants with the stem vowel e more often than speakers from Southern Germany (including East Franconia) and the Upper Saxon region (see Figure 1).

Moreover, sociolinguistic research has found that male and female speakers play different roles in linguistic variation and change. In stable linguistic variation or when change proceeds “above the level of consciousness”, women react to the overt prestige of the standard variant of a variable and avoid using incoming variants. However, when a variable changes “below the level of consciousness”, women are the leaders of change (Labov 1990; 2001). Due to this Gender Paradox and the fact that it has been described predominantly for phonological/phonetic variation and change (Labov 1990; 2001), female or male speakers may use a higher rate of analogical imperative singular forms of strong verbs with e/i-gradation with
the stem vowel \(e\), depending on the developmental stage of the analogical levelling process and speakers’ awareness of the variable (see 1.4.3).

A last demographic factor which may have an influence on the question whether speakers use the traditional stem vowel \(i\) or the analogical variant with the stem vowel \(e\) is their age. Variation between the two stem vowels has been commented for a longer time (see 1.2.1), but it does not seem to be a feature of linguistic variation or change in German which speakers are generally aware of. As will be shown below, the corpus used in the present study covers a relatively short time span and might thus not be able to uncover change-in-progress in imperative singular formation, even if it occurs. However, an examination of patterns according to speaker age may capture apparent-time change: this concept was introduced in sociolinguistics to describe a situation in which younger speakers use an innovative variant more often than older speakers from the same speech community. It is assumed that after a certain point during adolescence, speakers’ linguistic patterns remain relatively stable, so that a comparison of the language of speakers in their sixties and speakers in their twenties is close to sampling speech at one point in time and 40 years later (Meyerhoff 2011: 141). If differences in imperative formation were found to be patterned according to speakers’ age, it can be assumed that younger speakers use more of the innovative analogical e-stem variant of the imperative singular of strong verbs with e/i-gradation than older speakers in the corpus.

5.1.6 Traditional linguistic factors

Last but not least, the influence of “traditional” intralinguistic factors on imperative singular formation of strong verbs with e/i-gradation will be tested as well. These verbs fall into three of the seven ablaut classes of strong verbs in German (3b, 4, or 5); one or several of these classes may show a higher or lower general propensity to be affected by analogical levelling. The morphological structure of the verb might have an influence in so far as simplex and prefixed verbs may behave differently from particle verbs, or alternative groupings show up. Similarly, phonotactic constraints may explain the stronger or weaker tendency of some strong verbs with e/i-gradation to be used with a traditional or analogical imperative variant, dependent on i) the number of syllables (e.g. \(geben\) vs. \(übernehmen\)), ii) the preceding and/or following phonetic context (e.g. \(gib ich mir\) vs. \(geb es weg\)), iii) the occurrence of the imperative form sentence-initially, clause-initially or in a non-initial position in the sentence (for example, preceded by an adverb, like \(nun gib ...\) ‘now give ...’), or iv) the question whether the length of the stem vowel differs between the traditional and analogical imperative variants (e.g. \(nimm\) [nim] / \(nehmen\) [ne:m] vs. \(brich\) [brɪtʃ] / \(brech\) [brɛtʃ]). Finally, previous studies of morphosyntactic variation have shown that the factor negation can have a significant impact
on the choice of one variant of a linguistic variable over another (e.g. Poplack 2001). However, no specific or directional hypotheses with respect to the use of imperative singular variants of strong verbs with e/i-gradation can be provided for these factors.

5.2 Data and method

5.2.1 Corpus data

The selection and compilation of the primary corpus for the present study, the Walkthrough Corpus, was explained in Chapter 3, along with a description of the reference corpora which were consulted for the annotation of a number of frequency variables. The present section can only give a brief summary of how data were extracted from the Walkthrough Corpus; details are provided in Krause (2016) and in Appendix C.

The Walkthrough Corpus was searched for occurrences of all three imperative singular variants of strong verbs with e/i-gradation, e.g. *gib*, *geb* and *gebe*, either on POS- or on word level. The resulting dataset contained 2,446 observations; it was, however, reduced through two steps of manual post-correction. The same text (or parts thereof) may appear several times on the website for a number of reasons. As deduplication, i.e. the deletion of repeated material, was not implemented during corpus compilation, 198 duplicates and near-duplicates were sorted out manually. Finally, 309 of the remaining 2,248 imperatives were removed because they are part of original in-game instructions. Thus, the final dataset contained 1,939 observations of imperative singular forms of strong verbs with e/i-gradation.

All instances in the final dataset needed to be annotated for a number of frequency variables and other factors whose influence on imperative singular formation was tested in the present study. However, values for the frequency variables can be expected to be rather skewed in the Walkthrough Corpus due to the genre of the texts included: For example, avatars in games do not usually have the same eating habits as real human beings - *essen* 'to eat' being one of the verbs under investigation - which leads to a possible under- or overrepresentation of some verb lemmata. The numbers should therefore not be taken from the Walkthrough Corpus; the following five reference sources were consulted instead:

- two frequency dictionaries of German by Ruoff (1990) and Jones & Tschirner (2006)
- the online Duden dictionary (Bibliographisches Institut GmbH 2011-2019)
- the DeReWo lemma frequency list (Institut für deutsche Sprache 2012b) and
5.2.2 Annotation of variables and distribution

5.2.2.1 Dependent variables

The dependent variable, i.e. the form of the imperative singular, was annotated first. Instances received an annotation for STEM VOWEL (“e” or “i”) and SUFFIXATION (“suffixed” or “unsuffixed”), thus being able to differentiate all three possible forms *gib* (STEM VOWEL i, SUFFIXATION no), *geb* (STEM VOWEL e, SUFFIXATION no), and *gebe* (STEM VOWEL e, SUFFIXATION yes), and to attribute the influence of the predictor variables to either of the two dependent variables. A cross-tabulation of stem vowel and suffixation shows the following data distribution:

<table>
<thead>
<tr>
<th>SUFFIXATION</th>
<th>STEM VOWEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>suffixed</td>
<td>1299</td>
</tr>
<tr>
<td></td>
<td>43</td>
</tr>
<tr>
<td>unsuffixed</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>597</td>
</tr>
</tbody>
</table>

5.2.2.2 Predictor variables

As mentioned above, in a study of language change, the factor TIME is crucial. When the Walkthrough Corpus was compiled in 2013, unfortunately, the date of a walkthrough post was not displayed with the text on the website *spieletipps.de*; therefore, it was originally collected from the members’ profile pages, where the dates of the last four postings by that specific member are listed. Due to active membership of the author in the forum, the data points might not, however, be recorded among these last four postings. In these cases, two other pieces of information were gathered from the website: the release date of the game for which the walkthrough is written, and the date when the author had registered as a member of the website. Through a simple syllogism, the more recent date was recorded as the probable “post date”: an author cannot have posted the walkthrough for a game either before he has registered on the website or before the game has been published. The recorded and extrapolated dates were reduced to the respective year, thus yielding values for the variable POST.YEAR. Statistical models were run twice, either including all 1,939 observations or including only observations for which the original timestamp was available from the profile page (N=1,466, see Figure 9). No statistically significant differences were found between the regression models; thus, the uncertainty arising from extrapolation turned out non-problematic for the output of the analysis. Nevertheless, when the posting dates of texts on the website *spieletipps.de* were displayed from the beginning of the year 2017, all annotations for POST.YEAR in the corpus dataset were corrected and all statistical models were run again. This revision confirmed the results of
the previous analysis, yielding further evidence that reliance on partially extrapolated posting dates did not cause a major problem in terms of the results of the corpus analysis.

![Figure 9: Distribution of instances in the dataset over the variable POST.YEAR](image)

Frequency effects on language change are the over-arching topic of the present investigation. The first frequency variable whose influence should be tested is VERB LEMMA TOKEN FREQUENCY. As mentioned before, lemma frequencies were taken from five reference corpora. The frequency information from all sources should be incorporated into the regression analysis, but since the five scales are highly correlated, they could not have been entered as single predictors. Instead, the three absolute frequency scales from the frequency dictionaries and the Wortschatz corpus were log-transformed in order to ensure comparability with the logarithmic scales from Duden Online and DeReWo. These different scales were then transformed into a single scale by means of a principal components analysis. According to Harald Baayen (2008: 118-126), each of the five scales spans a plane in multi-dimensional space on which the values for the different verb lemmata are located. Principal components analysis rotates these planes, so that the variability of all five is transferred to a two-dimensional space. The scattering of values around the original planes of the five scales is captured by principal components (vectors of the values in the two-dimensional space), “ordered by how much variability they account for” (Baayen 2008: 120). He describes a rule-of-thumb that “only those principal components are important that account for at least 5 % of the variance” (Baayen 2008: 121). This rule applied to the first two principal components in the present analysis, with the first principal component (PC1.VERB) explaining 83.6 % and the second principal component (PC2.VERB) explaining 8.5 % of the variance. Values of PC1.VERB are highly correlated with four of the entered frequency scales (r > 0.9, p < 0.001;
cf. example verbs in left-hand panel of Figure 10); this component would thus have been sufficient to rotate these scales to a common two-dimensional space. The only frequency scale which is not as strongly correlated with PC1.VERB is the one from Ruoff’s (1990) frequency dictionary ($r = -0.74$, $p < 0.001$); at the same time, it is the only scale with a strong and significant correlation with the second principal component PC2.VERB ($r = 0.48$, $p < 0.001$).

The corpus underlying this frequency dictionary contains everyday language collected from a rural population (Ruoff 1990: 14), which leads to a slight over- or underrepresentation of some strong verbs with e/i-gradation like *melken ‘milk’. That means, the second principal component mainly serves the purpose to rotate the plane spanned by Ruoff’s frequency scale further to bring it into a common two-dimensional space with the remaining scales (see example verbs in right-hand panel of Figure 10). Although PC2.VERB was therefore not expected to have a significant effect on imperative formation, the values of both principal components were assigned to all 94 verbs in the dataset as measures for verb lemma token frequency, hereby distinguishing the lemma token frequency of, for example, the simplex verb *geben ‘give’, the prefixed verb *begeben ‘proceed toward/set off’ and the particle verb *achtgeben ‘look out’.

![Figure 10: Distribution of variables PC1.VERB and PC2.VERB in the dataset](image)

Note: Bars may represent observations for several verbs (including the respective example verb).

It was argued in Sections 1.3 and 5.1 that in German, *BASE VERB LEMMA TOKEN FREQUENCY may be the better predictor of the propensity of a strong verb with e/i-gradation to be affected by analogical levelling. In order to test this hypothesis, a second principal components analysis was performed for part of the original five scales, i.e. only for the 18 simplex and 21 prefixed verbs. Similarly as in the first analysis, the two first principal
components explained at least 5% variance: 85.2% by PC1.BASE, and 8.5% by PC2.BASE. Ruoff’s (1990) frequency scale is again less strongly correlated with the values of the first principal component (PC1.BASE) than the other four frequency scales ($r = 0.78$, $p < 0.001$) but more strongly with the second principal component (PC2.BASE) than the remaining scales ($r = 0.34$, $p = 0.065$). Like PC2.VERB, therefore, PC2.BASE is not expected to have a significant effect on imperative formation. Nevertheless, the values of both principal components were assigned as the measure for base verb lemma token frequency (PC1.BASE and PC2.BASE); particle verbs in the dataset were assigned the base verb lemma token frequency values of their respective simplex bases. For example, observations of the particle verb *achtgeben* received the values of its base verb *geben* (Figure 11; cf. Table 4).

![Figure 11: Distribution of variables PC1.BASE and PC2.BASE in the dataset](image)

*Note: Bars may represent observations for several verbs (including the respective example verb).*

The intraparadigmatic type frequency (INTRA.TYPE) of the stem vowel $e$ was counted in terms of whether the past participle of a verb in the dataset is formed with this stem vowel (e.g. *gegeben*) or with the stem vowel $o$, as in *gestorben*. 36 verbs in the dataset have a past participle with $e$ (representing 947 instances in the dataset), 58 verbs have a past participle with the stem vowel $o$ (corresponding to 992 observations).

In order to arrive at the relative intraparadigmatic token frequencies of the stem vowels $i$ and $e$ (INTRA.TOKEN.E and INTRA.TOKEN.I), data from the Portal Wortschatz Universität Leipzig (Leipzig University 1998-2019) was used. The corpus was searched for all inflected forms with the stem vowel $e$ or $i$ of all verbs recorded in the dataset, in the following three phases: First, the corpus was searched for forms of the prefixed verbs in the dataset: they
keep their prefix in all forms, so that the query output does not need to be manually analysed
and the displayed token frequency counts can be directly transferred to a frequency table. In a
next step, all unambiguous forms of simplex and particle verbs were searched, i.e. paradigm
forms in which particles are not stranded. In the past participle, for example, the particle is not
stranded, so that the query output of aufgegeben ‘given up’ can be noted down for the verb
ausgehen, and the output of the search for gegeben ‘given’ for the verb geben. This is not pos-
sible, for example, for the third person singular form, where gibt may be followed by a
stranded particle, such as auf. These ambiguous cases were searched last, and the output was
inspected for whether the verb was a simplex or a particle verb. All instances found for one
and the same cell in the conjugation paradigm of a verb were added up. This procedure re-
resulted in a table of token counts for all inflected forms of all verbs in the dataset, exemplified
in Table 13 for the verbs geben ‘give’ and sterben ‘die’:

| Table 13: Intraparadigmatic token frequencies for forms of the verbs geben ‘give’ and sterben ‘die’ |
|---------------------------------|-----------------|-----------------|-----------------|
| paradigm form(s)                | geben           | sterben         |
|                                 | form | N  | %   | form   | N  | %   |
| imp sg                         | gib   | 822 | 0.09 | stirb   | 344 | 0.55 |
| 2nd ps sg pres ind             | gbst  | 375 | 0.04 | sturbst | 82  | 0.13 |
| 3rd ps sg pres ind             | gibt  | 434,412 | 49.02 | sturbst | 5,884 | 9.42 |
| 1st ps sg pres ind, (geb)      | gebe  | 50,375 | 5.69 | sterbe  | 309 | 0.50 |
| 1st/3rd ps sg pres subj        | gebt  | 1,316 | 0.15 | sterbt  | 14  | 0.02 |
| imp pl                         | geben | 140,895 | 15.90 | sterben | 12,502 | 20.01 |
| 1st/3rd ps pl pres subj,       | gebend | 86  | 0.01 | sterbend | 65  | 0.10 |
| present participle             | gegeben | 79,475 | 8.97 | gestorben | 15,948 | 25.53 |
| past participle                | gab   | 157,236 | 17.74 | starb   | 17,301 | 27.70 |
| 1st/3rd ps sg pret ind         | gabst | 17  | 0.00 | starbst  | 3   | 0.01 |
| 2nd ps sg pret ind             | gbt   | 2   | 0.00 | starbt   | 0   | 0.00 |
| 1st/3rd ps pl pret ind         | gaben | 9,419 | 1.06 | starben  | 9,937 | 15.91 |
| 1st/3rd ps sg pret subj        | gäbe | 11,359 | 1.28 | stürbe   | 15  | 0.22 |
| 2nd ps sg pret subj            | gäbst | 1   | 0.00 | stürtbst | 0   | 0.00 |
| 1st/3rd ps pl pret subj        | gäbt | 1   | 0.00 | stürtbst | 0   | 0.00 |
| 1st/3rd ps pl pres subj        | gäben | 381 | 0.04 | stürben  | 64  | 0.10 |
| total                          | 886,172 | 99.99 | 62,468 | 100 |

Note: ps = person, sg = singular, pl = plural, pres = present, pret = preterite, ind = indicative, imp = imperative, subj = subjunctive

As mentioned in 1.3, the relative token frequencies for all forms in the paradigm of one verb
with the stem vowel e and i, respectively, were added up. This yields relative intraparadig-
matic token frequencies for the verb geben of 61.56 % for the stem vowel i, 27.21 % for the
stem vowel e and 0.09 % for the imperative singular and values of 17.95 %, 36.67 % and 0.55 % of the respective frequencies for the verb *sterben*.

In contrast to Szmrecsanyi (2005; 2006) who could annotate PERSISTENCE variables automatically, this option was not available in the Walkthrough Corpus interface; therefore, they were annotated manually in the present study. In the walkthrough texts, a window of 20 words of context to the left of the imperative singular form recorded in the dataset was inspected for whether, and if so, which kind of singular imperative forms (hence pretargets) occurred. Because persistence needed to be annotated manually, 20 words was considered a manageable window size, given the number of data points. All the relevant forms were distinguished according to two criteria, verb type and suffixation. The variable (verb) TYPE has three factor levels, “Sve/iG” (short for strong verb with e/i-gradation), “other” for all verbs not belonging to the class of strong verbs with e/i-gradation, and “none” for no pretarget of either type occurring before the particular instance. The variable SUFF(ixation) also has three possible factor levels: “suffixed”, “unsuffixed”, and “none” for cases in which no pretarget occurred before the target instance. Up to five pretargets could be found in the predefined window; this amounted to 10 variables of the form LAST.SUFF, LAST.TYPE (for the very last pretarget occurring before the target instance), SECOND.SUFF, SECOND.VERB (for the second last occurring pretarget) etc.:  

<table>
<thead>
<tr>
<th>pretarget number</th>
<th>.TYPE</th>
<th>.SUFF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SVe/iG</td>
<td>other</td>
</tr>
<tr>
<td>LAST.</td>
<td>112</td>
<td>950</td>
</tr>
<tr>
<td>SECOND.</td>
<td>76</td>
<td>372</td>
</tr>
<tr>
<td>THIRD.</td>
<td>22</td>
<td>119</td>
</tr>
<tr>
<td>FOURTH.</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>FIFTH.</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>none</th>
<th>other</th>
<th>SVe/iG</th>
</tr>
</thead>
<tbody>
<tr>
<td>none</td>
<td>none × none</td>
<td>none × other</td>
<td>none × SVe/iG</td>
</tr>
<tr>
<td>suffixed</td>
<td>suffixed × none</td>
<td>suffixed × other</td>
<td>suffixed × SVe/iG</td>
</tr>
<tr>
<td>unsuffixed</td>
<td>unsuffixed × none</td>
<td>unsuffixed × other</td>
<td>unsuffixed × SVe/iG</td>
</tr>
</tbody>
</table>

Not only the individual effects of the verb type and suffixation of the pretargets but also the interaction between suffixation and verb type should be tested in the statistical model; however, the corresponding interaction matrix includes a number of impossible combinations. These combinations are marked in grey in Table 15:

A non-existent pretarget can neither be unsuffixed or suffixed nor can it belong to any verb type. For this reason, the interaction between verb type and suffixation was captured in a new
recoded variable LAST.PRETARGET with five possible factor levels, “none”, “suffixed other”, “suffixed SVe/iG”, unsuffixed other”, “unsuffixed SVe/iG” (and correspondingly for the remaining four pretargets).

Consecutive uses of imperative singular pretargets can be expected to reinforce or alter the persistence effect. This assumption is neglected in Szmreczanyi’s (2005; 2006) analyses in which he only annotated for the last pretarget occurring before the target form, but it should be taken into account in the present analysis. However, four impossible combinations are contained in the matrix for the interaction between two pretargets (grey shaded cells in Table 16):
SOCIAL META INFORMATION included in the authors’ member profiles on the website spieletipps.de was recorded in the corpus. As expected, the majority of members did not reveal any details about themselves or only part of the information required for a sociolinguistic analysis of the phenomenon under investigation. As concerns the variable AGE, information was available for 126 out of 927 authors. Thus, 412 out of the total 1,939 instances in the dataset were assigned an indication of the respective author’s age, though at the time of data retrieval in July 2013. Following the annotation of the variable POST.YEAR, this age statement was therefore altered according to the difference between 2013 and the post year, in order to arrive at the age at the time of writing the walkthrough (AGE.WRIT). As mentioned before, members of spieletipps.de can hyperlink their Facebook or other personal homepage in their member profile. All members’ profiles in the dataset were inspected for whether a link was included in them, and if so, whether the member’s Facebook or other site yielded additional social metainformation. In the case of the variable speaker age, this was possible for 43 instances (18 authors), resulting in 455 instances with an age statement. Due to the fact that the extraction of metainformation from external sources is tainted with a level of uncertainty, these 455 annotations were kept separate from the original 412 as the variable EXTERNAL.AGE, in order to be able to test their impact separately. In a third stage, the members’ usernames were inspected for whether they contained numbers which could indicate the member’s year of birth and/or age. Thus, only usernames including two or four digits were taken into consideration, such as “username85” or “username1985” as an indication of year of birth, “username8517” as a combination of year of birth and age, or “username12” or “username0412” as an indication of age at the time of registration or the year of registration plus age. In order to avoid skewness in the data, age statements extrapolated in this way were only taken over as an annotation if they kept within the range attested from the original AGE.WRIT variable, i.e. between 15 and 35 years. This procedure added an indication of age for another 171 instances (85 authors), amounting to 626 occurrences (229 authors) in the dataset annotated for a variable called ALL.AGE.

A second sociolinguistic variable, the authors’ GENDER, was annotated on the basis of the first name which the members indicated in their profiles. This metainformation was available for 259 instances in the dataset (corresponding to 71 authors). By consulting the members’ profiles on Facebook and personal websites, another 67 instances (29 authors) in the dataset could be assigned an indication of gender. This amounts to a total of 85 male authors contributing 273 instances and 15 female authors contributing 53 instances. In contrast to the variable age, gender information added in this step was not kept separate as the Face-
book profile offered both the members’ first names and pictures of them which validated the gender assignment. The annotation of author gender on the basis of Facebook is thus not tainted with the same degree of uncertainty as it was for the variable age. However, a third step of gender assignment was employed which did present this difficulty, i.e. the annotation of authors’ gender on the basis of their usernames. The result of this step was again annotated in a separate variable ALL.GENDER with information for an additional 461 instances by 238 authors (787 instances by 309 authors in total). In the variable ALL.GENDER, 604 received the annotation “male” (277 authors), 183 the annotation “female” (61 authors).

Only a minority of members revealed any regional sociolinguistic information, i.e. their PLACE and COUNTRY of residence. Only 121 instances (by 46 authors) could be annotated for the variable PLACE on the basis of the spieletipps.de profiles; through access to the members’ Facebook profiles, another 43 instances were assigned information, amounting to 161 instances (by 65 authors) annotated for PLACE. Even though place names such as “Unbekannt” ‘unknown’ or “X-box 360 Gehäuse” ’Xbox 360 case’ were not taken over, indications of the countries Germany, Austria or Switzerland in the same profiles were; thus, 180 occurrences in the dataset (by 75 authors) were annotated for COUNTRY. Note that the two levels of data retrieval were not kept separate because the Facebook profiles were considered as reliable as the spieletipps.de profiles in this respect. The information in PLACE introduced an unnecessary scattering of instances across the map of Germany (differentiating between, for example, a city and a village close by); it was therefore replaced by the variable DIALECT whose annotation is based on the map of German dialects in Figure 1.

<table>
<thead>
<tr>
<th>DIALECT</th>
<th>N</th>
<th>COUNTRY</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Alemannic</td>
<td>3</td>
<td>Switzerland</td>
<td>3</td>
</tr>
<tr>
<td>Central Southern Bavarian</td>
<td>29</td>
<td>Austria</td>
<td>31</td>
</tr>
<tr>
<td>Low Alemannic</td>
<td>17</td>
<td>Germany</td>
<td>146</td>
</tr>
<tr>
<td>Swabian</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>North Bavarian</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Berlin</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thuringian</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hessian</td>
<td>37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>North Saxon</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Palatinate</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Central Franconian</td>
<td>24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>East Franconian</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eastphalian</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Westphalian</td>
<td>15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The High Alemannic dialect region is not included in the map in Figure 1 because it covers large parts of Switzerland, outside of the Federal Republic of Germany (cf. the number of observations in this category of the variable COUNTRY).

Table 17: Distribution of variables DIALECT and COUNTRY in the dataset

75
The MORPHOLOGICAL STRUCTURE of the verb lemma was annotated using three factor levels: simplex verb, prefixed verb, or particle verb. The dataset contains 903 instances of simplex verbs (e.g. *sprechen* ‘speak’), 166 instances of prefixed verbs (e.g. *versprechen* ‘promise’), and 870 instances of particle verbs (e.g. *ansprechen* ‘address’).

Strong verbs with e/i-gradation belong to three of the seven ABLAUT classes of Middle High German. The assignment of ablaut classes to verbs in the dataset followed the explanation given by Schmidt (2007: 243-244). There are 25 verbs from ablaut class III in the dataset (176 instances), 34 verbs belong to ablaut class IV (771 instances) and 37 verbs to ablaut class V (992 instances).

As explained in 5.1.6, the imperative singular of strong verbs with e/i-gradation may be subject to phonological conditioning. The variable NUMBER OF SYLLABLES was counted in terms of the number of syllables in the traditional imperative variant; thus, verbs like *brechen* ‘break’ (traditional imperative form *brich*) received a value of “1” (17 verbs in the dataset, 1,732 observations), 18 verbs have 2 syllables like *durchbrechen* ‘break through’ (traditional imperative *durch|brich*; 192 observations), and 5 verbs such as *unterbrechen* ‘interrupt’ (traditional form *un|ter|brich*) received a value of “3” (15 observations). The PHONETIC CONTEXT was operationalised as the annotation of the sounds in the coda and nucleus of the preceding syllable and in the onset and nucleus of the following syllable. These variables have too many levels to be presented here; this high number of levels could have led to an inadequate assessment of the explanatory value of the predictor (cf. Section 5.3.2 for the case of the persistence variables). In order to avoid this, the phonetic context variables were entered into the regression model with all their levels; the resulting model was inspected for whether groupings could be found, such that, for example, close vowels would favour the occurrence of the stem vowel *i* in the imperative, whereas open and open-mid vowels would favour analogical imperative variants with the stem vowel *e*. Levels of the context variables were conflated accordingly, to check whether the original or recoded variables had a significant effect on stem vowel choice or suffixation in the imperative singular. LOCATION OF IMPERATIVE: All instances in the dataset were also annotated for whether they occurred at the beginning of a sentence, at the beginning of a clause or within the running sentence, for example preceded by an adverb like *dann* ‘then’. Among the total 1,939 instances, 1,121 imperatives occurred sentence-initially, 242 were used clause-initially, and 576 imperatives occurred within a clause. Finally, all data points were annotated for whether stem VOWEL LENGTH differed between the traditional and analogical variants: in 561 cases, the length of
the stem vowel is identical ([brɪç] and [brɛç]), and 1,378 cases exhibit a difference in vowel length ([nɪm] vs. [ne:m]).

The factor NEGATION was annotated using two variables. The first variable, NEGATION in the narrower sense, was coded according to whether or not the imperative singular form in the dataset was negated by means of the German negation particle nicht ‘not’ (39 instances received the annotation “yes”). The second variable, NEGATOR, took into account negation in the wider sense; instances in the dataset received the annotation “yes” when they were negated by the particles nicht ‘not’, the determiner kein/keine/keinen ‘no’, the adverbials nie/niemals ‘never’ or the indefinite pronoun nichts ‘nothing’ etc. (49 instances).

5.2.2.3 Random variables

Firstly, the verbs whose imperative singular forms were recorded in the dataset could show individual tendencies with regard to the stem vowel and suffixation they preferred to take in the imperative singular. Secondly, the authors of the recorded texts could show preferences (idiosyncrasies) in the choice of one of the three possible imperative forms. However, as the impact of the predictor variables on the dependent variables shall be determined regardless of these idiosyncrasies and peculiarities, the influence of these two factors needs to be controlled in the analysis by turning them into random variables.

5.3 Results

The present analysis is faced with a fundamental problem: Theoretically, imperative singular formation in strong verbs with e/i-gradation would need to be treated in a statistical model allowing for a dependent variable with three levels (unsuffixed imperative with the stem vowel i, unsuffixed imperative with the stem vowel e, suffixed imperative with the stem vowel e). Such multinomial regression models exist; however, the available models either do not fulfill all requirements for the present analysis - the function multinom in the R package nnet, for example, does not incorporate random factors - or they require statistical modelling of a sort which has not been used in linguistics to the extent that it would be easily transferred to the phenomenon under investigation; they could thus not be used as the central means of analysing the present dataset. Instead, separate mixed-effects logistic regression models were fitted to the data for the dependent variables STEM VOWEL and SUFFIXATION because these statistical models are restricted to categorical dependent variables with exactly two factor levels. Theoretically, these models should also make it possible to disentangle which of the two dependent variables is influenced by which of the predictor variables. However, the
present dataset contains almost only observations of the traditional imperative variant with the stem vowel \(i\) (30.79 \%) and the suffixed analogical variant with the stem vowel \(e\) (66.99 \%; 97.78 \% in total). This explains why the same predictor variables were found to have a significant influence on both stem vowel choice and suffixation and why their estimates are almost identical in the mixed-effects regression models for both dependent variables. Therefore, the following section first introduces the results of an additional analysis of the dataset in multinomial regression models which made it possible to attribute the influence of the predictors to either stem vowel choice or suffixation in the imperative singular. The results of the mixed-effects regression analysis are presented in Section 5.3.2, followed by the explanation and discussion of the effects of the individual predictors in the remainder of the chapter.

5.3.1 Results of the multinomial regression analysis

In the following, two functions for multinomial regression models in the programme R will be briefly introduced in order to clarify why multinomial regressions were rejected as the sole means of statistical analysis in the present study in favour of mixed-effects logistic regressions. More importantly, the results will make it possible to disentangle which of the predictors in the mixed-effects regression models (see Tables 20 and 21 in Section 5.3.2) has a significant influence on which aspect of variation in the dependent variable, i.e. stem vowel choice or suffixation in the imperative singular of strong verbs with e/i-gradation.

MULTINOM

The first of these multinomial regression functions is called `multinom` (R package nnet, version 7.3-12, February 2016; Venables & Ripley 2002). It allows for a categorical dependent variable with more than two levels; thus, the model can be specified to take up the dependent variable IMPERATIVE with three levels: the traditional unsuffixed imperative singular forms of strong verbs with e/i-gradation with the stem vowel \(i\) (the reference level), unsuffixed analogical forms with the stem vowel \(e\) and suffixed analogical forms with the stem vowel \(e\). The same predictors were entered which turned out to have a significant effect in the mixed-effects logistic regressions: base verb lemma token frequency (PC1.BASE), persistence (LAST.SECOND.7\(^{21}\)) and time (POST.YEAR). Due to the fact that multinom does not incorporate random effects, the idiosyncratic effects of VERB and AUTHOR could not be controlled for. The regression table below (Table 18) presents the estimates for the use of the two analogical imperative singular variants in comparison to the traditional imperative variant with the stem vowel \(i\). The results of the multinom model show that with varying base verb

\(^{21}\) Please see Section 5.3.2 for an explanation of the variable LAST.SECOND.7 and its factor levels.
lemma token frequency and posting year, the probabilities for the use of both analogical imperative singular variants are significantly lower than that of the reference level, the traditional imperative variant. Thus, these two independent variables influence the stem vowel of the imperative singular.

In contrast, whereas the probability for the use of the suffixed analogical imperative variant increases significantly when suffixed pretargets occur in the preceding context (“suffixed”, “suffixed SVe/iG”), the probability for the unsuffixed analogical variant is either lower (“suffixed”) or not significantly different (“suffixed”) from that for the likewise unsuffixed

Table 18: Results of the multinom regression analysis of imperative singular formation (Krause-Lerche 2019a: 20)

<table>
<thead>
<tr>
<th>PREDICTOR</th>
<th>Factor level</th>
<th>IMPERATIVE</th>
<th>estimate</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>unsuff e</td>
<td>-3.158</td>
<td>0.0000 ***</td>
<td></td>
</tr>
<tr>
<td>(Intercept)</td>
<td>suff e</td>
<td>-0.323</td>
<td>0.1647</td>
<td></td>
</tr>
<tr>
<td>PC1.BASE</td>
<td>unsuff e</td>
<td>-0.215</td>
<td>0.0196 *</td>
<td></td>
</tr>
<tr>
<td>PC1.BASE</td>
<td>suff e</td>
<td>-0.150</td>
<td>0.0000 ***</td>
<td></td>
</tr>
<tr>
<td>bs(POST.YEAR, knots=c(2005, 2009), degree=1)1</td>
<td>unsuff e</td>
<td>-0.840</td>
<td>0.3130</td>
<td></td>
</tr>
<tr>
<td>bs(POST.YEAR, knots=c(2005, 2009), degree=1)1</td>
<td>suff e</td>
<td>-0.128</td>
<td>0.6714</td>
<td></td>
</tr>
<tr>
<td>bs(POST.YEAR, knots=c(2005, 2009), degree=1)2</td>
<td>unsuff e</td>
<td>-0.708</td>
<td>0.2816</td>
<td></td>
</tr>
<tr>
<td>bs(POST.YEAR, knots=c(2005, 2009), degree=1)2</td>
<td>suff e</td>
<td>-0.516</td>
<td>0.0175 *</td>
<td></td>
</tr>
<tr>
<td>bs(POST.YEAR, knots=c(2005, 2009), degree=1)3</td>
<td>unsuff e</td>
<td>-0.421</td>
<td>0.6525</td>
<td></td>
</tr>
<tr>
<td>bs(POST.YEAR, knots=c(2005, 2009), degree=1)3</td>
<td>suff e</td>
<td>-0.460</td>
<td>0.1087</td>
<td></td>
</tr>
<tr>
<td>LAST.SECOND.7</td>
<td>unsuff e</td>
<td>reference level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>no pretarget</td>
<td>suff e</td>
<td>reference level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>last unsuffixed, second suffixed</td>
<td>unsuff e</td>
<td>1.186</td>
<td>0.0737</td>
<td></td>
</tr>
<tr>
<td>last unsuffixed, second suffixed</td>
<td>suff e</td>
<td>-0.787</td>
<td>0.0589</td>
<td></td>
</tr>
<tr>
<td>last suffixed, second unsuffixed</td>
<td>unsuff e</td>
<td>0.818</td>
<td>0.2076</td>
<td></td>
</tr>
<tr>
<td>last suffixed, second unsuffixed</td>
<td>suff e</td>
<td>-0.700</td>
<td>0.0596</td>
<td></td>
</tr>
<tr>
<td>suffixed</td>
<td>unsuff e</td>
<td>-0.182</td>
<td>0.6791</td>
<td></td>
</tr>
<tr>
<td>suffixed</td>
<td>suff e</td>
<td>0.734</td>
<td>0.0000 ***</td>
<td></td>
</tr>
<tr>
<td>suffixed SVe/iG</td>
<td>unsuff e</td>
<td>-5.791</td>
<td>0.9281</td>
<td></td>
</tr>
<tr>
<td>suffixed SVe/iG</td>
<td>suff e</td>
<td>3.329</td>
<td>0.0000 ***</td>
<td></td>
</tr>
<tr>
<td>unsuffixed</td>
<td>unsuff e</td>
<td>1.267</td>
<td>0.0012 **</td>
<td></td>
</tr>
<tr>
<td>unsuffixed</td>
<td>suff e</td>
<td>-0.478</td>
<td>0.0223 *</td>
<td></td>
</tr>
<tr>
<td>unsuffixed SVe/iG</td>
<td>unsuff e</td>
<td>-5.926</td>
<td>0.8203</td>
<td></td>
</tr>
<tr>
<td>unsuffixed SVe/iG</td>
<td>suff e</td>
<td>-1.009</td>
<td>0.1032</td>
<td></td>
</tr>
</tbody>
</table>
traditional imperative variant in these surroundings. This is to show that the predictor persistence has direct influence only on the suffixation of the imperative singular forms.

Even though multinom needed to be rejected as the sole means of analysing the present dataset because it does not generalise over the idiosyncratic effects of authors and verbs, on the basis of the regression results it is possible to attribute effects directly to each of the three imperative singular variants.

MCMCGLMM
Another function for multinomial modelling of data in the programme R is called \textit{MCMCglmm}; this function and the corresponding R package of the same name were developed by the evolutionary biologist Jarrod Hadfield (2010). It allows for a categorical dependent variable and incorporates random effects for specified variables. Thus, it does in theory fulfil the necessary criteria for the present investigation, but it requires statistical modelling which has not or seldom been used in the field of linguistics so far. It fits Generalised Linear Mixed Models (hence “glmm”) using Markov Chain Monte Carlo (“MCMC”) sampling, a procedure known from Bayesian statistics. It is aimed at biological, e.g. phylogenetic, research, where both statistical procedures are widely used. In linguistics, however, Bayesian statistics and MCMC sampling techniques have only just gained a foothold\textsuperscript{22}. The function MCMCglmm, for example, requires the definition of priors; whether or not these have been set correctly can be assessed by an inspection of the posterior distribution of the effects. Both concepts are not used in “classical” (restricted) maximum likelihood estimation (REML). This serves to show that familiarisation with the procedures and proper testing of MCMCglmm models would go well beyond the scope of the present project. Although the following example MCMCglmm model was set up with the help of Jarrod Hadfield’s reply on the R-sig-mixed-models mailing list, this illustration must therefore be treated with caution.

As was done in the multinom regression model before, the MCMCglmm model was specified to take up the dependent variable IMPERATIVE, the predictors base verb lemma token frequency (PC1.BASE), persistence (LAST.SECOND.7) and time (POST.YEAR), but the random variables AUTHOR and VERB could be added in this model. The results of the model are presented in Table 38 in Appendix D; they are largely identical with the results of the multinom model above, but they are more conservative because the model takes into account the effect of the random variables VERB and AUTHOR. For the predictors PC1.BASE and POST.YEAR, the model shows that the estimates for both analogical imperative singular

\textsuperscript{22} Notable exceptions are papers by Gray et al. (2009), investigating language phylogenies in the Pacific, and Johnson et al. (2007) on probabilistic context-free grammars.
variant have a negative sign: with varying base verb lemma token frequency or time, the probability for the use of both analogical variants decreases. Thus, the model shows that these two variables affect the stem vowel of the imperative singular forms. In contrast, the estimates have different signs for the two analogical variants for almost all levels of the variable persistence (LAST.SECOND.7). As in the results of the multinomial regression, suffixed contexts ("suffixed", "suffixed SVe/iG") favour the use of the suffixed analogical variant, whereas the probability for the unsuffixed analogical variant does not increase in these contexts. This serves to show that in both multinomial regression models, persistence directly affects only the suffixation of the imperative singular forms.

5.3.2 Results of the mixed-effects regression analyses

Two regression models were fitted to the dataset of imperative singular observations in the Walkthrough Corpus, one for the dependent variable STEM VOWEL and one for the SUFFIXATION of the imperatives. The random variables entered in both models are the AUTHOR of the text and the VERB whose imperative singular is recorded in the dataset. As anticipated in the previous section, the predictors which have a significant effect on stem vowel choice or suffixation are time (POST.YEAR), base verb lemma token frequency (PC1.BASE) and persistence (LAST.SECOND.7), i.e. the interaction between the last two pretargets occurring within a window of 20 words of context to the left of the target instance. Based on the results of these multinomial regression models, the final mixed-effects regression tables will only report results for predictors which have a direct influence on the respective aspect of variation in the imperative singular of strong verbs with e/i-gradation, i.e. base verb lemma token frequency and time in the case of stem vowel choice, and persistence in the case of suffixation.

The final mixed-effects regression models in the present study were arrived at in a forward selection process. First, two baseline regressions were fitted to the data, with only the respective dependent variable, the random variables and the model intercept. When the fixed factor POST.YEAR was added, this significantly improved the quality of the models in explaining the variation in the dependent variables (stem vowel: p < 0.01; suffixation: p < 0.05), but it turned out upon further inspection that this linear modelling of the data did not do justice to the real temporal development. Section 5.3.5 will explain how and why the influence of time was therefore modelled using a so-called B-spline, i.e. a non-linear, transforma-

---

23 In Table 20 and Table 21, some factor levels of the categorical predictor LAST.SECOND.7 do not show significant contrasts with the reference level, and not all components of the B-spline modelled predictor time (POST.YEAR) are significant. However, this does not mean that the respective predictor itself does not have a significant effect on the dependent variable. For categorical and non-linear predictors, the latter question can only be answered by means of (log-likelihood) comparisons between regression models with and without the respective predictor (cf. Baayen 2008: 167; Levshina 2015: 267).
tion of POST.YEAR, which further improved the quality of the regression models ($p < 0.001$). The quality of the models was not improved by adding the fixed factor verb lemma token frequency (either PC1.VERB or PC2.VERB), but adding the variable base verb lemma token frequency (PC1.BASE) did improve them ($p < 0.001$).

The influence of persistence was tested in several stages: First, the variables capturing suffixation and verb type of the individual pretargets (e.g. LAST.TYPE, LAST.SUFF) were added to the regression models for both dependent variables. This showed that only the suffixation and verb type of the last two pretargets in the predefined window had a significant influence on stem vowel choice and suffixation. The interactive effect of suffixation and verb type of these two pretargets was tested by replacing the variables LAST.SUFF and LAST.TYPE by LAST.PRETARGET and the variables SECOND.SUFF and SECOND.TYPE by SECOND.PRETARGET ($p < 0.001$).

Finally, in order to find out whether the presence and form of two consecutive pretargets reinforces or alters the persistence effect, the variables LAST.PRETARGET and SECOND.PRETARGET were replaced by LAST.SECOND. This did not significantly improve model quality, however, mainly because the contrasts between many of the 21 factor levels of LAST.SECOND were not significant. Therefore, as illustrated in Figure 12 and Table 19, some very similar factor levels of this variable were conflated into one of seven factor levels in a recoded variable LAST.SECOND.7. The final regression models with LAST.SECOND.7
as a fixed factor performed significantly better at explaining the variation in the dependent variables than the previous models, which contained the last and second last pretarget as separate fixed factors and thereby ignored the linear order between them and with respect to the target imperative form.

The influence of the individual significant predictors in the regression models will be described and discussed in Chapters 5.3.3 to 5.3.5. On the basis of the results of the multinomial regression analyses, base verb lemma token frequency (PC1.BASE) and time (POST.YEAR) will be argued to have a direct influence on the stem vowel of the imperative singular forms of strong verbs with e/i-gradation in the dataset, and only an indirect impact on their suffixation. For the persistence variable (LAST.SECOND.7), the opposite argumentation is presented, i.e. direct influence on suffixation, indirect impact on the stem vowel.

### Table 19: Conflation of factor levels of variable LAST.SECOND in variable LAST.SECOND.7

<table>
<thead>
<tr>
<th>Level index</th>
<th>Levels of variable LAST.SECOND</th>
<th>N</th>
<th>Distinguishing criterion in LAST.SECOND.7</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>last and second unsuff SVe/iG</td>
<td>6</td>
<td>two unsuff pretargets, at least one of them is a form of a SVe/iG</td>
<td>27</td>
</tr>
<tr>
<td>2</td>
<td>last unsuff SVe/iG, second unsuff other</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>last unsuff other, second unsuff SVe/iG</td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>last and second unsuff other</td>
<td>52</td>
<td>at least one unsuff pretarget, no more than one of them is a form of a SVe/iG</td>
<td>197</td>
</tr>
<tr>
<td>5</td>
<td>last unsuff other, no second</td>
<td>99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>last unsuff SVe/iG, no second</td>
<td>46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>last unsuff SVe/iG, second suff SVe/iG</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>last unsuff other, second suff other</td>
<td>43</td>
<td>last unsuff, last pretarget unsuff, second pretarget suff</td>
<td>55</td>
</tr>
<tr>
<td>9</td>
<td>last unsuff SVe/iG, second suff other</td>
<td>10</td>
<td>last pretarget suff, second pretarget suff</td>
<td>63</td>
</tr>
<tr>
<td>10</td>
<td>last unsuff other, second suff SVe/iG</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>no pretarget</td>
<td>877</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>last suff other, second unsuff SVe/iG</td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>last suff SVe/iG, second unsuff other</td>
<td>1</td>
<td>last suff, last pretarget suff, second pretarget suff</td>
<td>63</td>
</tr>
<tr>
<td>14</td>
<td>last suff other, second unsuff other</td>
<td>37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>last suff SVe/iG, second unsuff SVe/iG</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>last suff SVe/iG, no second</td>
<td>17</td>
<td>at least one suffix pre- target, no more than one of them is a form of a SVe/iG</td>
<td>680</td>
</tr>
<tr>
<td>17</td>
<td>last suff other, no second</td>
<td>452</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>last and second suff other</td>
<td>211</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>last suff other, second suff SVe/iG</td>
<td>15</td>
<td>two suffix pre-targets, at least one of them is a form of a SVe/iG</td>
<td>40</td>
</tr>
<tr>
<td>20</td>
<td>last suff SVe/iG, second suff other</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>last and second suff SVe/iG</td>
<td>11</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Level index - see x-axis Figure 12, N = number of observations, last = last pretarget, second = second pre-target, suff = suffixed, unsuff = unsuffixed, SVe/iG = strong verb with e/i-gradation.
Table 20: Results of mixed-effects regression for dependent variable STEM VOWEL (Krause-Lerche 2019a: 18)

<table>
<thead>
<tr>
<th>Fixed factors</th>
<th>estimate</th>
<th>standard error</th>
<th>z value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.619</td>
<td>0.547</td>
<td>-1.132</td>
<td>0.2576</td>
</tr>
<tr>
<td>PC1.BASE</td>
<td>-0.503</td>
<td>0.125</td>
<td>-4.016</td>
<td>0.0001 ***</td>
</tr>
<tr>
<td>bs(POST.YEAR, knots=c(2005,2009), degree=1)1</td>
<td>1.089</td>
<td>0.641</td>
<td>1.699</td>
<td>0.0894</td>
</tr>
<tr>
<td>bs(POST.YEAR, knots=c(2005,2009), degree=1)2</td>
<td>-0.774</td>
<td>0.465</td>
<td>-1.666</td>
<td>0.0957</td>
</tr>
<tr>
<td>bs(POST.YEAR, knots=c(2005,2009), degree=1)3</td>
<td>0.260</td>
<td>0.659</td>
<td>0.394</td>
<td>0.6934</td>
</tr>
</tbody>
</table>

Random factors

<table>
<thead>
<tr>
<th>variance</th>
<th>standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUTHOR</td>
<td>4.742</td>
</tr>
<tr>
<td>VERB</td>
<td>1.362</td>
</tr>
</tbody>
</table>

Measures of model quality

| N         | 1,931              |
| AIC       | 1928.1              |
| BIC       | 2000.5              |
| degrees of freedom | 13 |
| R^2 marginal | 0.1139 |
| R^2 conditional | 0.6897 |

Table 21: Results of mixed-effects regression for dependent variable SUFFIXATION (Krause-Lerche 2019a: 19)

<table>
<thead>
<tr>
<th>Fixed factors</th>
<th>Factor levels</th>
<th>estimate</th>
<th>standard error</th>
<th>z value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>no pretarget</td>
<td>-1.135</td>
<td>0.632</td>
<td>-1.797</td>
<td>0.0723</td>
</tr>
<tr>
<td>LAST.SECOND.7</td>
<td>last unsuffixed, second suffixed</td>
<td>-1.252</td>
<td>0.716</td>
<td>-1.748</td>
<td>0.0805</td>
</tr>
<tr>
<td></td>
<td>last suffixed, second unsuffixed</td>
<td>-0.427</td>
<td>0.594</td>
<td>-0.719</td>
<td>0.4723</td>
</tr>
<tr>
<td></td>
<td>suffixed</td>
<td>1.227</td>
<td>0.228</td>
<td>5.379</td>
<td>0.0000 ***</td>
</tr>
<tr>
<td></td>
<td>suffixed SVe/iG</td>
<td>3.808</td>
<td>0.857</td>
<td>4.445</td>
<td>0.0000 ***</td>
</tr>
<tr>
<td></td>
<td>unsuffixed</td>
<td>-0.082</td>
<td>0.364</td>
<td>-0.224</td>
<td>0.8226</td>
</tr>
<tr>
<td></td>
<td>unsuffixed SVe/iG</td>
<td>0.772</td>
<td>0.987</td>
<td>0.782</td>
<td>0.4341</td>
</tr>
</tbody>
</table>

Random factors

<table>
<thead>
<tr>
<th>variance</th>
<th>standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUTHOR</td>
<td>6.646</td>
</tr>
<tr>
<td>VERB</td>
<td>1.963</td>
</tr>
</tbody>
</table>

measures of model quality

| N         | 1,925              |
| AIC       | 1802.7              |
| BIC       | 1875.0              |
| degrees of freedom | 13 |
| R^2 marginal | 0.1336 |
| R^2 conditional | 0.7605 |
The impact of sociolinguistic variables on the formation of the imperative singular observations in the dataset needed to be tested in separate regression models; the reasons for this procedure and the results of the respective statistical models are presented in Section 5.3.6.

5.3.3 Frequency effects

From previous usage-based research, it was assumed that the frequency of a verb should have an influence on the trajectory of analogical levelling in strong verbs with e/i-gradation: Due to their weaker entrenchment, verbs of low token frequency succumb to analogical levelling first, whereas verbs with a high token frequency have stronger mental representations and thus resist analogical replacement longer, i.e. they exhibit a “conserving effect” (Bybee & Thompson 1997: 380). This is exactly the effect illustrated in Figure 13, computed from the regression model for stem vowel choice (cf. Table 20): the verbs in the dataset with the highest token frequency show the lowest probability of occurring in analogical imperative singular variants with the stem vowel e, and this probability increases gradually along the frequency continuum up to the verbs with the lowest token frequency. Note, however, that it is not verb lemma token frequency which has this effect but base verb lemma token frequency; only the latter was shown to have a significant influence on stem vowel choice in the regression model.

Figure 13: Effect of base verb lemma token frequency (PC1.BASE) on stem vowel choice in imperative singular (Krause-Lerche 2019a: 22)

Notes: The higher the value of PC1.BASE (x-axis), the higher the base verb lemma token frequency (see 5.2.2). Grey shaded area indicates confidence intervals.
The predictors PC2.BASE, PC1.VERB, PC2.VERB, INTRA.TYPE, INTRA.TOKEN.E, INTRA.TOKEN.I and INTRA.TOKEN.IMP were entered into the regression model as well; none of them had a significant effect on the dependent variable.

5.3.4 Persistence effects

Bearing in mind the conserving effect of high token frequency identified in the regression analysis, it is interesting to see that a number of very high frequent base verbs nevertheless occur with analogical imperative singular forms in the present dataset. One way of accounting for this observation may be the counter-influence of persistence (Szmrecsanyi 2005; 2006).

The formation of the imperative singular instances of strong verbs with e/i-gradation in the dataset is assumed to be influenced by the form of preceding imperative singular forms: when previous imperative singular items are used with a suffix, the author might develop a routine by which they wish to suffix imperative singular forms of strong verbs with e/i-gradation as well (e.g. *gebe*). The reverse phenomenon can be expected for unsuffixed pretargets, resulting in a higher probability of use for unsuffixed target imperative variants (e.g. *gib* or *geb*).

The results of the regression model for suffixation in Table 21 show that the form of two preceding imperative singular forms indeed has a significant effect on the suffixation of the target imperative singular form of a strong verb with e/i-gradation recorded in the dataset.

![Figure 14: Interactive effect of last and second last pretarget (variable LAST.SECOND.7) on suffixation of target imperative singular (Krause-Lerche 2019a: 24)](image)

Note: Bars indicate confidence intervals; last = last pretarget, second = second pretarget, SV e/iG = strong verb with e/i-gradation

This effect is illustrated in Figure 14: when no imperative singular form occurs in a window of 20 words before the target imperative singular form (x-axis level “no pretarget”), then the
probability for the target imperative to be suffixed is very low (7%). When all preceding imperative singular form are unsuffixed (“unsuffixed”, “unsuffixed SVe/iG”) or when both suffixed and unsuffixed imperative forms precede the target imperative (“last unsuffixed, second suffixed” or “last suffixed, second unsuffixed”), then the latter is similarly not very likely to be suffixed (between 1.9 and 12.5% probability).

However, the situation changes when only suffixed pretargets precede the target imperative form of a strong verb with e/i-gradation. In the condition “suffixed”, i.e. when only one suffixed pretarget occurs or two suffixed pretargets which are imperative singular forms of “other” verbs (not strong verbs with e/i-gradation), the probability for suffixation of the target imperative form rises to 20.4%. Except for the contrast with the level “unsuffixed SVe/iG”, which comprises only 27 data points, this probability is significantly higher than in conditions with unsuffixed pretargets or without pretargets (p < 0.01). When two suffixed pretargets precede the target imperative of which at least one is a form of a strong verb with e/i-gradation (level “suffixed SVe/iG”), then the probability for the target imperative to be suffixed is 77.2%; this probability is significantly higher than in all other conditions (p < 0.05).

Up to five pretargets occurred in the specified window of 20 words of context to the left of the target imperative singular form; however, as mentioned in Section 5.3.2, the suffixation and verb type of the third, fourth and fifth pretarget do not have a significant effect on the form of the target imperative singular instances in the dataset.

5.3.5 Development over time - variation or change?
Perhaps most importantly, the present project set out to determine whether analogical leveling in imperative singular formation of strong verbs with e/i-gradation must be regarded as a genuine change-in-progress or “simple” linguistic variation, hereby preventing that “a diachronic trend is read into a situation which merely shows variable or fluctuating usage” (Mair 2008: 1111). Arguably, the present data set is not particularly well-suited to answer this question, but given that it is the largest data set which can be extracted from existing corpora of German, this section will examine the development across the 13-year time span covered by the data.

To this end, as explained in Section 5.3.2, the variable POST.YEAR was originally entered into the regression model for the dependent variable STEM VOWEL as a linear predictor, as is done per default in (generalised) linear regression models. The effect of the variable POST.YEAR in this model suggests an unexpected temporal development which is illustrated in Figure 15:
Contrary to expectations, the regression model seems to reveal that instead of stable variation between the stem vowels $i$ and $e$ in the imperative singular of strong verbs with $e/i$-gradation or a change-in-progress from the former to the latter, the change is in fact moving in the opposite direction. Thus, it appears that speakers at the beginning of the century did use a number of analogical imperative forms; however, they slowly but surely went back to using the traditional imperative variant with the stem vowel $i$. This suggests an irregularisation trend instead of the expected regularising process of analogical levelling, but this impression is likely to be an artefact of the statistical model applied in the analysis.

**B-SPLINE TRANSFORMATION**

It is important to recall that the dataset has been analysed using a generalised linear mixed-effects regression. Such regression models assume a linear development for all numeric predictor variables; however, this kind of modelling of the temporal development does not do justice to the fluctuations in use of the traditional and analogical imperative singular variants over time which become obvious in Figure 16. This figure displays the effect of the variable POST.YEAR when it is entered into the regression as a categorical predictor (see Gries & Hilpert 2008, 2010, 2012 for a similar bottom-up approach for the identification of developmental stages in diachronic data). It suggests an increase in the probability of the use of analogical imperative singular forms until approximately the year 2005; this probability decreases until it reaches a low peak around the year 2009, from where it seems to increase again.
Among the different algorithms and functions available for modelling non-linear developments are breakpoint regressions (Baayen 2008: Ch. 6.4), logarithmic and polynomial transformations of predictors and restricted cubic splines (e.g. Baayen 2008: 30-32, Ch. 6.2.1; Hastie & Pregibon 1993: 219-220), B-splines and natural splines (Hastie 1993: Ch. 7.2.4; Venables & Ripley 2002: Ch. 8.7). As explained in Section 5.3.2, adding POST.YEAR as a linear predictor to the baseline regression models fitted to the Walkthrough Corpus data significantly improved model quality. In an attempt to represent the fluctuations in stem vowel choice displayed in Figure 16 more accurately in the regression model, however, all the non-linear transformations and regressions were tested and compared with the linear model variant. In the present study, the regression analyses with the best fit contained POST.YEAR modelled as a B-spline term (using the function bs() in R package splines). Instead of the uniform linear behaviour assumed per default for numerical predictors in generalised linear models, the relationship between the numerical predictor and the dependent variable is thus modelled as a “piecewise polynomial [with] knots [emphasis original] or breakpoints at prechosen places in the range of the predictor. These knots separate disjoint regions in the data, and the regression function is modeled as a separate polynomial piece in each region” (Hastie 1993: 271). In other words, the development of stem vowel choice in the imperative singular of strong verbs with e/i-gradation over time (POST.YEAR) is modelled as non-linear insofar that the temporal curve exhibit so-called knots or breakpoints. Up to these knots, there is a certain (linear, quadratic, cubic etc.) relation between the predictor and the dependent variable in question; around the knot, the data show a discontinuity. After the breakpoint, the relation

**Figure 16:** Probability of stem vowel e in imperative singular over time (POST.YEAR as a categorical predictor)
between the two variables may be different in slope, e.g. a steeper or slower increase or decrease, or even direction, perhaps an increase before and a decrease after the knot.

The function \texttt{bs()} for B-splines in R(Studio) allows the researcher to specify the location of the knots or the number of degrees of freedom, and the degree of the spline; these different options were tested in the present study to arrive at the most appropriate statistical model of the corpus data. In the final model, the years 2005 and 2009 were specified as the internal knots of the curve, and the degree parameter was set to 1 instead of the default 3, resulting in a linear instead of a cubic spline. That means, the relationship between time and stem vowel choice is defined as non-linear insofar as the regression line is not a single straight line, but a line consisting of separate linear pieces whose direction and/or slope may change at the knots. This procedure significantly improved the quality of the regression model in explaining the variation in the dependent variable compared to the default linear regression (p < 0.005). Log-likelihood comparisons also show that the model reported in Table 20 is significantly better than a model with only two degrees of freedom (corresponding to 1 knot; p < 0.001). Thus, it is not enough to assume only one breakpoint in the temporal development, for example, a slight decrease in the use of analogical variants until approximately 2007, and stable variation or fluctuation in use from then on. The reported final regression model is also significantly better (p < 0.001) than a model in which the degrees of freedom of the spline (instead of the knots) are specified so that the function \texttt{bs()} chooses the knots autonomously and “uniformly along the range of the predictor” (Hastie 1993: 271), for example in the years 2004 and 2008.

Figure 17 below compares the regression line for \texttt{POST.YEAR} computed from the final non-linear regression model reported in Table 20 with the regression line for this predictor computed from a linear variant of the final model. It spells out clearly that not only statistically speaking, the non-linear regression is a significant improvement. It also provides a more accurate picture of the temporal development of imperative singular formation in strong verbs with e/i-gradation: analogical imperative singular variants formed with the stem vowel $e$ are on the increase between 2000 and 2005 and between 2009 and 2013, but a temporary decrease in their use can be observed between 2005 and 2009.
5.3.6 Influence of Sociolinguistic Variables

As explained in Chapter 5.2.2, sociolinguistic metainformation was not available for all data points. As a result, separate regression models were fitted on subsets of the data frame which contained the necessary information in order to determine the influence of the sociolinguistic variables on imperative singular formation of strong verbs with e/i-gradation. The random factors VERB and AUTHOR and the predictors base verb lemma token frequency (PC1.BASE), persistence (LAST.SECOND.7) and time (POST.YEAR as a B-spline) from the regression models reported in Tables 20 and 21 were retained. The regressions for the variables AGE, GENDER and COUNTRY showed that none of these factors has a significant impact on either stem vowel choice or suffixation of the imperative singular observations in the reduced datasets. This outcome can be explained as an effect of the scarcity of sociolinguistic meta information in the dataset. Moreover, a high amount of variance in the dependent variables is explained by the random factors AUTHOR and VERB (cf. Tables 20 and 21), but leaving them out of the regression models would mean that the results were tainted by the individual preferences of speakers or verbs.

In the case of the variable DIALECT, the regression model needed to be modified because the random factor AUTHOR presented a problem. As explained in Section 5.2.2, the
latter parameter was introduced in all regression models to control for personal preferences by individual authors for either of the two imperative singular stem vowels of strong verbs with e/i-gradation or a general inclination toward or against suffixation of the imperative singular of these verbs. To some extent, this preference or inclination may be caused by the dialect which the authors speak; thus, by entering AUTHOR as a random variable in a regression model, dialect-induced preferences are controlled. This problem is illustrated in the caterpillar plot in Figure 18 for the dependent variable STEM VOWEL: While in Central Franconian, Hessian and Westphalian dialect regions, some authors prefer to use the stems vowel e in the imperative singular of strong verbs with e/i-gradation and others prefer the stem vowel i, authors from the Berlin area, the Palatinate and the North Saxon dialect region show a clear preference for imperative singular forms with the stem vowel e.

![Figure 18: Intercept adjustment for random factor AUTHOR removes dialect-induced stem vowel preferences](image)

Notes: Each dot stands for one author. Positive x-values indicate a preference for the use of the stem vowel e by the respective author, negative values indicate a preference for the stem vowel i.

Therefore, the regression model for the predictor DIALECT was run without the random factor AUTHOR (but retaining the random factor VERB and the predictors). Nevertheless, this sociolinguistic variable turned out not to have a significant impact on stem vowel choice or suffixation of the imperative singular instances in the reduced dataset, either (p = 0.0956 for dependent variable STEM VOWEL p = 0.1556 for SUFFIXATION).

According to the dialectological literature (cf. Section 1.4.2), speakers of Lower German and West Middle German dialects and Thuringian (black points in Figure 19 below) use the stem vowel e in the imperative singular of strong verbs with e/i-gradation, while speakers of Upper German dialects and Upper Saxon (grey points in Figure 19) use the stem vowel i.
However, the only group which uses significantly more imperative singular forms with the stem vowel e than others (e.g. Low Alemannic, p < 0.05) in the present regression model are speakers from the North Saxon dialect region. Thus, apart from a lack of meta data about authors, the reason for the nonsignificant effect of the predictor DIALECT may lie in the observation that the majority of authors in the dataset use the standard imperative singular variant with the stem vowel i.

Figure 19: Effect of DIALECT on dependent variable STEM VOWEL

5.4 Discussion of results

Stepwise regression modelling showed that the variable verb lemma token frequency (PC1.VERB and PC2.VERB), distinguishing between the token frequency of, for example, the simplex verb geben, the prefixed verb übergeben and the particle verb aufgeben, does not have a significant effect on the formation of the imperative singular of strong verbs with e/i-gradation, whereas base verb lemma token frequency (PC1.BASE) has a significant effect on stem vowel choice in the imperative singular (see Table 20). Thus, it must be concluded that speakers do not differentiate between simplex verbs and simplex bases of particle verbs when forming an imperative singular. In the light of the syntax of German particle verbs, this finding is not surprising: as illustrated in example (15), the finite component of the imperative singular form of a simplex verb (e.g. geben) and particle verbs derived from it (like wiedergeben or zurückgeben) is identical, whereas the particle component is always non-finite and can be stranded in an imperative clause as far as necessary.


‘Give me the book back (tomorrow) (after lunch) (in the office).’
The second principle component of base verb lemma token frequency (PC2.BASE) did not show a significant effect. This finding does not undermine the overall effect of base verb lemma token frequency, however, as the second principle component mainly served the purpose to rotate the frequency scale from Ruoff’s (1990) frequency dictionary into a common two-dimensional space with the remaining reference frequency scales (cf. Section 5.2.2.2). Thus, this finding merely serves to show that genre- or register-specific frequencies (for example from rural language) do not have an effect on imperative singular formation. In the present analysis, none of the intraparadigmatic frequency variables (INTRA.TYPE, INTRA.TOKEN.E, INTRA.TOKEN.I, INTRA.TOKEN.IMP) exhibit a significant effect on imperative formation, either; that means that speakers recorded in the corpus are neither driven by the fact that some verbs have a higher number of forms with the stem vowel e available in their paradigm, nor are they oriented towards the stem vowel which accounts for the majority of tokens of a verb.

These findings strongly suggest that speakers generally act according to the productive verb inflection pattern or schema when they wish to use an imperative singular form of a strong verb with e/i-gradation. They access the entry of the respective base verb in their mental lexicon and, instead of orientating towards the prevailing stem vowel in a paradigm, they apply the pattern for forming an imperative singular which they know from the majority of verbs: [[INFINITIVE STEM](-e)]. However, the higher the lemma token frequency of a base verb, the less likely it becomes that the imperative singular is formed according to this productive pattern. Instead, the irregular form is retrieved from the mental lexicon as a unit. Thus, the conserving effect of high token frequency takes effect in ongoing morphological change in German as well: depending on the frequency of a verb, speakers are more or less likely to use the traditional imperative singular variant with the stem vowel i or an analogical variant with the stem vowel e. It has also become apparent that even in more highly inflected language like German, speakers do not keep track of the number of stem vowel forms available in a paradigm or the number of times they encounter these forms, i.e. they do not seem to register intraparadigmatic frequencies; instead, they are “only” influenced by the number of times they encounter the particular (base) verb.

The fact that base verb lemma token frequency (PC1.BASE) also featured as a significant predictor in the regression analysis of suffixation (cf. Section 5.3.2) suggests that the variable has the same effect on the suffixation of the imperative singular forms as it does on the stem vowel; this effect, however, must be interpreted as an indirect one. In all texts contained in the dataset, the imperative singular forms of verbs not belonging to the group of
strong verbs with e/i-gradation were inspected for whether they were suffixed or not, e.g. *laufe vs. lauf ‘run/walk’: on average, 81.5% of them carried a suffix. Thus, when the imperative singular of the strong verbs with e/i-gradation is formed in analogy to other verb paradigms, in the vast majority of cases this will include suffixation, just as the majority of “other” verbs form a suffixed imperative singular. When base verb lemma token frequency is sufficiently high to inhibit analogical levelling, the traditional irregular imperative singular variant will not be suffixed because a suffixed imperative singular variant with the stem vowel i does not exist (e.g. *gibe). Along these lines, the strong link between the stem vowel e and suffixation and between the stem vowel i and a lack of suffixation can be understood.

It was hypothesised further that two different types of persistence have an impact on imperative singular formation of strong verbs with e/i-gradation. The results of the regression analysis of suffixation indicate that speakers indeed develop a routine by which they suffix consecutive imperative singular forms. This practice is extended to include imperative singular forms of strong verbs with e/i-gradation, which traditionally do not occur in a suffixed form. The level “suffixed” of the variable LAST.SECOND.7 captures mainly β-persistence like that in the repeated example 14, i.e. cases in which the suffixation of target imperative singular forms of strong verbs with e/i-gradation in the dataset is triggered by the suffixation of pretarget imperatives which do not belong to the class of strong verbs with e/i-gradation. Only 17 out of 680 data points in the “suffixed” category are observations with a suffixed last pretarget of a strong verb with e/i-gradation and no preceding second last pretarget (cf. Table 19). The effect of pretargets in this “suffixed” condition is significantly different from most conditions with mainly unsuffixed pretargets and without pretargets, but the probability for the target imperative to carry a suffix is still well below 50% (cf. 5.3.4).

(14) Wenn der Reiter nach links reitet, renne bis zu Sandigmann’s Mühle und verstecke dich im Mühlhof. Werte mit Steinen in Richtung Dorf und wenn der Reiter dort hineireitet, rennst du über die Brücke weg. [herr-ringe-gefahrcnten/2350311]


(12) Wenn du dort bist, breche mit dem Dietrich den unordentlichen Schreibtisch auf und nehme den Schlüssel für das Hafenbüro (Werkzeug) aus der Schublade. [clou-2/4268011]

By contrast, the level “suffixed SVe/iG” of LAST.SECOND.7 represents cases of α-persistence like those in (11) and (12), i.e. the suffixed analogical imperative variant of a strong verb with e/i-gradation is repeated or promotes the use of the suffixed analogical vari-
t of another strong verb with e/i-gradation. The effect of \( \alpha \)-persistence in the present study is stronger than that of \( \beta \)-persistence: when imperative singular forms of strong verbs with e/i-gradation in the dataset are preceded by at least one imperative singular of the same or another strong verb with e/i-gradation, the probability for the target imperative to be suffixed rises to 77.2%. An additional regression analysis, in which the verb type of the pretarget was distinguished into “no pretarget”, “other” (not strong verb with e/i-gradation), “SVe/iG” (strong verb with e/i-gradation) and “same SVe/iG” (same strong verb with e/i-gradation), showed that it does not matter in this context whether the pretarget and target are imperative singular forms of the same or different strong verbs with e/i-gradation. Thus, the form *nehme* is equally probable to follow *breche* and *nehme*.

The finding in the regression model for stem vowel choice that the probability for the use of the analogical stem vowel \( e \) in the imperative singular increases when pretargets are suffixed (omitted in Table 20) can be explained by the fact that a suffixed version of the traditional imperative singular variant of strong verbs with e/i-gradation with the stem vowel \( i \) is not attested. Therefore, when speakers wish to attach a suffix to an imperative singular form, they need to adjust its stem vowel to \( e \). The multinomial regression analysis of the data (Chapter 5.3.1) corroborates this interpretation of the effect of persistence: it directly affects the suffixation of the imperative singular forms, and only indirectly their stem vowel.

Finally, it needs to be noted that there is no significant interactive effect of base verb lemma token frequency (PC1.BASE) and persistence (LAST.SECOND.7) on imperative singular formation, but the individual effects of both variables may add up. Thus, the probability for the use of suffixed analogical imperative singular variants is highest when low frequency verbs are preceded by suffixed pretargets which are imperative forms of strong verbs with e/i-gradation, for example *bestehle* ‘steal from (sb.)’ following *gebe*. By contrast, the probability for the use of the traditional (unsuffixed) imperative variant with the stem vowel \( i \) to be used in high frequency verbs like *geben* and *nehmen* is higher when they are preceded by no or only unsuffixed pretargets (in a window of 20 words of left context) than when they are preceded by suffixed imperative singular forms. This additive effect of the variables PC1.BASE and LAST.SECOND.7 also explains the occurrence of the unsuffixed analogical imperative singular variant of strong verbs with e/i-gradation: they are imperative forms of medium to low frequency verbs, and only in a minority of cases (8 out of 43) they are preceded exclusively by suffixed pretargets; thus, their suffixation is not promoted.

The B-spline regression analysis of the temporal development of imperative formation has shown that the use of analogical imperative singular variants of strong verbs with e/i-
gradation increases from 2000 to 2005 and from 2009 to 2013; however, this development is interrupted by a decrease in the probability of the use of analogical variants between the years 2005 and 2009. A burning question raised by this observation is what causes the temporary decline. Part of the reason lies in the data basis, the website spieletipps.de, which interferes with the analysis in a way which the present investigation could not have controlled for: The introduction of spell checking in web browsers can have a strong influence on the way users write their texts and, more importantly, form the imperative singular of strong verbs with e/i-gradation. For example, in the Firefox browser, the most widely used browser in Germany from May 2008 until May 2016 (Papenbrock 2006-2019; StatCounter 1999-2019), inline spell checking was implemented for the first time in version 2.0, released in 2006. Similarly, spell checking was implemented in the Safari browser before 2008, in Opera in 2009, in the Internet Explorer in version 10 (2011/2012) and in Google Chrome in 2012 (cf. browser version histories on Wikipedia (Wikimedia Foundation, Inc. 2019)).24 Since then, spelling errors are marked in every input field with more than one line, including the field for typing a walkthrough on the website spieletipps.de. The traditional imperative singular variants with the stem vowel i and the suffixed analogical variants with the stem vowel e, e.g. gib and gebe, are accepted by such a spell checker (the suffixed analogical variant is identical with the form of the first person singular present indicative, see Table 1). By contrast, unsuffixed analogical imperative singular variants of strong verbs with e/i-gradation like geb are marked as wrong, i.e. they are underlined in red as in text processing programmes like Microsoft Word. Taking into consideration the advice given to walkthrough authors to read through the written text carefully before submitting it25, it is not unlikely that they replace the unsuffixed analogical imperative forms marked as wrong by the traditional variants of the respective verbs with the stem vowel i. If the authors of walkthroughs indeed replace “wrong” imperatives on the basis of the recommendations of the spell checker, they can also be expected to correct genuine spelling errors in the same texts. In order to test this assumption, a sample of 10% of the walkthrough texts in the dataset from the time span between 2005 and 2009 was entered in the original walkthrough submission field on the spieletipps.de website in the Firefox browser, and the number of errors marked by the spell checker in each text was recorded. The mean number of errors in texts per year was found to be highly correlated with the proportion of unsuffixed analogical imperative variants among the three imperative singular variants of

24Third-party spell checking applications, which could be used as a browser add-on, were available all the while (e.g. Red Egg Software 2002-2010; Schwarz 2005)
25 “Schaue dir deinen Text vor dem Absenden noch einmal an und veröffentliche anschließend deinen Beitrag” (http://www.spieletipps.de/cheatcontrib/edit/)
strong verbs with e/i-gradation used in the same years: $r = 0.593$ (p < 0.05). Thus, the decrease in the probability for the use of analogical imperative singular variants with the stem vowel e between 2005 and 2009 can be at least partly explained as an artefact of the introduction of spell checking in web browsers since 2006. The reason for the renewed increase in the probability for the use of analogical imperative singular variants of strong verbs with e/i-gradation after the year 2009 can be understood as a token that spell checking had taken full effect by then (web browsers with inbuilt spell checking had a market share of approximately 70 %, Papenbrock 2006-2019). Thus, by the year 2009, all users of different German web browsers were back on equal terms, and the incipient change in the imperative singular of strong verbs with e/i-gradation could start to progress more naturally again.

The explanation for the decrease in the use of suffixed analogical imperative singular variants can be found in the distribution of persistence contexts between 2005 and 2009. As explained above, previous uses of suffixed imperative singular forms promote the use of suffixed imperative singular variants of strong verbs with e/i-gradation. In the time before 2005 and after 2009, the mean proportion of observations with unsuffixed last pretargets in the dataset is 12.3 %, that of observations with suffixed last pretargets is 42.3 %, and 45.4 % of observations are not preceded by imperative singular forms in the specified window. Between the years 2005 and 2009, the proportions of observations with unsuffixed last pretargets and without pretargets increase to 22.8 % and 54.6 %, respectively, while that for suffixed last pretargets temporarily drops from 42.3 % to 24.7 %. The temporary decline in the use of analogical imperative singular variants of strong verbs with e/i-gradation is thus in part triggered by a sudden increase in the proportion of environments which favour the use of the traditional imperative singular variant with the stem vowel i. The correlation between the proportion of observations with suffixed pretargets and the proportion of analogical imperative variants among all imperative singular variants of strong verbs with e/i-gradation is strong and significant in the time span between 2005 and 2009 ($r = 0.91$, p < 0.05) but not in the years before 2005 and after 2009 ($r = -0.05$, p = 0.92). This low correlation provides further evidence that the increases in the probability of the use of analogical variants before 2005 and after 2009 are genuine signs of a change-in-progress rather than, for example, the side-effects of an uneven distribution of pretargets.

No specific hypotheses were formulated in connection to the “traditional” linguistic predictors like ablaut class, number of syllables, morphological structure and negation. In contrast to Benedikt Szmrecsanyi’s investigation of five features of variation in English (2005; 2006), in which both persistence variables and “conventional intralinguistic variables”
like morphological structure, length in syllables and negation (Szmrecsanyi 2006: 73, 117) were found to have a significant influence, in the present study none of the annotated linguistic variables had a significant impact on either stem vowel choice or suffixation of imperative singular forms of strong verbs with e/i-gradation in the dataset. Thus, all three ablaut classes (3b, 4, 5) and verb structure classes (simplex, prefixed, particle) are affected by change to the same extent, and imperative singular formation is not conditioned phonologically, among others, by the length of the verb (in terms of number of syllables) or morphosyntactically by the negation of the verb in the form of the particle *nicht* or other negators.

In contrast to these intralinguistic factors whose influence was disproved for the complete dataset, only subsets of the data could be annotated for sociolinguistic variables. Thus, although none of the latter variables was found to be significant in the regression models, this does not necessarily mean that they do not have an influence on imperative singular formation in strong verbs with e/i-gradation. An equally likely conclusion is that their influence could not be systematically examined due to a lack of data.

### 5.5 Summary

On the basis of a dataset of 1,939 instances of imperative singular forms of strong verbs with e/i-gradation extracted from a web-based corpus, the present study has shown that, in line with previous investigations of analogical change (among others Hooper 1976; Smith 2001; Bybee 2007), token frequency has a significant conserving effect on imperative singular formation in strong verbs with e/i-gradation: lower frequency verbs succumb to analogical leveling, while higher frequency verbs resist change. However, in this case of German ongoing morphological change, the conserving effect is not exerted by verb lemma token frequency but *base verb* lemma token frequency; thus, a simplex verb like *geben* and particle verbs derived from it, such as *aufgeben*, behave alike. The question whether this conserving effect is a reflection of the mental entrenchment of the competing traditional and analogical imperative variants of a verb in the mental lexicon will be tackled in a reading experiment (Chapter 6). The results of the corpus study are taken into account in the compilation of stimulus sentences for this experiment in that the base verb lemma token frequency of stimulus verbs, not their verb lemma token frequency, is varied.

In addition, the suffixation and indirectly the stem vowel of imperative singular forms recorded in the dataset are affected by the presence and form of imperative singular forms used in the immediate previous context: when no or mainly unsuffixed imperative singular forms occur in a window of 20 words before the target imperative in the dataset, the latter is
very likely to occur in the traditional imperative singular variant with the stem vowel \( i \). When \( \beta \)-persistence (Szmrecsanyi 2005; 2006) applies, i.e. when suffixed imperative singular forms of verbs that do not belong to the class of strong verbs with e/i-gradation precede the target imperative, the probability for the use of the suffixed analogical imperative variant (e.g. \( \text{gebe} \)) increases significantly. The effect of \( \alpha \)-persistence (ibid.) is strongest; that means, when the preceding suffixed imperatives are forms of strong verbs with e/i-gradation, the probability for the suffixation of the target imperative increases to 77.2\%. These findings emphasise that the influence of persistence should be taken into consideration not only in all studies of linguistic variation but also in investigations of (ongoing) language change. Benedikt Szmrecsanyi explains that production priming is a likely cause of such persistence patterns (2006: 3); it would thus be interesting to test whether the persistence patterns identified in the present study translate into priming effects when imperative singular forms are processed by speakers. As the experimental study (Chapter 6) focuses on testing the relation between frequency and entrenchment rather than that between persistence and priming, however, the potential priming effect of the form of pretarget imperatives was instead controlled in the experiment.

Over the largest part of the relatively short time span covered by the corpus basis (2000 to 2013), a slight increase in the use of analogical imperative singular variants of strong verbs with e/i-gradation was observed; thus, this case of analogical levelling seems to have moved past the stage of variation to a change-in-progress from the traditional imperative singular variant with the stem vowel \( i \) to analogical variants with the stem vowel \( e \).

Due to a lack of available sociolinguistic meta information, the respective variables did not feature significantly in regression models of subsets of the data with the necessary documentation. Judging from these regression analyses, a speaker’s gender does not seem to influence stem vowel choice or suffixation of the imperative singular. Female and male participants in the experimental test might, however, react differently to the different imperative variants. They should therefore be evenly represented, and the data obtained from the experiment will be annotated for the variable participant gender in order to assess the influence of the variable more accurately than was possible in the present corpus study. Similarly, the geographical background of authors turned out not to have a significant effect on imperative singular formation in the corpus study, mainly due to a lack of data and the high proportion of standard imperative singular variants used by authors from different dialect regions. Therefore, the experimental study involves participant groups from two regions in Germany.
6 Experimental Study

The present chapter aims to test the explanation of the conserving effect of high frequency in analogical levelling on the basis of their cognitive entrenchment. In the Walkthrough Corpus study (Chapter 5), such a conserving effect was discovered in the change-in-progress affecting the formation of the imperative singular of German strong verbs with e/i-gradation. The generally accepted explanation of this effect on the grounds that high frequency linguistic items are so strongly entrenched in speakers’ minds that they resist levelling cannot, however, be tested in a corpus study alone. Therefore, the competing traditional and analogical imperative variants were presented in verbs of different token frequency to participants in a self-paced reading with recall experiment. In this design, the entrenchment of the stimulus imperatives is measured in terms of reading times and recall accuracies.

The following Section 6.1 will outline the basic assumptions of the present study, i.e. the logic behind measuring entrenchment of competing forms in terms of reaction times and recall accuracies, the general idea of an apparent-time model and the specific method used in the current experiment, along with hypotheses and expected results. In Section 6.2, the experiment design and procedure are introduced, including the cover story, instructions and software used and an explanation of how data and metadata were obtained in the actual test and through the use of a background questionnaire. Section 6.3 explains how the assumptions and hypotheses were translated into variables, i.e. experimental stimulus conditions and participant groups (6.3.1 and 6.3.2). It also summarises how the reading-time and recall-accuracy data were annotated for predictor and random variables and meta information from the background questionnaire (6.3.3 and 6.3.4). Section 6.3.5 describes how the dataset obtained from the experiment was separated into subsets of data for participants who were deceived by a cover story, which served to distract participants from the aim of study, and participants who detected the cover story. The results of the analyses of the subsets of data are presented and discussed in Sections 6.4 and 6.5 for these two groups of participants. Sections 6.6 and 6.7 conclude with a general discussion and summary of the experiment results and their contribution to usage-based accounts of frequency effects on entrenchment.

6.1 Assumptions and hypotheses

The present experiment, its design and methodology and the choice of participants, relies on a number of assumptions and hypotheses. These are based on the results of the Walkthrough

26 Parts of the present chapter have been published in *Cognitive Linguistics* (Krause-Lerche 2019b).
Corpus study and findings from previous work by other researchers which will be outlined in the following sections.

6.1.1 **Measuring the entrenchment of competing forms in analogical levelling in terms of processing latencies**

**THEORETICAL ASSUMPTIONS**

As outlined in Section 1.4, the reason which is generally provided in the usage-based literature to account for the conserving effect of high token frequency in analogical change is the *lexical strength of words* (Bybee 1985: 117) or *entrenchment of linguistic units* (Langacker 1987: 59). While these general descriptions are arguably ambiguous between referring to the lemma level or the level of individual forms in a paradigm or individual variants of a linguistic variable, explanations of frequency patterns in analogical change are almost unanimously referring to the entrenchment or strength of lexical representation of linguistic *forms*:

“the more a *form* [emphasis added] is used, the more its representation is strengthened” (Bybee & Thompson 1997: 380);

“the notion of lexical strength … accounts for the maintenance of irregularity and suppletion in high-frequency *forms*. Conversely, the proposal that infrequently-used *forms* fade accounts for the tendency to regularize infrequent irregular *forms*, for an irregular *form* that is not sufficiently reinforced will be replaced by a regular formation” (Bybee 1985: 119; emphasis added throughout);

“High token frequency of *specific items* [emphasis added], especially irregular ones such as *went, told, or spoke*, has a conserving effect on their morphological form” (Schmid 2016b: 15).

These statements imply that irregularity is concentrated in high frequency lemmata not because of the strong entrenchment of the respective lemma, but because of the strong entrenchment of the irregular forms themselves, induced by their high frequency of occurrence. To quote Bybee’s famous example (Hooper 1976: 99-100; Bybee & Thompson 1997: 380 inter alia), the reason that *slept* is retained in its irregular form can be argued to be caused by the fact that this particular form is highly entrenched in speaker’s minds, whereas the analogically formed variant *slepted* is weakly entrenched. As infrequent irregular forms like *wept* are
typically affected first by analogical levelling because they are weakly entrenched, the regularised forms of the same lemmata like *weeped* should increase in frequency; they should therefore be entrenched to an increasing degree in speakers’ minds (Bybee & Thomson 1997; Langacker 1987), to the point that they are more frequent and, therefore, more strongly entrenched than their irregular counterparts. Thus, when a change has progressed sufficiently, among higher frequency lemmata the conservative or traditional irregular forms will have a higher frequency than their analogical or regularised variants, whereas the opposite is true for lower frequency lemmata, where the analogical variants will be more frequent.

**Figure 20:** Sketch of processing latencies reflecting frequency and entrenchment of innovative and conservative variants of the same lemma competing during analogical leveling (Krause-Lerche 2019b: 6)

Left panel: competing variants are entrenched to different degrees; middle panel: competing variants are entrenched to same degree; right panel: different entrenchment of competing variants in stable variation

In order to test whether entrenchment is the reason for the resistance of frequent irregular forms to change, it is not enough to monitor speaker’s reactions to irregular forms only. Speakers are likely to process irregular forms of high frequency lemmata (e.g. *slept*) faster than those of low frequency lemmata (e.g. *wept*) because they can access the former faster in their mental lexicon than the latter; however, this may simply be a reflection of entrenchment on the lemma level. If innovative and conservative variants of lemmata undergoing analogical levelling are indeed entrenched to a variable extent, an interaction between frequency and irregularity should be observable when the competing forms of higher and lower frequency lemmata are presented to speakers, as illustrated in the left panel of Figure 20. If, by contrast, the entrenchment of the competing variants of the same lemma only de-
pended on the frequency of the respective lemma, both variants should be processed equally fast (middle panel of Figure 20). A third pattern could be observed in stable variation where the traditional or conservative forms are more frequent, and thus more strongly entrenched, than the innovative analogical forms across the lemma token frequency range; this situation is outlined in the right panel of Figure 20. Note that two concepts of frequency are contrasted here and in the remainder of the study which may not be confused: i) the token frequency of the lemma which is affected by analogical levelling (see x-axis in Figure 20) and ii) the absolute frequency of occurrence of one variant/form in a particular lemma token frequency region (as reflected by vertical location with respect to the y-axis).

A second division of frequency and entrenchment becomes relevant in this context which was propagated in cognitive linguistics, among others, by Schmid (2010). He elaborates on a division of types of frequency suggested by Hoffmann (2004), on the basis of which he introduces a division of entrenchment into cotextual and cotext-free entrenchment. He acknowledges that the frequency of any linguistic item, its absolute frequency, must be kept apart from its frequency in a given construction, its relative frequency. These two types of frequency reflect two different types of entrenchment in speakers’ minds:

For example, high-frequency nouns like time, point or way are most likely more entrenched than less frequent ones (like disinclination or unwillingness), irrespective of their actual linguistic environment. ... [T]his means that there is after all an absolute, cotext-free type of entrenchment, which correlates with absolute frequency of occurrence. ... If absolute frequency translates into cognitive system as cotext-free entrenchment, it seems reasonable to think of its relative counterpart as reflecting cotextual entrenchment. ... [C]otextual entrenchment can be seen as the tendency of one linguistic element or unit to trigger the (co-)activation of one or more other linguistic units or structures in language users’ minds, if the former significantly co-occurs with the latter in actual discourse (Schmid 2010: 120)

The aim of the present study is to test whether the different absolute frequencies of occurrence of innovative and conservative variants of the same lemma result in different degrees of cotext-free entrenchment, in Schmid’s terms. In order to measure this type of entrenchment and prevent interference from cotextual entrenchment of the competing forms due to their higher or lower relative frequency in a given context, reactions to the two variants should be
contrasted in the same context. Returning to the example above, when forms like slept, slepted, wept, and weeped are presented to speakers in the same context, such as She slept/slepted/wept/weeped all night, processing differences can yield evidence for whether innovative and conservative forms of the same lemma are entrenched in speakers’ minds to the same extent because their entrenchment only depends on the frequency of the lemma, as illustrated in the middle panel of Figure 20, or whether different forms of the same lemma are entrenched to a variable extent, as outlined in the left panel, because of their different frequency of occurrence. The following section will translate this theoretical construct into hypotheses and expected results with regard to speakers’ reactions to the traditional and analogical imperative variants of strong verbs with e/i-gradation in the present experiment.

EXPECTED READING-TIME RESULTS FOR STRONG VERBS WITH E/I-GRADATION

Figure 21 replicates the effect of base verb lemma token frequency (PC1.BASE) on the probability of imperative singular forms to occur with the analogical stem vowel e in the Walk-through Corpus data (cf. Section 5.3.3). Traditional imperative singular variants with the stem vowel i of high frequency verbs, like nimm, have a high frequency and probability of occurrence, indicated by the proximity of this label to the regression line in Figure 21. By contrast, the traditional imperative singular variants of low frequency verbs like zertritt have a low frequency and probability of occurrence, as illustrated by the distance of the respective label from the regression line. The same two concepts of frequency as in Figure 20 are contrasted here: i) the token frequency of the base verb lemma whose imperative occurs in the corpus (on the x-axis) and ii) the absolute frequency of occurrence of one imperative singular variant/form in a particular verb token frequency region (indicated by the regression line).

The widely-accepted explanation of the conserving effect of frequency on the basis of entrenchment holds that irregular high frequency forms are strongly entrenched in speakers’ minds, “which makes them resistant to paradigmatic analogical pressure and change” (Schmid 2016: 15; cf. Bybee 1985). The traditional irregular imperative forms of strong verbs with e/i-gradation with the stem vowel i can thus be expected to be strongly entrenched in higher frequency verbs but weakly entrenched in lower frequency verbs. When presented in the same sentence (16), nimm will thus be read faster than zertritt.

(16) Nimm/Zertritt alle Gegenstände in deiner Nähe (mit)!
   ‘Pick up/ Tread down all items near you.’
Figure 21: Frequency and probability of occurrence of imperative singular variants with the stem vowel \(i\) dependent on base verb lemma token frequency (Krause-Lerche 2019b: 9)

However, measuring processing times for the two verb forms in sentence (16) is only part of the picture: If forms like \(nimm\) are read faster (regardless of word length), this only replicates findings from previous studies that higher frequency lemmata are generally recognised faster than lower frequency lemmata (e.g. Cortese & Balota 2012: 164). Hence, in order to genuinely test the entrenchment explanation of the conserving effect of frequency in analogical levelling, the stimulus material needs to be extended. As Figure 22 illustrates, the frequency and thus probability of occurrence of the analogical imperative singular variants of strong verbs with \(e/i\)-gradation (with the stem vowel \(e\)) similarly depends on the frequency of the base verb: they are higher in verbs with a lower than with a higher token frequency. For example, the forms \(zertret\) and \(zertrete\) have a high absolute frequency and probability of occurrence, whereas \(nehm\) and \(nehme\), the analogical imperative singular variants of a high frequency verb, have a low frequency and probability of occurrence (indicated by the distance of labels from the regression line in Figure 22). Accordingly, the analogical variants of higher frequency verbs are assumed to be weakly entrenched, which can be expected to lead to delayed access to these forms in the mental lexicon, as reflected by slower processing, in comparison to analogical variants of low frequency verbs which, owing to their high frequency and probability of occurrence, should be processed faster by speakers of German and, hence, be more strongly entrenched. Thus, in addition to the imperative singular forms used in sentence (16), analogical imperative singular variants of the same verbs with the stem vowel \(e\) were added to the set of stimuli in the present experiment, as illustrated in example (17), in
order to test whether corpus frequencies are really a reflection of the mental entrenchment of forms.

(17) *Nehm(e)/Zertret(e) alle Gegenstände in deiner Nähe (mit)!*

Entrenchment in the present study was measured in terms of processing, or more precisely, reading times of stimuli. The analogical imperative singular variants of high frequency verbs (e.g. *nehm(e)*) as well as the traditional irregular variants of lower frequency verbs (e.g. *zertritt*) were found to have a low frequency and probability of occurrence in the Walkthrough Corpus study; they should thus be weakly entrenched in speakers’ minds and in turn require relatively high reading times. By contrast, traditional imperative singular variants of higher frequency strong verbs with e/i-gradation with the stem vowel *i* (e.g. *nimm*) have a high frequency and probability of occurrence which is assumed to reflect their strong mental entrenchment; these forms should thus be read relatively fast. Finally, on the basis of their frequency of occurrence in the corpus, analogical imperative singular variants of low frequency verbs like *zertret(e)* can be expected to be more strongly entrenched than the traditional irregular variants of the same verbs like *zertritt* and the analogical variants of high frequency verbs like *nehm(e)*. However, they can be expected to be processed slower than the traditional imperative variants of high frequency verbs (e.g. *nimm*) because the latter have been in use for centuries. Figure 23 captures these predictions, taking into account that higher frequency verbs are in general processed more quickly than lower frequency verbs.
Thus, it is assumed that the conservation of traditional irregular imperative variants of high frequency strong verbs with e/i-gradation like *nimm* and *gib* is caused by the high frequency of occurrence of these individual forms and their resulting strong entrenchment in speaker’s minds, not by the frequency and entrenchment of the respective verb lemmata. If this assumption is wrong, both imperative variants of each verb (e.g. *nimm* and *nehm*) should be processed equally fast, corresponding to the same degree of entrenchment (cf. left and middle panel of Figure 20).

### 6.1.2 Introducing an off-line measure of entrenchment: recall accuracy

**THEORETICAL ASSUMPTIONS**

In addition to testing how speakers process the imperative singular variants of strong verbs with e/i-gradation by measuring reaction times to stimulus imperatives, the present study introduces an off-line measurement of entrenchment. This idea is based on an observation by Joan Bybee (Hooper) in her seminal 1976 paper:

For infrequent paradigms adults may not be exactly sure of all the forms, or the mere infrequency of a suppletive paradigm makes an analogical formation more acceptable. For instance, as I mentioned earlier, *creeped* is not standard, but I would not flinch if I heard it, and I might even produce it myself, although I know *crept* is “correct.” However, *keeped* would definitely cause a negative reaction, because the form *kept* is so solidly established, due to its frequency. (Hooper 1976: 101)
She assumes that the traditional irregular variants of high frequency verbs, e.g. *kept*, are so strongly entrenched in speakers’ minds that analogical variants of the same verbs are quickly rejected by speakers as ineligible realisations, for example of the past tense of the verb *keep*. In contrast, irregular variants of low frequency verbs are weakly entrenched; when speakers encounter analogical variants of these verbs, it is not unlikely that these variants go unnoticed or that speakers even reproduce them, for example the analogical past tense form *creeped*. Bybee’s intuition is translated into a measure of entrenchment in the present experiment in the form of a recall task. In contrast to “classical” self-paced reading experiments in which participants are presented with stimulus sentences and comprehension questions in order to ensure that they pay attention to the content of the sentences, after having read a stimulus sentence, participants in the present experiment were asked to recall it. Assuming that the frequency of linguistic items is correlated with their mental entrenchment, the finding that a stimulus is repeated in the form in which it was presented can be treated as a sign that the presented stimulus form is to some extent entrenched in the mental lexicon of the speaker. If, however, a stimulus is not repeated, for example if the form *keeped* is replaced by the form *kept*, this would show that the presented form is weakly entrenched in the speaker’s mental lexicon. Thus, as Figure 24 illustrates, an interaction between irregularity and frequency is expected to apply to recall accuracies, as hypothesised for reading times. However, different outcomes of the experiment are conceivable with respect to recall accuracies as well, depending on whether the object of study is a case of ongoing analogical change from a traditional irregular formation to an innovative analogical formation or an instance of stable variation between the competing variants. Note that Figure 24 contains inverted y-axes for frequency/entrenchment and, therefore, recall accuracies for easier comparison between hypotheses related to reading times and recall accuracies (cf. Figure 20).

If the assumption is true that the absolute frequencies of the innovative and conservative variants competing during analogical levelling are a reflection of entrenchment on the level of the individual form or variant, and irregularity is retained in high frequency lemmata because of the high entrenchment of these irregular high frequency *forms*, then irregular variants of high frequency lemmata (like *kept*) should be replaced in no or only very few cases when they are presented to speakers, as illustrated in the left panel of Figure 24. However, the innovative analogical variants of the same high frequency verbs (*keeped*) should be replaced by their conservative irregular counterparts because of their low frequency and entrenchment. Since lower frequency lemmata occur in analogical variants more often than in traditional variants, analogical forms like *creeped* can be expected to be replaced less often than analogi-
cal variants of high frequency lemmata, whereas irregular forms of low frequency lemmata like crept should be replaced more often than the irregular variants of high frequency lemmata.

If, in contrast, the conservation of irregular formations only depended on the frequency of the particular lemma, there should be no difference in recall accuracy between the traditional and analogical variants of one and the same verb. Since under this assumption high frequency lemmata are strongly entrenched in speakers’ mental lexicons irrespective of the form in which they occur, both competing variants (e.g. kept and keeped) should be recalled and repeated to a high degree. On the basis of their lower frequency and entrenchment, speakers can be expected to be more insecure about lower frequency lemmata; therefore, they will replace a higher number of the presented traditional variants of these lemmata by their analogical counterparts (e.g. crept by creeped), and vice versa. In other words, as displayed in the middle panel of Figure 24, there should only be a main effect of lemma frequency on recall, and little to no effect of variant type.

Figure 24: Sketch of replacement rates reflecting frequency and entrenchment of innovative and conservative variants of the same lemma competing during analogical leveling
Left panel: competing variants are entrenched to different degrees; middle panel: competing variants are entrenched to same degree; right panel: different entrenchment of competing variants in stable variation

Finally, if the case of analogical levelling under investigation has not yet moved past a stage of stable variation between traditional and analogical variants, the frequency of the lemma presented to speakers should not matter, and only the variant type can be expected to have an effect on recall accuracies. Since in this scenario (illustrated in the right panel of Fig-
ure 24) traditional variants (like kept and crept) are generally more frequent and therefore more strongly entrenched than analogical variants, they can be expected to be recalled in the vast majority of cases, irrespective of the frequency of the lemma in which they are presented, while the analogical variants (such as kepeed and crepeed) will be replaced by traditional variants to a relatively high degree among higher and lower frequency lemmata.

EXPECTED RESULTS IN TERMS OF RECALL ACCURACIES OF STRONG VERBS WITH E/I-GRADATION

As hypothesised for the dependent variable reading time, recall accuracies of imperative singular forms of strong verbs with e/i-gradation in the present experiment are expected to directly reflect the frequency and thus entrenchment of the individual forms or variants; thus, the frequency of the verb lemma presented to speakers and the variant type in which they are presented (traditional irregular vs. innovative analogical) should exhibit an interactive effect on recall. As explained above, the results of the Walkthough Corpus study have shown that the traditional irregular imperative singular variants of high frequency strong verbs with e/i-gradation, like nimm, have a high frequency and probability of occurrence (see Figure 22). These forms are assumed to be strongly entrenched in the mental lexicons of native speakers of German, which means that participants in the present experiment can be expected to repeat them.

Figure 25: Expected results for dependent variable recall accuracy

Analogical imperative singular variants of high frequency strong verbs with e/i-gradation like nehm(e), in contrast, have a low frequency of occurrence. If this reflects the weak entrenchment of these forms in speakers’ mental lexicons, then the analogical stem vowel e in such
stimulus imperatives is likely to be replaced by the traditional irregular stem vowel *i*. More or less the opposite pattern can be assumed for low frequency strong verbs with *e/i*-gradation: their imperative singular occurs more frequently with the analogical stem vowel *e* than in the traditional irregular formation with the stem vowel *i*. Thus, the analogical imperative variants of low frequency verbs (e.g. *zertret(e)*) ought to be repeated as presented in the stimulus sentences, reflecting strong entrenchment, whereas traditional imperative singular variants of low frequency verbs with the stem vowel *i* (e.g. *zertritt*) are not unlikely to be replaced by analogical variants with the stem vowel *e*.

Thus, recall accuracies are assumed to be different for the competing traditional and analogical imperative variants of strong verbs with *e/i*-gradation, which would lend further weight to the explanation of the conserving effect of high token frequency on the basis of the entrenchment of individual forms/variants instead of on the lemma level.

6.1.3 Applying the apparent-time method in a study of language processing

Variation between the traditional imperative singular variant of strong verbs with *e/i*-gradation and an analogically formed variant with the stem vowel *e* in non-standard, vernacular German has been commented at least since the 1970s, for example in the Duden Grammar (Grebe 1973; cf. 1.2.1). The high probability for the analogical variants to occur not only in the lowest frequency base verbs which has been found in the Walkthrough Corpus study (Chapter 5) similarly suggests that the ongoing change in the formation of the imperative singular of these verbs has developed for much longer than the time span covered by the corpus (2000-2013). An apparent-time method should, therefore, be implemented in the present experiment in order to confirm the classification of the phenomenon under investigation as a change-in-progress across a longer stretch of time than was possible in the corpus study.

REAL TIME, APPARENT TIME AND AGE GRADING

The apparent-time methodology was developed in sociolinguistics, where it is contrasted with real-time studies of language change (cf. Meyerhoff 2011, chapter 7; Blondeau 2013). In *real-time* investigations, the use of a certain linguistic feature or variant is observed in a speech community at two or more different points in time. These longitudinal studies are diachronic or historical linguistic studies in a narrower sense, which require a close matching of the conditions of the original study in terms of the speech community and the sample of speakers drawn from it when the study is replicated after a certain number of years, decades or centuries. This becomes particularly complicated when, instead of matching a new sample of speakers on the composition of an earlier sample in so-called *trend studies* (Trudgill 1988;
Pope et al. 2007), (a part of) the same sample of speakers is observed at both times in a panel study (Sankoff & Blondeau 2007; Buchstaller et al. 2017), perhaps by the same researcher. This type of investigation would not have been possible in the present case: the Walkthrough Corpus was the only source with a sufficient number of tokens of imperative singular forms of strong verbs with e/i-gradation to determine the stage of the change-in-progress in the first place, and no older corpora exist, for example from the 1970s or 1980s, which would yield enough data to make it possible to monitor the change (see 3.1.2.1). Several sociolinguists have run into similar problems of data scarcity, time constraints and the like. As a result, they constructed the apparent-time hypothesis, which assumes that once speakers are past a certain age or developmental stage, i.e. past adolescence (Chambers 2003), their early teens (Lenneberg 1967) or “after the critical period, which delimits the formative period of language acquisition, individual speakers’ linguistic behavior changes minimally over the course of the life span, and individual vernaculars remain essentially stable” (Blondeau 2013: 501). Thus, a potential language change can be assessed “by comparing the speech of speakers of different ages within a community at a single point in time” (Meyerhoff 2011: 140). The first examples of such synchronic studies of language change were William Labov’s investigations on Martha’s Vineyard and in New York City (1965; 1966); later apparent-time studies include Tagliamonte’s investigations of non-standard was and were in existential sentences in Yorkshire (1998) and work by Cameron on the use of verbs of quotation in Puerto Rican Spanish (1998). They have shown that in change-in-progress, “typically, a variant in the parents’ speech will occur in the speech of their children with greater frequency, and in the speech of their grandchildren with even greater frequency” (Chambers 2003: 185).

Even though real-time tests have affirmed the usefulness of the apparent-time method (Trudgill 1988; Sankoff & Blondeau 2007), a number of studies have been criticised because they claim to have discovered an ongoing change in the use of a certain linguistic feature or variant which other researchers interpret as an instance of age-graded variation. This term refers to the fact that, while the language of an individual remains relatively stable across their lifespan, changes may be observed in particular when they enter and leave the so-called linguistic market (Bourdieu & Boltanski 1975). Sankoff and Laberge, who introduced the latter term in sociolinguistics, “noticed that there is often a peak in use of the standard variant in people as they reach their early twenties, and then a subsequent decline in the frequency of that same variant among speakers in later middle age” (1978b; Sankoff 2004; qtd. in Meyerhoff 2011: 154). Thus, the use of a standard variant in these cases may be explained as a consequence of “marketplace pressures”.
APPARENT TIME IN THE PRESENT EXPERIMENT

The apparent-time method is taken up in the present experiment and applied to language processing: it will be assumed that the reactions of speakers from two different generations towards the competing imperative singular variants of strong verbs with e/i-gradation reflect their experience with different stages of the ongoing change in imperative formation of these verbs. More precisely, older speakers are expected to react more conservatively towards the analogically formed variants with the stem vowel e than younger speakers who have had more contact with them. The two generations of participants were chosen in such a way that the potential disturbing effect of age-grading on the results could be minimised. The younger participants should have entered the linguistic marketplace, i.e. be between 18 and 30 years of age (born 1985-1997), while the older participants should still be part of it at 45 to 60 years of age (born 1955-70), with a distance of 20-30 years between them. In this way, both generations are subject to the same or similar degrees of “marketplace pressures” and thus affected by the same or similar concerns for language correctness, also known as social desirability effects.

APPARENT TIME IN THE WIDER SENSE

In addition to covering a relatively short time span, unfortunately, the Walkthrough Corpus contained little sociolinguistic meta information about the authors of texts. This made it hard to impossible to determine dialectal influences on the variation in imperative singular formation of the strong verbs with e/i-gradation which would be expected on the basis of dialectological research (cf. Section 1.4.2). Therefore, in addition to two age groups of participants from one dialect region, an additional group should be tested from a second dialect region. Not only does this method yield information of whether dialect differences with regard to imperative singular formation described in the literature are reflected in the reactions of speakers with a different regional background. In similar “apparent-time studies in the wider sense” (Mair 2008: 1118), dialects have repeatedly been used as “a window on the past”: Under this heading, Sali Tagliamonte explains that “contrastive, inter-community patterns provide evidence for the progression of linguistic change, i.e. how a change is taking place, or pinpoint the stage of development represented by the variety in question” (2012: 186). It is not unlikely that analogical levelling in the imperative singular of strong verbs with e/i-gradation is not progressing uniformly across Germany; participants from different regions can thus be expected to react differently to the competing imperative singular variants of these verbs.
For logistic reasons, the experiment was run in Freiburg im Breisgau in the very southwest of Germany and in Leipzig, in the middle east, with Freiburg representing the Low Alemannic dialect region and Leipzig the Upper Saxon dialect area (cf. Figure 1). Both dialects generally form the imperative singular of strong verbs with e/i-gradation with the stem vowel i, e.g. geben > gib! (cf. 1.4.2), but Harald Noth implicitly explains that for lower frequency verbs, the analogical imperative singular with the stem vowel e is used in the Kaiserstuhl region and its surroundings (1993: 316-317). Thus, if differences are found in the reactions of the dialect groups, the participants from Leipzig are expected to react more conservatively towards the analogical imperative forms with the stem vowel e presented to them in the experiment than the participants from Freiburg, who are better accustomed to the forms, at least in lower frequency verbs. Apart from the fact that such analogical formations of the imperative singular of strong verbs with e/i-gradation are not attested for Upper Saxon in the literature, this hypothesis is reinforced by the fact that the Upper Saxon “dialect” is in fact a “supradialectal colloquial variety [which is today] characterised by the loss of almost all the most localizable dialect features” (Bergmann 2013: 309), while Low Alemannic is “still relatively alive and well preserved and is used in the domain of the family, among friends and even at the work-place” and the vernacular in this region is “characterized by numerous dialect interferences at the phonological, morpho-syntactical and lexical level” (Philipp & Bothorel-Witz 2013: 313). For reasons of time, limited funding and increased difficulties in finding older participants for the experiment, it would not have been possible to run the experiment with younger and older participants from Leipzig. Participants from this region will therefore match the age range of the younger group from Freiburg only (18-30 years of age, born 1985-1997); if the younger participants from Leipzig are found to behave more conservatively towards the analogical imperative singular variants than those from Freiburg, older speakers from the same dialect region can be expected to react even more conservatively. Thus, the benefits of testing such a fourth group would not necessarily have outweighed the costs of doing so. For similar reasons, the experiment is not run in any other major dialect region in Germany.

6.2 Method

In the self-paced reading with recall design resulting from the operationalisation of the theoretical assumptions outlined above, participants are presented with stimulus sentences on a computer screen in centred self-paced reading, i.e. they see one word of the sentence at a time in the middle of the screen. After each sentence, they are asked to repeat the sentence which
they have just read. This oral repetition is recorded with the help of a headset. After the repetition, the next stimulus sentence is displayed, it is repeated, and so on.

The experiment was created on the basis of the software application *Linger* (version 2.94), “a Tcl/Tk application that performs a variety of language experiments, particularly self-paced reading”, developed by Douglas Rohde (2001-2003). The software comes with a few preprogrammed experiment types, such as “Masked Self-Paced Reading/Listening”. The software scripts for two of these experiment types were modified in order to create a new experiment type, viz. the required self-paced reading with recall. To this end, parts of the scripts for “Centered Self-Paced Reading/Listening” and “Speak and Listen” were combined and a few additional settings were changed.

In the experiment, participants are first provided with an instruction which informs them about the procedure of the test. The exact instructions for the experiment can be found in Appendix E; they include the following main points:

1. Participants will read sentences from a video game instruction.
2. These sentences are presented either in the original order of the text they are taken from, or in a random order. The order may influence participants’ responses.
3. Participants use the space bar on the keyboard to proceed to the next screen.
4. Sentences are displayed word by word in the middle of the screen.
5. After each sentence, participants are asked to repeat the sentence they have just read.
6. Breaks are integrated at an interval of 5 stimulus sentences.
7. Participants will see some practice sentences in order to get used to the procedure.

After the instructions, seven practice sentences are presented, two very general ones like “Dieser Satz ist ein kurzer und leichter Übungssatz” ‘This sentence is a short and easy practice sentence’, and five sentences which are similar to the actual test sentences. This procedure was inspired by the example instructions provided by Douglas Rohde.

8. Participants can ask questions after the instructions or after the practice part.

If participants did not have any questions after the practice sentences, they could proceed to the actual test, in which, however, they were first shown another three practice sentences in
order to avoid any possible warm-up effects on the reading and recall of the target sentences. They were then presented an equal number of alternating target and filler sentences. The order of presentation of items is randomised by the Linger software. Which of the conditions of the target sentences is presented to which participant is determined in a Latin square (cf. Section 6.3.1), likewise implemented in the software. Each participant received a participant number which serves two purposes: on the one hand, this number decides which of the lists in the Latin square is allotted to which participant; on the other hand, this method ensures anonymisation of the data, as all data obtained from the testing of a particular participant is stored under this participant number. After the end of the test was indicated by a message on the screen, participants filled out a background questionnaire about metadata and the procedure of the test. They were debriefed about the cover story, hereby making sure whether they had detected the aim of the study. Finally, they signed a statement agreeing that the data obtained in the test may be used for the purposes of the present investigation. All original materials can be found in Appendix E.

A control group was introduced which performed a self-paced reading task of the same stimuli; this group, however, did not repeat the sentences but answered comprehension questions about the stimuli. In this way, the impact of the recall task on reading times could be compared to a baseline (see 6.3.2 for details of the control task and the control group and 6.4.1 for the impact of the recall task on reading times).

Each participant was tested in a separate session. The experimenter was present until the participant started the test on the computer, and she returned when they filled out the background questionnaire. The information about the cover story and the proper aim of the study were stated orally by the experimenter. During the actual test, the experimenter stayed in an adjacent room.

6.3 Data

6.3.1 Stimuli

Participants were presented 10 practice sentences, followed by 30 target sentences, i.e. five times the number of conditions (see below), and 30 filler sentences in the main part of the test. As explained in Section 6.1, the present study measured context-free entrenchment by varying the frequency of stimulus imperative forms while holding the context in which they occur constant. To this end, target stimulus sentences were constructed which contained a slot for an imperative singular form. As illustrated in Table 23 below, the frequency and probability constellations found in the Walkthrough Corpus study (Chapter 5) were captured by filling this slot with differ-
ent imperative singular variants of strong verbs with e/i-gradation with a higher or lower base verb lemma token frequency. The length of the stimulus sentences was controlled, at approximately 12 words on average (12.32 for all stimulus sentences, 12.41 for target sentences and 12.23 for filler sentences). This should make it possible but not too easy to repeat the sentences, so that participants would subconsciously process them.

CONDITIONS OF TARGET SENTENCES
Each target sentence contained a slot for an imperative singular form. The specific form presented to a participant belonged to one of the six conditions following from the multiplication of the factor levels of the two main predictors in the experiment design imperative form (IMP.FORM) and base verb lemma token frequency (BVLTF, cf. 1.3 and 6.1).

Imperative singular forms in the target sentences occurred in three forms. The first two forms were the unsuffixed imperative singular variants of a strong verb with e/i-gradation, e.g. *gib* and *geb* ‘give’ (cf. conditions a, b, d, and e in Table 22). A control condition was introduced in the form of the imperative singular of a verb from a verb class other than the strong verbs with e/i-gradation, e.g. *mach* ‘do’ (conditions c and f). Thus, the variable imperative form (IMP.FORM) has three levels: “stem vowel i”, “stem vowel e”, “control verb”.

The base verb lemma token frequency (BVLTF) of the verbs whose imperative singular forms occurred in the target sentences was varied as well; that means, participants saw forms of verbs with higher or lower token frequency. Values for the variable were taken from the DeReWo frequency list offered by the Institute for the German language (Insitut für deutsche Sprache 2012b; cf. Chapter 3.1), however, only for simplex and prefixed verbs. If a particle verb should occur in a target sentence, the value of the corresponding simplex verb was referred to. Verbs in the target stimuli for the experiment were taken from two frequency bins, i.e. higher vs. lower frequency. The bins were determined by the median of the base verb lemma token frequency class of items in a list of all strong verbs with e/i-gradation which could be used in the stimulus sentences (see Appendix A). High frequent verbs should stem from a verb token frequency class above the median value of 12, low frequent verbs from below the median (frequency classes in the DeReWo list are ordered decreasingly from 0 “highest frequency” to 29 “lowest frequency”). In addition, the distance between the high frequency verb and the low frequency verb should be at least three frequency classes in the DeReWo list, e.g. a higher frequency verb from class 10 was contrasted with a lower frequency verb from class 14, so that verbs from a frequency class just above the median would not be compared to verbs just below the median. Control verbs are taken from the same fre-
quency class as the corresponding strong verb with e/i-gradation (with minor deviations for the sake of plausibility, see below).

The multiplication of the levels of these independent variables yields the number of conditions which the target sentences need to fulfil. In the present $2 \times 3$ design, each sentence is realised in 6 different versions:

Table 22: Conditions of target sentences in the main experiment

<table>
<thead>
<tr>
<th>Imperative singular form</th>
<th>Base verb lemma token frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>high</td>
</tr>
<tr>
<td>unsuffixed imperative with stem vowel $i$, strong verb with e/i-gradation</td>
<td>a</td>
</tr>
<tr>
<td>unsuffixed imperative with stem vowel $e$, strong verb with e/i-gradation</td>
<td>b</td>
</tr>
<tr>
<td>unsuffixed/suffixed imperative, control verb</td>
<td>c</td>
</tr>
</tbody>
</table>

All original stimulus sentences used in the experiment are found in Appendix E (conditions are indicated by the lower case letters used in Table 22). The realisations of the imperative slot in target sentence 11 serve as an illustration of the conditions:

Table 23: Realisation of stimulus sentence conditions in target sentence 11

*Auch wenn du nicht weiter kommst,*

‘Even if you are stuck, do not ...’

<table>
<thead>
<tr>
<th>Imperative singular form</th>
<th>Base verb lemma token frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>high</td>
</tr>
<tr>
<td>unsuffixed imperative with stem vowel $i$, strong verb with e/i-gradation</td>
<td><em>gib (auf)</em> ‘give up’</td>
</tr>
<tr>
<td>unsuffixed imperative with stem vowel $e$, strong verb with e/i-gradation</td>
<td><em>geb (auf)</em> ‘give up’</td>
</tr>
<tr>
<td>unsuffixed imperative, control verb</td>
<td><em>lass (sein)</em> ‘break off’</td>
</tr>
</tbody>
</table>

*die Aktion auf gar keinen Fall (...).*

the move/campaign on any account.’

As can be seen in Table 23, imperative slots in the target sentence were realised by four different verbs each. A difference in plausibility of the realisations could affect both the reading and the recall task and should therefore be avoided. To this end, all sentence conditions were judged by five native speakers of German (2 linguists, 3 non-linguists) and adjusted until all of them were perceived as equally plausible realisations of the respective sentence. Some verbs were used in more than one sentence: the 60 slots for strong verbs with e/i-gradation in the 30 target sentences were realised by 48 verbs, and the 60 slots for control verbs were realised by 52 verbs.

The location of the slots for the imperative singular form was controlled by two criteria: i) they must not be located at the beginning of a sentence, and ii) they may only occur at the beginning of a main clause with a preceding dependent clause. Even though imperative singular forms usually occur clause-initially, this location needed to be avoided in the stimu-
lus sentences: particularly the analogical imperative singular variants, which are identical with the first person singular forms of the same verbs, would be ambiguous. In order to prevent that participants have to disambiguate some of the stimulus words, slots for the stimulus imperatives were either preceded by adverbials or a dependent clause. In addition, the context of imperative singular slots in the stimuli for the main experiment should fulfil one further criterion: the two words following the imperative singular slot should be identical in all conditions, so that spillover effects on reading times can be analysed.

**WHY IS THERE NO “gebe” CONDITION?**

In preparation of the main experiment, a pilot study was conducted in order to i) test the feasibility of the combined self-paced reading with recall task, ii) assess the average time needed to complete a given number of stimulus items and iii) evaluate the quality of the experiment conditions and stimulus sentences. In this pilot study, imperative singular forms in the target sentences occurred in four forms.

**Table 24: Conditions of target sentences in the pilot experiment**

<table>
<thead>
<tr>
<th>Imperative singular form</th>
<th>Base verb lemma token frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Higher</td>
</tr>
<tr>
<td>unsuffixed imperative with the stem vowel $i$, strong verb with e/i-gradation</td>
<td>a</td>
</tr>
<tr>
<td>unsuffixed imperative with the stem vowel $e$, strong verb with e/i-gradation</td>
<td>b</td>
</tr>
<tr>
<td>suffixed imperative with the stem vowel $e$, strong verb with e/i-gradation</td>
<td>c</td>
</tr>
<tr>
<td>unsuffixed/suffixed imperative, control verb</td>
<td>d</td>
</tr>
</tbody>
</table>

In addition to the conditions of the main experiment (cf. Table 22), the fourth form was the suffixed analogical imperative singular variant of a strong verb with e/i-gradation, e.g. *gebe* ‘give’ (conditions c and g in Table 24). Thus, in contrast to the main test, the pilot experiment had a 2x4 design, so that each sentence was realised in 8 different versions (cf. Appendix F).

With regard to the interactive effect of base verb lemma token frequency and imperative form of the stimuli on participants’ reading times, the unsuffixed imperative variants of strong verbs with e/i-gradation (“stem vowel $i$” and “unsuffixed stem vowel $e$” in Figure 26) show the expected cross-over effect: Whereas the analogical variants with the stem vowel $e$ are read faster in lower than in higher frequency verbs, the traditional variants with the stem vowel $i$ are read faster in higher than in lower frequency verbs. Surprisingly, however, the suffixed analogical imperative variants (“suffixed stem vowel $e$”, solid line in Figure 26) behave quite different from their unsuffixed equivalents. Instead of being read faster among

---

27 In the pilot experiment, stimulus verbs were not taken from bins with high and low frequency verbs; the only requirement was that the distance between a higher and a lower frequency verb should be at least three frequency classes in the DeReWo list (as in the main experiment).
lower frequency verbs, they are consistently read slower than the traditional variants with the stem vowel \(i\) throughout the frequency continuum.

The recall data display a further strong difference between the reactions to the unsuffixed and suffixed analogical imperative variants of the strong verbs with \(e/i\)-gradation. Whereas the unsuffixed variants display reactions in line with the hypotheses, i.e. repetition among lower frequency verbs and increasing replacement among higher frequency verbs (reaching almost 100% replacement in the highest frequency classes), the recall accuracy of the suffixed forms is higher throughout the frequency range (see “unsuffixed stem vowel \(e\)” and “suffixed stem vowel \(e\)” in Figure 27). The explanation for the different reactions to suffixed and unsuffixed analogical variants is captured in the recordings of the repetition of the respective stimulus sentences: in many of them, the repetition of the suffixed stimulus imperatives is preceded by a pause or hesitations or the forms themselves are stressed or accentuated, while this is not the case for the unsuffixed variants. Thus, it is highly probable that these forms were repeated in the recall task only because participants perceived them as unusual (some participants reported this impression explicitly). As participants are supposed to process the imperative stimuli subconsciously, an omission of the suffixed analogical variants from the stimulus set strongly suggested itself.

**Figure 26:** Interactive effect of base verb lemma token frequency and imperative form on residual reading times in the pilot experiment
The main goal of the experimental study is to find out whether the same or similar conserving effect of high token frequency can be found in perception as known from the production (corpus) data. In production, verb token frequency was shown to have a direct influence only on the stem vowel of the imperative singular forms of strong verbs with e/i-gradation, not on their suffixation (cf. Section 5.3). Thus, the suffixed variants are not needed as stimuli in the experiment in order to test the conserving effect in perception, and the respective test conditions (c and g in Table 24) were omitted in the main experiment.

FILLER SENTENCES
Filler sentences did not contain imperative forms (neither singular nor plural) for two main reasons: On the one hand, participants should be distracted from the fact that the aim of the study was to measure response latencies and recall accuracies of imperative forms. Including them in the filler sentences would have led to such a high proportion of imperative sentences in the experimental setup (more than 50%) that participants could have discovered the deception more easily and paid attention to the existence of imperative forms in general, and to their morphological form in particular. On the other hand, the persistence effects found in the corpus study could have translated into a potentially harmful priming effect in the experimental study: Participants in previous studies have been shown to be primed not only by whole words; their reactions are also affected by previous exposure to words sharing a suffix (suffix priming, e.g. Duñabeitia et al. 2008). In order to eliminate the influence of priming, 20 words to the left of the target imperative singular form should not contain any imperative forms.

![Figure 27: Interactive effect of base verb lemma token frequency and imperative form on recall accuracy in the pilot experiment](image-url)
(which equals the window of left context examined in the corpus study). The stimulus sentences were at least 10 words long, and target and filler sentences alternated with one another. As every sentence is repeated, there is a distance of at least 20 words between the presentation of one imperative singular stimulus and the next.

LATIN SQUARE DISTRIBUTION

The allocation of conditions of the target sentences to participants was based on a Latin square. This procedure is commonly used for incomplete factorial designs, i.e. experiment designs in which participants are not presented with all conditions of a stimulus (cf. Bortz & Schuster 2010, Ch. 20).

Table 25: Latin square for stimulus conditions (a-f) in the main experiment

<table>
<thead>
<tr>
<th>List</th>
<th>Sentence 1</th>
<th>Sentence 2</th>
<th>Sentence 3</th>
<th>Sentence 4</th>
<th>Sentence 5</th>
<th>Sentence 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>a</td>
<td>b</td>
<td>c</td>
<td>d</td>
<td>e</td>
<td>f</td>
</tr>
<tr>
<td></td>
<td>unsuffixed imperative with stem vowel i, high frequency strong verb with e/i-gradation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>b</td>
<td>c</td>
<td>d</td>
<td>e</td>
<td>f</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>unsuffixed imperative with stem vowel e, high frequency strong verb with e/i-gradation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>c</td>
<td>d</td>
<td>e</td>
<td>f</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>unsuffixed imperative, high frequency control verb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>d</td>
<td>e</td>
<td>f</td>
<td>a</td>
<td>b</td>
<td>c</td>
</tr>
<tr>
<td></td>
<td>unsuffixed imperative with stem vowel i, low frequency strong verb with e/i-gradation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>e</td>
<td>f</td>
<td>a</td>
<td>b</td>
<td>c</td>
<td>d</td>
</tr>
<tr>
<td></td>
<td>unsuffixed imperative with stem vowel e, low frequency strong verb with e/i-gradation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>f</td>
<td>a</td>
<td>b</td>
<td>c</td>
<td>c</td>
<td>e</td>
</tr>
</tbody>
</table>

The Latin square for the present design is shown in Table 25: Participant #1 is allocated to list #1; thus he/she is presented sentence #1 in condition a, i.e. the imperative singular slot is realised with an imperative singular form with the stem vowel i of a high frequency strong verb with e/i-gradation. He/she reads another imperative singular form of a strong verb with e/i-gradation with the stem vowel i in sentence #4, this time a form of a low frequency verb. Participant #2 also sees two imperative singular forms of strong verbs with e/i-gradation with the stem vowel i but in sentences #3 and 6. As a multiple of the number of conditions is tested, in terms of both the number of sentences (30) and the number of participants\(^{28}\), the list is re-

\(^{28}\) The exact number of participants varies between the participant groups (see Chapter 6.3.2).
peated from the top, so that participant #7 reads the same sentence realisations as participant #1, and sentence #7 is displayed to all participants in the same condition which they were allocated for sentence #1.

6.3.2 Participants

Three groups of participants took part in the main experiment. These groups were designed so as to be able to test the influence of the independent variables participant age and dialectal region. Within each group, an equal number of male and female participants was aimed at; however, a slight overrepresentation of female participants could not be avoided (approx. 60 percent, as compared to approx. 40 percent male participants). All participants were native speakers of German, and all participants received a payment of 10 Euros for taking part in the experiment.

Participants were recruited with the help of leaflets and notices distributed in the university libraries and canteens in Freiburg im Breisgau and Leipzig and bulletin board announcements\(^{29}\). This worked very well for the younger participant groups which mainly comprise students from the local universities. As soon as it became apparent that the older target participant group was not as responsive, employees of the University of Freiburg\(^{30}\) were emailed directly and asked to take part in the experiment. In order to disguise the fact that the data from participants from two dialectal regions in Germany should be compared, the announcement addressed interested participants “from Baden-Württemberg” instead of speakers “from the Low Alemannic region”. However, only participants from the respective dialectal regions were selected to take part in the experiment.

Due to the fact that some participants detected the cover story (see Chapter 6.2), they were not evenly distributed across the lists in the Latin square (see above). As the experiment proceeded, participants were marked for the list which they had been shown and whether or not they had detected the cover story, so that new participants could be assigned to previously underrepresented lists. The numbers of participants in each list who did not detect the cover story are presented in the tables below for each of the groups. The total number of participants in each group is listed in brackets (i.e. including those who detected the cover story). The reasons and procedure for separating the data of participants who detected the cover story from those who did not are explained in Section 6.3.5 below.

\(^{29}\)“Das Schwarze Brett Leipzig”, http://www.dsble.de/in/leipzig

\(^{30}\)Only employees from non-philological departments were emailed, for example from the IT, the economics and the history departments.
YOUNGER PARTICIPANTS FROM BADEN-WÜRTTEMBERG

The first group comprises 44 participants from the federal state Baden-Württemberg at the age of 18 to 30 years. 13 of these participants detected the cover story, so that the data obtained from them could not be used for the main analysis. Participants are distributed across lists as displayed in Table 26.

<table>
<thead>
<tr>
<th></th>
<th>#1</th>
<th>#2</th>
<th>#3</th>
<th>#4</th>
<th>#5</th>
<th>#6</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>female</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>19 (26)</td>
</tr>
<tr>
<td>male</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>12 (18)</td>
</tr>
<tr>
<td>total</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>31 (44)</td>
</tr>
</tbody>
</table>

OLDER PARTICIPANTS FROM BADEN-WÜRTTEMBERG

The second group of participants comprises 30 speakers between 43 and 75 years of age from Baden-Württemberg, 27 of whom did not detect the cover story. Their distribution across participant lists is shown in Table 27 below.

<table>
<thead>
<tr>
<th></th>
<th>#1</th>
<th>#2</th>
<th>#3</th>
<th>#4</th>
<th>#5</th>
<th>#6</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>female</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>16 (18)</td>
</tr>
<tr>
<td>male</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>11 (12)</td>
</tr>
<tr>
<td>total</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>27 (30)</td>
</tr>
</tbody>
</table>

It might seem problematic to present participants from two very different age groups with the same self-paced reading with recall task. However, the effect of “cognitive slowing” with age (Craik et al. 1995: 215) on reading times will be accommodated in the statistical analysis (see 6.4.2.1), and several studies have shown that the ability to repeat sentences is not influenced by age (Alloway & Alloway 2013; Meyers et al. 2000; Craik et al. 1995: 214-216). Thus, differences between this and the younger participant groups can be treated as differences in linguistic, not neuropsychological, behaviour.

YOUNGER PARTICIPANTS FROM SAXONY

Finally, 46 participants from the federal state Saxony between 18 and 30 years of age are in the third participant group. The data obtained from 18 of them needed to be analysed separately because these participants detected the cover story.
Table 28: Distribution of participants in group 3 (younger, from Saxony) across participant lists

<table>
<thead>
<tr>
<th>list</th>
<th>#1</th>
<th>#2</th>
<th>#3</th>
<th>#4</th>
<th>#5</th>
<th>#6</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>female</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>14 (28)</td>
</tr>
<tr>
<td>male</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>14 (18)</td>
</tr>
<tr>
<td>total</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>28 (46)</td>
</tr>
</tbody>
</table>

CONTROL GROUP

As explained before, the present study uses a combined experiment design, i.e. self-paced reading with recall. Participants’ reading times can be expected to be influenced by the fact that at the same time as reading a sentence, they also try to memorise it. Aaronson and Scarborough are among the first researchers who used this “sentence repetition” design and report a significant difference in reading times between “subjects who must later recall a sentence and subjects who must simply comprehend it” (1976: 56). Their results cannot be fully applied to the present study, however, as they investigated language processing in English and their method is slightly different: the recall task consisted in participants writing down the sentence instead of orally repeating it. Kuhn similarly reports a significant difference between a ‘verbatim repetition’ (“wörtliche Wiedergabe”) and a ‘comprehension’ (“Verstehen”) group in her data obtained from native speakers of German (2011: 97); again, speakers wrote down the sentences in the recall task. Section 6.4.1 will show that, even though their experiment design is almost identical, the authors report very different figures for the impact of recall on reading times. Due to these and other methodological differences between the two papers and the present study, a control group was introduced in the current experiment in order to be able to assess the impact of the recall task on reading times in this particular experiment.

Participants in the control group fulfil the same criteria as those in participant group 1, i.e. they are native speakers of German from the federal state Baden-Württemberg between 18 and 30 years of age. They read the same stimulus sentences as those in the three participant groups above; however, instead of repeating each sentence after having read it, they were presented comprehension questions, to be answered with “yes” or “no” (the original comprehension questions are included in Appendix E with the stimulus sentences they refer to). Comprehension questions are related to content at the beginning, in the middle, or towards the end of the stimulus sentence, or they relate to the whole sentence content; these categories were evenly distributed across the stimulus sentences. The correct answers alternate between “yes” for odd-numbered and “no” for even-numbered stimulus sentences; randomisation of stimulus sentences by the Linger software ensures randomisation of positive and negative answers.
12 participants make up the control group, none of whom detected the cover story so that they could be evenly distributed across the participant lists (see Table 29). The reading-time data from these speakers is compared only to the data from participant group 1.

**Table 29:** Distribution of participants in the control group across participant lists

<table>
<thead>
<tr>
<th></th>
<th>#1</th>
<th>#2</th>
<th>#3</th>
<th>#4</th>
<th>#5</th>
<th>#6</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>female</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>male</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>total</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>12</td>
</tr>
</tbody>
</table>

### 6.3.3 Annotation of variables

#### 6.3.3.1 Dependent variables

During the experiment, reading times were measured and the repetitions of stimulus sentences were recorded with the help of a headset.

**READING TIME (RT)**

While the participants read the stimulus sentences, reaction time was measured for all individual words in milliseconds from the point at which the word was presented on the screen to the next key press. A data frame containing all observations for every participant was automatically created by the software Linger, including the annotation for the first dependent variable reaction time (RT). Thus, the dataset for participant groups 1-3 comprises 63,669 data points in total, of which 31,562 are reading times of stimulus words in filler sentences and 32,107 are reading times of words in target sentences. Out of the latter subset, 7,740 data points are reading times of the stimulus imperatives and the two stimulus words following them in the respective sentence (2,580 data points each). The dataset for the control group comprises 8,896 data points, among them 4,416 for filler sentences and 4,480 for stimulus words in target sentences.

All reading times obtained in the experiment were used in the statistical analysis, regardless of whether the imperative singular form was repeated as presented or replaced by a different form, or whether the comprehension question was answered correctly or incorrectly. In other self-paced reading studies (e.g. Jiang 2012; Pappert et al. 2007; Spivey-Knowlton 1992), reading times for trials are often excluded when comprehension questions are answered incorrectly because the incorrect answers are interpreted as a sign that the participant did not pay full attention while reading the stimulus sentence. In the present study, however, participants’ attention during reading is ascertained by the subsequent recall task. Moreover, researchers like Contemori et al. provide evidence that the exclusion of reading times for incorrect trials does not necessarily affect results: “For the self-paced reading task, the data
from trials in which the comprehension question was answered accurately or inaccurately were included in the analysis. We also analyzed the data excluding the trials based on comprehension accuracy, but the overall pattern of results did not change” (2015: 268). After all, participants may have misunderstood a sentence during reading and still answer the related comprehension question correctly, perhaps by chance, or recall it correctly once they have accessed its meaning as a whole; these “correct” trials would not be removed by the above exclusion procedure. Thus, reading times for all trials were entered into the regression analysis in order to avoid a loss of valuable data; outlier data points were removed after the effect of predictor variables and random factors had been assessed (see Appendix B).

RECALL ACCURACY (REC.ACC)

The recordings of the subsequent repetition of stimulus sentences by participants were inspected for whether the target imperative form in the sentence was repeated as presented. This second dependent variable REC.ACC has several factor levels:

<table>
<thead>
<tr>
<th>factor level</th>
<th>criterion</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>“yes”</td>
<td>exact repetition</td>
<td>2,109</td>
</tr>
<tr>
<td>“no”</td>
<td>presented imperative stem vowel was replaced,</td>
<td>174</td>
</tr>
<tr>
<td></td>
<td>e.g. geb by gib or vice versa</td>
<td></td>
</tr>
<tr>
<td>“unclear”</td>
<td>unclear which stem vowel was used by the participant in recall</td>
<td>10</td>
</tr>
<tr>
<td>“other”</td>
<td>verb of the stimulus imperative was replaced by a different verb</td>
<td>103</td>
</tr>
<tr>
<td>“error”</td>
<td>defective recordings,</td>
<td>184</td>
</tr>
<tr>
<td></td>
<td>participants could not recall the relevant parts of the sentence</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td></td>
<td>2,580</td>
</tr>
</tbody>
</table>

The annotations for this variable were added to the existing data frame. However, only instances annotated with “yes” or “no” were used in the statistical analysis of recall accuracies.

It must be noted that, for convenience, the annotation of the second dependent variable REC.ACC was not based on measurements in a speech analysis programme, such as PRAAT\(^{31}\), but on the auditory impression of the experimenter. Following recommendations given by Bayerl and Paul for testing the “trustworthiness of [manual] annotations” (2011: 719-720), a random five percent sample\(^{32}\) of the relevant\(^{33}\) recordings was given to four different raters (German native speakers; linguists but not phoneticians). The sample consisted of 107 instances: in 55 of them, stimulus imperatives with the stem vowel \(i\) had been presented

\(^{31}\) http://www.praat.org (Paul Boersma & David Weenink, University of Amsterdam)

\(^{32}\) The authors do not make a recommendation with regard to the size of the sample for testing inter-coder agreement. In the present study, a five percent sample was considered a feasible size for assessing agreement while still asking a favour of the other raters.

\(^{33}\) Instances marked by the experimenter as “other” and “error” were deemed unproblematic; thus, only instances of the factor levels “yes”, “no” and “unclear” of the variable REC.ACC occurred in the sample.
to participants, and in 52 cases, stimulus imperatives with the stem vowel e. After familiarising themselves with the type of material by listening to 10 practice recordings, the raters were asked to assess the sample in three categories: “i” for when the imperative they heard was used with the stem vowel i, “e” for an imperative with the stem vowel e, and “unclear” for when they were unsure about the vowel quality in the imperative. Raters saw the context to the left and right of the target imperative, but they were blind to the verb whose imperative they heard (to exclude any potential priming effects caused by reading the infinitive stem, e.g. geben) and, self-evidently, to the experiment condition in which the sentence was originally presented to the participant.

According to Bayerl and Paul (2011: 703), “agreement percentages still [seem] to be one of the most important indicators for annotator agreement”. In the present study, intercoder agreement was measured in terms of agreement percentages according to the so-called “consensus” method, in which “agreement is only given if all annotators agree in their annotation of an instance” (Bayerl & Paul 2011: 718). Even though this makes it the most conservative method for calculating agreement percentages, a very high figure can be reported for the presented study: 101 of the 107 instances in the sample were rated equally by the experimenter and all other raters, resulting in 94.39 % agreement. In 18 of the 107 instances in the sample, the experimenter detected a replacement of the presented stem vowel e in the imperative by the stem vowel i, or vice versa; this annotation was confirmed by all other raters for 16 instances (88.89 %).

However, many researchers also use kappa as an indicator for agreement because its formula corrects for chance agreement (assuming that raters guess when they are in doubt). As in the present task instances are rated in two dominant categories (“i” and “e”), the danger of chance agreement is reasonably high. Therefore, Fleiss’ kappa was calculated as well for the dataset of 107 recordings annotated by all 5 raters: $\kappa = 0.941$, p<.001 (function kappam.fleiss() in package irr for R, version 0.84; Gamer et al. 2012). Both the percent agreement and Fleiss’ kappa value are very close to the maximum possible values of 100 and 1, respectively; thus, the annotation of recordings by the experimenter on the basis of auditory impression can be considered trustworthy, and the annotations can be used as the basis for a statistical analysis of participants’ recall accuracy.

---

34 Cohen’s kappa for pair-wise comparisons (Cohen 1960), Fleiss’ kappa for more than two raters (Fleiss 1971)
3.3.2 Predictor variables

VERB LEMMA TOKEN FREQUENCY
Although a binary distinction between higher and lower frequency verbs formed the basis of stimuli creation, the original frequency classes of verbs in the DeReWo list were entered into the data frame as the annotation for the variable lemma token frequency. On the basis of the results of the Walkthrough Corpus study, it was assumed that base verb lemma token frequency (BVLT) has a significant effect in the regression analysis of the experiment data, while verb lemma token frequency (VLTF) should not have a significant effect on reading times or recall accuracies. In order to test this assumption, all data points were annotated for both variables: simplex and prefixed verbs have the same values for both variables, whereas particle verbs received the BVLT values of their simplex bases and the value of VLTF which was listed for them in the DeReWo list. Values of both variables range from “6” for the most frequent verb geben to “23” for the least frequent verbs zerdreschen and zerwerfen in the dataset (cf. Appendix A).

IMPERATIVE FORM
As explained in Section 6.3.1, imperative singular stimuli occurred in three forms: “stem vowel i”, “stem vowel e”, “control verb”; these are also the three factor levels of the variable IMP.FORM. Data points could be automatically annotated for this variable on the basis of the stimulus sentence conditions which were recorded in the dataset compiled by the experiment software Linger: reading times and recall accuracies were obtained for imperative singular forms with the stem vowel i in conditions a and d, for imperatives with the stem vowel e in conditions b and e, and for imperatives of control verbs in conditions c and f (see Table 22).

PARTICIPANT GROUP
Each participant received a participant number by which the obtained data could be attributed to them. In a dataset with participants’ metadata obtained from the background questionnaire (see 6.3.4), the group to which a participant belonged (see previous section) was recorded. In this way, the regression analysis could test whether reading times or recall accuracies are significantly different between groups. Comparisons between the data from younger and older groups from Baden-Württemberg test whether a change from the traditional imperative singular formation with the stem vowel i towards analogical formation with the stem vowel e can be traced along a longer time span than was possible in the Walkthrough Corpus study. The
data from these two groups are in turn compared to those from the participants from Saxony, in order to track whether change is proceeding throughout Germany.

**WORD LENGTH**

The variable word length was also (automatically) annotated for all data points in the data set by counting the number of letters in a stimulus word. This variable is not a predictor in the same sense as the above-named variables: as word frequency and word length are highly correlated, its purpose is mainly to control that a frequency effect is not an effect of word length in disguise, especially in the regression analyses of reading times.

### 6.3.3.3 Random variables

A number of random variables were entered into the regression model, in order to generalise over their effects on the dependent variables. Individual differences in the performance of the participants, i.e. overall differences in reading pace, slope (familiarisation or fatigue effects across the stimulus sentence and experiment) and sensitivity to word length, word frequency and imperative form are accommodated with the help of the participant identification (PART; POS.IMP|PART; POS.ITEM|PART; W.LEN|PART; BVLTF|PART; IMP.FORM|PART). Individual items (sentences) and the stimulus verbs may also have idiosyncratic effects on reading times; thus, these variables were likewise taken up as random predictors (ITEM.ID; BASE.VERB). Position effects can be expected both for the sentences themselves (POS.ITEM), i.e. the order in which they are presented, and for the position of the imperative singular form in the sentence (POS.IMP): they may be read and recalled differently towards the beginning of a sentence than towards the end (cf. Aaronson & Scarborough 1976; Baayen 2008: 244).

### 6.3.4 Meta data from the background questionnaire

All participants filled in a background questionnaire after they had completed the self-paced reading with recall test. It contained questions about participants’ metadata: their AGE, GENDER, whether they are dialect speakers (variable DIALECT), etc. The answers to these questions were entered in a separate dataframe to which information like the participant number and the participant list in the Latin square which was allocated to the participant (variable LIST) were added; this data frame was used to test if any of these variables had an influence on whether participants detected the cover story and thereby the aim of the experiment (cf. 6.5.1). The metainformation dataframe was then merged with the dataframe containing the data about the dependent variables reading times and recall accuracies, the predictor variables like verb token frequency and imperative form, and the random variables for all obtained data
points. In this way, regression analyses of reading times and recall accuracies could also take into account the possible influence of sociolinguistic variables.

6.3.5 Separation of data from participants who did or did not detect the aim of the study

As explained in Section 6.3.2, some of the participants in the experiment detected the cover story, i.e. they recognised at some point during the experiment that its purpose was to test reactions to forms of the imperative singular. This presents a problem for the analysis of their reading time and recall accuracy data because they develop unintended expectations; to put it differently, they might wait for the next instance of an “unusual” imperative singular form instead of processing and repeating stimulus sentences subconsciously, as envisaged in the present study. Nevertheless, the data obtained from these participants was not excluded from the analysis altogether, since they can provide valuable insights into i) speakers’ awareness of the ongoing change under investigation and ii) the behaviour towards the innovative analogical variants of speakers who become aware of the competition in the imperative singular. These participants’ data were therefore annotated in the same way as outlined above but separated from those of the participants who did not detect the cover story on the basis of the background questionnaire and debriefing:

In addition to enquiring about participants’s meta data, one part of the questionnaire triggered them to consciously reflect upon their participation in the experiment. For instance, they should indicate whether they had developed an intuition of what the experiment aimed at and when they had done so. The annotation for whether a participant had detected the aim of the experiment, however, was not exclusively based on the answer to this question. A few participants indicated in the questionnaire that they had not developed an intuition. Nevertheless, after having been debriefed about the aim of the experiment, they answered the question ‘Have you “guessed” the proper object of study of the test without this explanation?’ with “yes”. The annotation in the dataframe for whether participants detected the cover story (variable DETECTED) was therefore based on the answer to the debriefing question. In contrast, annotations for when participants felt they had developed the intuition (variable WHEN.DETECTED) were taken from the background questionnaire because the answers to this question should not be affected by having heard the debriefing summary. The question which can be answered on the basis of the annotation for this variable is whether participants read and repeated sentences differently before and after they detected the aim of the experimental study.

35 “Haben Sie den eigentlichen Untersuchungsgegenstand des Tests auch ohne diese Erklärung ‘erraten’?” (see Appendix E)
6.4 Results for participants who did not detect the aim of the study

In the following, separate sections (6.4.2 and 6.4.3) will be devoted to the analysis of the reading times and recall accuracies of those participants who did not detect the cover story. Section 6.5 will present the analysis of data obtained from those participants who detected the aim of the experiment.

6.4.1 Impact of the recall task on reading times

As explained in previous sections, a control group was presented with the same stimulus sentences as the participant groups 1-3, but they answered comprehension questions about the sentences instead of repeating them; in this way, it should be possible to assess the impact of the recall task on reading times. As Table 31 shows, Aaronson and Scarborough (1976) and Kuhn (2011) report very different figures for the difference between “comprehension”, referring to “classical” self-paced reading with comprehension questions after sentences, and “repetition”, i.e. self-paced reading with recall.

Table 31: Reading times in comprehension and repetition condition in Aaronson & Scarborough (1976) and Kuhn (2011)

<table>
<thead>
<tr>
<th></th>
<th>mean reading times</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>comprehension</td>
<td>repetition</td>
<td>difference</td>
</tr>
<tr>
<td>Aaronson &amp; Scarborough (1976: 56)</td>
<td>402 msec</td>
<td>583 msec</td>
<td>181 msec</td>
</tr>
<tr>
<td>Kuhn (2011: 97)</td>
<td>463.72 msec</td>
<td>785.05 msec</td>
<td>321.33 msec</td>
</tr>
</tbody>
</table>

In the present study, the reading times of the first participant group, i.e. participants from Baden-Württemberg between 18 and 30 years of age, were compared to the reading times of the control group who meet the same criteria. The comparison was carried out by means of a Mann-Whitney U test (also called Wilcoxon rank sum test), a variant of the t-test for data that are not normally distributed. Participants in the control group show a mean reading time of 432.32 msec per word, which is significantly lower than the mean 606.39 msec per word observed for participants in participant group 1 (U = 125674260; p < 0.001); thus, the difference in reading time between “comprehension” and “repetition” groups in the present study is 174.07 msec. Even though the language under investigation in their study is English, the mean reading times reported by Aaronson and Scarborough (1976) are closer to the figures in the present study than Kuhn’s (2011) who also investigated German. This finding can in part be attributed to the stimuli used in the three studies. Participants in Kuhn’s study read sentences of 16 to 20 words length, with a mean of 17.8 words per sentence (2011: 72); in Aaronson and Scarborough’s study, sentences were between 9 and 19 words long (mean = 14.4; 1976: 58). In the present experiment, participants read sentences between 10 and 15 words length (mean = 12.32, standard deviation = 0.84). Mean reading times for both
groups in the present study are slightly higher than in Aaronson and Scarborough’s (1976) study although stimulus sentences are shorter. Thus, the fact that words in German tend to be longer is reflected in the reading times. As participants in Kuhn’s (2011) study read longer sentences, this increasing cognitive demand is reflected in the even higher reading times in “comprehension”; the immense difference in reading times in “repetition” can be explained by increasing demands as well, coupled with a generally higher word length in German.

The reading times reported for participant groups 1-3 in the following sections are thus not comparable to data obtained from participants in “classical” self-paced reading (without recall). However, the analysis of the data obtained from the control group illustrates how the perhaps surprisingly high reading times in all three groups do not reflect demands imposed by the German language as such, by complex syntax in stimulus sentences or the like but rather multitasking demands.

6.4.2 Reading times

6.4.2.1 Residualisation

Many studies have shown that both the frequency of occurrence of any given language item and its length (for example, in terms of characters) are correlated with the time needed to read/process it (among others, Whaley 1978; Hauk & Pulvermüller 2004; Cortese & Balota 2012). The more frequent or the shorter the word, the lower is the reading time on that particular word; in other words, the effects of both factors on reading times are confounded. In order to be able to assess the influence of frequency independently of a word length effect, both factors could be entered into the mixed-effects regression model as fixed factors. The factor word length would then “claim” as much of the variation in reading times as it can explain and leave the rest to the remaining fixed factors and interactions in the model, among them frequency. Alternatively, the influence of word length on reading times can be determined in a separate regression. In such a regression model, log-transformed reading times would be entered as the dependent variable and word length as a predictor (fixed factor). Random effects could be added for participants’ intercepts and word length slopes. The residuals of this model can be understood as values for the variable reading time which are corrected for the influence of word length: they represent the share of variation in reading times which is not accounted for by the variation in word length in the stimuli. Figure 28 illustrates the regression of word length on (log-transformed) reading times obtained in the present experiment, with arrows indicating residuals, i.e. deviations of data points from the regression line.
These residual reading time values could be used as the dependent variable in the actual regression models in order to be able to determine the influence of frequency irrespective of the influence of word length. However, the present dataset poses another difficulty which is illustrated in Figure 29.

Not only are reading times affected by the length of stimulus words, indicated by ascending regression lines, but this influence is not uniform in all three participant groups. The older participants from Baden-Württemberg (“bawü.old” in Figure 29) generally show higher reading times (mean = 850.64, sd = 33.19) than the two younger participant groups (“bawü.young”: mean = 606.39, sd = 49.56; “sax.young”: mean = 634.51, sd = 36.45); there-
fore, their regression line in Figure 29 has a higher intercept. In other words, when participant
group is added as a fixed factor to the above-sketched residualisation model, its effect is sig-
ificant.

The central research question of the experiment is whether the form of a presented im-
perative singular and/or the frequency of the verb in which it is presented affect the reactions
of the three participant groups differently, not whether the groups generally read faster or
slower. Contrary to the effect of word length on reading times which could be captured in a
fixed factor in the final regression model, these differing participant group intercepts could
not be controlled for easily. Thus, if in the final regression models the original (log-
transformed) reading times were used as the dependent variable, the effect of frequency
would not be undermined by the influence of word length, but the main effect of the group
variable might be overestimated. In a test run with a variant of the final regression model
which contains the group variable as a predictor and the original log-transformed reading
times as the dependent variable, the contrast between younger and older participants from
Baden-Württemberg is significant: \( t(4) = 3.14; p < 0.01 \); this is not the case when the residual-
ised reading times are entered as the dependent variable: \( t(4) = 1.359; p = 0.17505 \). This dif-
fERENCE must be interpreted as an artefact of the generally higher reading times obtained from
the older participant group.

In order to control for the influence of word length and allow differing intercepts for
participant groups, preliminary regression models were therefore fitted on subsets of the data
for the three participant groups, assessing only the effect of word length on reading times. For
convenience, reading times were log-transformed (cf. Baayen 2008: 31); random variables
entered into the models were individual intercepts for participants \( (1|\text{PART}) \), and random
slopes for the sensitivity of participants to word length \( (1+\text{W.LEN}|\text{PART}) \). The values for
residual log-transformed reading times \( (\text{RESID. LOG.RT}) \) from these regression models were
used as the dependent variable in all following regression models. In order to allow for differ-
ent processing patterns in the three participant groups without overestimating the significance
of interactive effects, separate regression models were fitted on the three subsets of data.

6.4.2.2 Results of the regression analysis

In the actual regression models assessing the influence of predictors on reading times of par-
ticipants who did not detect the aim of the study, the variable \( \text{RESID.LOG.RT} \) was entered as
the dependent variable. Random intercepts were introduced for stimulus verbs
\( (1|\text{BASE.VERB}) \), stimulus sentences \( (1|\text{ITEM.ID}) \) and participants \( (1|\text{PART}) \); participant-
specific random slopes were entered for the position of the stimulus in the experiment (1+POS.ITEM\|PART), the position of the imperative in the stimulus sentence (1+POS.IMP\|PART) and participants’ sensitivity to word length (1+W.LEN\|PART), base verb lemma token frequency (1+BVLTF\|PART) and imperative form (IMP.FORM\|PART).

The influence of fixed factors was tested in stepwise regression modelling. The baseline models included only the model intercept, the necessary fixed factors POS.ITEM and POS.IMP\textsuperscript{36} and the random factors. The fixed factors whose effect should be tested, i.e. base verb lemma token frequency (BVLTF) and imperative form (IMP.FORM), were successively added to the regression models, and an interactive effect of the two predictors on reading times was tested. Log-likelihood tests determined whether the single predictors and interactions significantly improved the performance of the regression models in predicting the variation in the dependent variable reading time. The analysis revealed that BVLTF and IMP.FORM have a significant interactive effect on reading times obtained from participants from Baden-Württemberg (p < 0.05), but the two predictors only have an additive effect on reading times obtained from participants from Saxony (p < 0.01).

For the sake of comparability in log-likelihood tests, intermediate models were not corrected for outliers. The final regression models were corrected for outliers, whereby 4.76 percent of the data points obtained from speakers from Saxony were removed from analysis, and 4.82 and 4.52 percent of the data points in the case of the older and younger speakers from Baden-Württemberg, respectively. The results of the final regression models are displayed in Table 32.

Apart from the predictors which were specifically varied in the experiment design, two other fixed factors show up with a significant effect in the regression models. POS.ITEM and POS.IMP both have a significant main effect on reading times. POS.ITEM, the position of the stimulus in the experiment, has a negative estimate; thus, all participants read faster as the experiment proceeds. This is a classical familiarisation effect: participants get used to the experiment task and can accomplish it faster with every stimulus item. The positive estimate of POS.IMP, position of imperative in stimulus sentence, can be explained by the cognitive demand on participants which increases with every word they encounter in a sentence and which they need to incorporate into the sentence structure. Therefore, imperatives are read slower the later they occur in the stimulus sentence; the effect of this predictor, however, is only significant in the analysis of data obtained from older speakers from Baden-Württemberg.

\textsuperscript{36} When participant-specific random slopes are entered into a regression model, the grouping variables (in this case, POS.IMP and POS.ITEM) should be entered as fixed factors as well (cf. Baayen 2008: 249)
Table 32: Results of regression analysis of residual reading times (Krause-Lerche 2019b: 27)

<table>
<thead>
<tr>
<th>variable factor level</th>
<th>Saxony (18-30 years)</th>
<th>Baden-Württemberg (43-75 years)</th>
<th>Baden-Württemberg (18-30 years)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>estimate standard z value p-value</td>
<td>estimate standard z value p-value</td>
<td>estimate standard z value p-value</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.038 0.032 1.167 0.249</td>
<td>-0.031 0.030 -1.049 0.300</td>
<td>-0.037 0.025 -1.458 0.151</td>
</tr>
<tr>
<td>POS.IMP</td>
<td>0.006 0.009 0.640 0.526</td>
<td>0.019 0.007 2.706 0.010 *</td>
<td>0.009 0.005 1.649 0.109</td>
</tr>
<tr>
<td>POS.ITEM</td>
<td>-0.004 0.001 -4.793 0.000 ***</td>
<td>-0.005 0.001 -6.111 0.000 ***</td>
<td>-0.003 0.001 -5.352 0.000 ***</td>
</tr>
<tr>
<td>BVLTF (numeric)</td>
<td>0.007 0.003 2.073 0.045 *</td>
<td>0.003 0.004 0.604 0.547</td>
<td>-0.001 0.004 -0.280 0.781</td>
</tr>
<tr>
<td>IMP.FORM</td>
<td>reference level = e</td>
<td>reference level = e</td>
<td>reference level = e</td>
</tr>
<tr>
<td></td>
<td>-0.093 0.027 -3.433 0.002 **</td>
<td>-0.024 0.026 -0.927 0.363</td>
<td>0.012 0.020 0.589 0.557</td>
</tr>
<tr>
<td>BVLTF×IMP.FORM</td>
<td>reference level = BVLTF×e</td>
<td>reference level = BVLTF×e</td>
<td>reference level = BVLTF×e</td>
</tr>
<tr>
<td>control</td>
<td>-0.078 0.028 -2.804 0.009 **</td>
<td>0.020 0.024 0.836 0.410</td>
<td>-0.010 0.020 -0.492 0.624</td>
</tr>
<tr>
<td>BVLTF×i</td>
<td>0.004 0.005 0.848 0.397</td>
<td>0.008 0.005 1.543 0.123</td>
<td>0.012 0.005 2.418 0.018 *</td>
</tr>
<tr>
<td>BVLTF×control</td>
<td>0.009 0.005 1.745 0.081</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>random effects variance standard deviation</th>
<th>variance standard deviation</th>
<th>variance standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1</td>
<td>ITEM.ID)</td>
<td>0.0105 0.1024</td>
</tr>
<tr>
<td>(1</td>
<td>BASE.VERB)</td>
<td>0.0017 0.0414</td>
</tr>
<tr>
<td>(1</td>
<td>PART)</td>
<td>0.0005 0.0129</td>
</tr>
<tr>
<td>(1+POS.IMP</td>
<td>PART)</td>
<td>0.0002 0.0213</td>
</tr>
<tr>
<td>(1+POS.ITEM</td>
<td>PART)</td>
<td>0.0000 0.0029</td>
</tr>
<tr>
<td>(1+W.LEN</td>
<td>PART)</td>
<td>0.0002 0.0131</td>
</tr>
<tr>
<td>(1+BVLTF</td>
<td>PART)</td>
<td>0.0001 0.0087</td>
</tr>
<tr>
<td>(1+IMP.FORM</td>
<td>PART) i</td>
<td>0.0074 0.0859</td>
</tr>
<tr>
<td>(1+IMP.FORM</td>
<td>PART) control</td>
<td>0.0062 0.0785</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>measures of model quality</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>800</td>
<td>771</td>
<td>888</td>
</tr>
<tr>
<td>AIC (df)</td>
<td>259.897 (28)</td>
<td>198.787 (30)</td>
<td>119.467 (30)</td>
</tr>
<tr>
<td>marginal R2</td>
<td>0.0845</td>
<td>0.1718</td>
<td>0.0815</td>
</tr>
<tr>
<td>conditional R2</td>
<td>0.2907</td>
<td>0.2599</td>
<td>0.1365</td>
</tr>
</tbody>
</table>

Notes: control = control verb, e = imperative singular of strong verb with e/i-gradation with stem vowel e, i = imperative singular of strong verb with e/i-gradation with stem vowel i
With respect to the influence of the focal predictors base verb lemma token frequency (BVLTF) and imperative form (IMP.FORM), the processing patterns of the three participant groups are largely divergent, as illustrated in Figure 30 (note that BVLTF on the x-axis is measured in classes on a decreasing scale; thus, the higher the frequency class, the lower the base verb lemma token frequency). The only obvious similarity between the three participant groups is in reading times of imperatives of control verbs: they generally increase with decreasing base verb lemma token frequency, and the slight difference in the slope of the regression lines across the three participant groups is not significant (p > 0.05). This pattern matches the one sketched in the middle panel of Figure 20: as there is no variability in the imperative singular of the control verbs, processing of these stimuli only depends on the base verb lemma token frequency of the stimulus imperative. Residual reading times for imperatives of strong verbs with e/i-gradation, by contrast, strongly depend on the group to which a participant belongs. Whereas the two predictors BVLTF and IMP.FORM only have significant main effects on residual reading times obtained from speakers from Saxony, they also have an interactive effect on residual reading times obtained from both participant groups from Baden-Württemberg.

In the reading time data obtained from participants from Saxony, the regression lines for both imperative variants of strong verbs with e/i-gradation increase almost parallel; i.e. reading times increase with decreasing base verb lemma token frequency, and the analogical imperative variants with the stem vowel e are generally read slower than the traditional va-
riants with the stem vowel i (this difference is significant, p < 0.01). In contrast, both groups from Baden-Württemberg read the traditional imperative singular variants of strong verbs with e/i-gradation faster in (base) verbs with a higher lemma token frequency than the analogical imperative variants of the same verbs. The opposite is true for the lower base verb lemma token frequency range: here, the analogical imperative singular variants are read faster than the traditional variants of the same verbs. Thus, whereas the data of speakers from Baden-Württemberg are very similar to the expectations sketched in Figure 23, the results for speakers from Saxony are far more similar to the hypothesised reading time results for a case of stable variation between traditional and analogical imperative variants in the right panel of Figure 20.

Two important findings are true for all participant groups: All three groups show significant processing differences between the competing traditional and analogical imperative singular variants of the same strong verb with e/i-gradation. Moreover, out of the four stimulus sentence conditions involving a strong verbs with e/i-gradation, the traditional imperative variants of high frequency verbs are read fastest by all groups, i.e. faster than their analogical counterparts and faster than both traditional and analogical variants of low frequency verbs. This suggests that the frequency of occurrence of individual linguistic forms or variants is reflected in how fast speakers can access the respective forms in their mental lexicons, regardless of the speed of access to the verb lemmas.

The reading-time patterns of both groups from Baden-Württemberg seem to reflect ongoing change in the imperative singular variants of strong verbs with e/i-gradation as well. Whereas older speakers from Baden-Württemberg read traditional imperative variants faster than analogical ones in verbs belonging to all BVLTF classes up to class 18, younger speakers from the same region read traditional imperatives faster only in very highly frequent verbs up to BVLTF class 10. This result can be understood as evidence of a change-in-progress in the imperative singular of strong verbs with e/i-gradation. According to the results of the Walkthrough Corpus study, the use of the analogical imperative singular variants of these verbs has increased over time; therefore, the younger participants may have encountered these forms more often than the older speakers in the sample. The higher frequency of occurrence of analogical imperative variants during the lifetime of the younger speakers is reflected in the observation that they process such forms faster than the older speakers who have witnessed a lower frequency of occurrence of analogical imperative variants in lower frequency verbs during their lifetime.
6.4.2.3 Discussion of the reading time results

As mentioned before, it has been shown in a number of previous studies that the token frequency of a given linguistic item is negatively correlated with the time needed to read it (e.g. Whaley 1978; Hauk & Pulvermüller 2004). Thus, the finding that the token frequency of verbs has a significant influence on reading times in the present experiment comes as no surprise. However, the above regression model only tested for an effect of base verb lemma token frequency, the factor whose values were varied during stimulus compilation. The decision to vary base verb lemma token frequency was based on the results of the analysis of the Walk-through Corpus data which had revealed that this variable, rather than verb lemma token frequency, determines the choice of the stem vowel in the imperative singular of strong verb with e/i-gradation. It was therefore hypothesised that it would likewise influence participants’ reading times of the two imperative stem variants. In order to test whether verb lemma token frequency really has no effect on the processing of the imperatives, variants of the final regression models reported in Table 32 were fitted on the data. In these regression models, reading times were residualised for word length of the verb lemma (instead of the base verb lemma, cf. Section 6.4.2.1). The fixed factor BVLTF was exchanged for verb lemma token frequency (VLTF); the random intercept for base verb (1|BASE.VERB) was likewise replaced by one for verb lemma (1|VERB) and the random slope for participants’ sensitivity to base verb length (1+W.LEN|PART) by one for verb lemma length (1+LEMA.LEN|PART). Log-likelihood tests revealed that, compared to the original base verb regressions, the performance of the verb lemma variants of the regressions in explaining the variation in reading times was significantly lower: p < 0.05). Thus, in accordance with the findings for language production from the walkthrough corpus study, base verb lemma token frequency explains the processing of the imperative singular forms of strong verbs with e/i-gradation significantly better than verb lemma token frequency does.

It was hypothesised further that reading times depend on two more predictor variables, imperative form (IMP.FORM) and participant group (GROUP). The analysis indeed revealed an interactive effect of BVLTF and IMP.FORM in the data obtained from speakers from Baden-Württemberg and an additive effect of both predictors on reading times obtained from participants from Saxony; this amounts to a three-way interaction between BVLTF, IMP.FORM and participant GROUP. Participants from Baden-Württemberg read traditional imperatives with the stem vowel i faster in higher frequency verbs and thus have them more strongly entrenched in their mental lexicons than their analogical counterparts. In the lower verb frequency range, by contrast, they read traditional imperative variants read slower and
thus have them less strongly entrenched than the analogical imperative variants. While data obtained from participants from Baden-Württemberg display this cross-over effect, the reading times for imperative variants with the stem vowel \( e \) measured from Saxonian participants are generally higher than those for imperative variants with the stem vowel \( i \); thus, the entrenchment of the traditional imperative variants in the mental lexicons of these speakers is generally stronger than that of the analogical variants. This result confirms one of the hypotheses formulated in Section 6.1.1: the different imperative variants of a verb are not entrenched in speakers’ mental lexicons to the same extent. If they were, there should be no such profound (and statistically significant) differences in the processing of alternative forms of one and the same verb as can be observed in the reaction times of all three participant groups.

As mentioned before, the traditional imperative variants of high frequency strong verbs with \( e/i \)-gradation are read fastest by all three participant groups: these forms therefore are most strongly entrenched in their minds. This result can be taken as converging evidence for the prevalent explanation of the conservation of irregularity in high frequency linguistic items: The high frequency of occurrence of traditional imperative singular variants of high frequency verbs in the corpus study (cf. Figure 22) is reflected in the finding that exactly these forms are processed faster and are thus more strongly entrenched in the mental lexicon of speakers than the alternative analogical imperative variants of the same high frequency verbs and more strongly than the imperatives of lower frequency verbs in general. Owing to their high token frequency and, as demonstrated in the present experiment, their strong entrenchment, these forms are so easily retrievable from memory that they resist replacement by analogical formation patterns.

Taken together, the order of the panels in Figure 30 (the arrangement of participant groups) could represent an apparent-time development captured by participants’ age and regional background. Younger participants from Baden-Württemberg have become used to the analogical variants of the imperative singular of strong verbs with \( e/i \)-gradation to such an extent that they have started processing them faster in all lower frequency verbs up to frequency class 10. Only in higher frequency base verbs do they have the traditional imperative variants so strongly entrenched that they take more time to process their analogical variants. The older participants from Baden-Württemberg similarly process traditional imperative variants of low frequency verbs slower than the traditional variants; nevertheless, they do not exhibit the same interaction effect as the younger generation. It seems that they are not familiarised with the analogical imperative variants in lower frequency verbs to the same extent as the younger participants have, which is why they only process these forms faster in base verbs.
with a frequency class below 19. Thus, the difference between the younger and older participants can be understood as further evidence of a change-in-progress in the imperative singular of strong verbs with e/i-gradation. The use of the analogical imperative singular variants of strong verbs with e/i-gradation has increased over time; therefore, the younger participants have encountered these forms more often and have become familiarised with them, which is reflected in their reading times as compared to those of older speakers. Participants from Saxony, by contrast, generally read the analogical imperative forms slower than the traditional (standard) variants with the stem vowel \( i \) because they generally have analogical imperative variants more weakly entrenched in their mental lexicons than the traditional imperative variants. This finding matches well with the developmental stage of analogical levelling in the strong verbs with e/i-gradation in Saxony documented in the dialectological literature, where no indication of change in the verb class can be found (cf. Section 1.4.2). In a stage of stable variation, which therefore seems to prevail in this area, it is to be expected that the traditional imperative singular variants with the stem vowel \( i \), which will be encountered and used more often, are more strongly entrenched than analogical imperatives which still represent the minority formation pattern for this verb class (cf. right panel of Figure 20).

Finally, it should be noted that regression models were fitted for spillover regions as well, testing the influence of the three predictors on residual log-transformed reading times of the two words following the stimulus imperatives. The analysis showed that the reading times on the first word following the target imperative are not determined anymore by an interactive effect of BVLTF, IMP.FORM and GROUP; in fact, the latter two predictors do not have a significant effect on these reading times at all. Only the predictor BVLTF has a significant effect, similar to that on reading times of the stimulus imperatives (\( p < 0.05 \)). That means, the exact same word occurring after the imperative of a low frequency base verb is read significantly slower than if it occurs after the imperative of a higher frequency base verb. In the analysis of the reading times of the second word following the stimulus imperative, none of the predictors is shown to have any significant effect anymore. Thus, reading times on the stimulus imperative itself are determined by a three-way interaction effect between base verb lemma token frequency, imperative form and participant group. This interaction effect does not spill over onto the reading times of the following word; only the main effect of base verb lemma token frequency is found to linger on in this region. By the time that the second word following the target imperative is read, the main effect of base verb lemma token frequency has vanished as well.
6.4.3 Recall accuracy

6.4.3.1 Results of the regression analysis

Observations for all imperative singular forms in the stimulus sentences were annotated for whether they were repeated as presented or not. The dataset contains observation for five factor levels of the dependent variable recall accuracy (REC.ACC), of which only instances annotated with levels “yes” or “no” were used in the present statistical analysis. The factor level “yes”, exact repetition of the presented stimulus imperative, was treated as the reference level. The factor level “no”, replacement of the presented stimulus imperative stem vowel by a different stem vowel (e by i and vice versa), is the response level of the logistic regression model. Random factors entered into the regression are the same as for the dependent variable reading time: random intercepts for stimulus verbs (1|BASE.VERB), stimulus sentences (1|ITEM.ID) and participants (1|PART), and participant-specific random slopes for the position of the stimulus in the experiment (1+POS.ITEM|PART), the position of the imperative in the stimulus sentence (1+POS.IMP|PART) and participants’ sensivity to word length (1+W.LEN|PART), base verb lemma token frequency (1+BVLTF|PART) and imperative form (IMP.FORM|PART). As for the dependent variable RESID.LOG.RT, separate regression models were fitted to the data for the three participant groups, in order to allow for group-specific recall patterns and to avoid overestimating interactive effects of predictors. In contrast to reading times, recall accuracies could not be residualised in order to assess the influence of base verb lemma token frequency irrespective of the effect of word length. Therefore, the variable word length (W.LEN) was included as a fixed factor in all regression models for REC.ACC, so that the share of variation in recall accuracies explained by this variable would be captured by this parameter.

As in the case of reading times, the final regression models for recall accuracies were arrived at in a forward selection process. The baseline regression models for all participant groups contained the model intercepts, the fixed factors stimulus position in experiment (POS.ITEM) and imperative position in stimulus sentence (POS.IMP), and the random factors. The influence of base verb lemma token frequency (BVLTF) and imperative form (IMP.FORM) was tested by adding the individual predictors and an interaction parameter to the baseline regressions and running log-likelihood comparisons. In the analysis of the data obtained from participants from Saxony, both predictors have a significant interactive effect on recall accuracy not unlike the sketch in Figure 25 (p < 0.001), whereas only IMP.FORM has a significant effect on recall accuracies obtained from participants from Baden-Württemberg (p < 0.001).
<table>
<thead>
<tr>
<th>variable</th>
<th>Saxony (18-30 years)</th>
<th>Baden-Württemberg (43-75 years)</th>
<th>Baden-Württemberg (18-30 years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>factor level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1.617 1.001</td>
<td>-1.104 1.270</td>
<td>-0.805 0.978</td>
</tr>
<tr>
<td>POS.IMP</td>
<td>-0.064 0.094</td>
<td>0.077 0.119</td>
<td>-0.171 0.090</td>
</tr>
<tr>
<td>POS.ITEM</td>
<td>0.007 0.011</td>
<td>0.015 0.015</td>
<td>0.006 0.012</td>
</tr>
<tr>
<td>W.LEN</td>
<td>0.009 0.131</td>
<td>-0.380 0.166</td>
<td>-0.138 0.126</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>variable</th>
<th>Saxony (18-30 years)</th>
<th>Baden-Württemberg (43-75 years)</th>
<th>Baden-Württemberg (18-30 years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BVLTF (numeric)</td>
<td>-0.227 0.071</td>
<td>-3.562 1.098</td>
<td>-18.991 2057.9</td>
</tr>
<tr>
<td>IMP.FORM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i</td>
<td>-8.153 3.426</td>
<td>-2.380 0.0173</td>
<td>-3.393 0.870</td>
</tr>
<tr>
<td>control</td>
<td>-21.806 305.58</td>
<td>-0.071 0.9431</td>
<td>-18.737 1716.5</td>
</tr>
<tr>
<td>BVLTF×IMP.FORM</td>
<td>0.230 0.208</td>
<td>1.107 0.2683</td>
<td>0.230 0.208</td>
</tr>
<tr>
<td>BVLTF×control</td>
<td>-0.159 232.65</td>
<td>-0.001 0.9995</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>random effects</th>
<th>variance standard deviation</th>
<th>variance standard deviation</th>
<th>variance standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1</td>
<td>ITEM.ID)</td>
<td>0.4839 0.6956</td>
<td>0.9717 0.9858</td>
</tr>
<tr>
<td>(1</td>
<td>BASE.VERB)</td>
<td>0.0000 0.0007</td>
<td>0.0003 0.0161</td>
</tr>
<tr>
<td>(1</td>
<td>PART)</td>
<td>0.0092 0.0958</td>
<td>0.0000 0.0066</td>
</tr>
<tr>
<td>(1+POS.IMP</td>
<td>PART)</td>
<td>0.0001 0.0094</td>
<td>0.0111 0.1051</td>
</tr>
<tr>
<td>(1+POS.ITEM</td>
<td>PART)</td>
<td>0.0000 0.0009</td>
<td>0.0010 0.0321</td>
</tr>
<tr>
<td>(1+W.LEN</td>
<td>PART)</td>
<td>0.0389 0.1973</td>
<td>0.1139 0.3375</td>
</tr>
<tr>
<td>(1+BVLTF</td>
<td>PART)</td>
<td>0.0020 0.0446</td>
<td>0.0000 0.0005</td>
</tr>
<tr>
<td>(1+IMP.FORM</td>
<td>PART) i</td>
<td>9.4454 3.0733</td>
<td>6.9927 2.6440</td>
</tr>
<tr>
<td>(1+IMP.FORM</td>
<td>PART) control</td>
<td>1.0123 1.0061</td>
<td>0.7965 0.8924</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>measures of model quality</th>
<th>Saxony (18-30 years)</th>
<th>Baden-Württemberg (43-75 years)</th>
<th>Baden-Württemberg (18-30 years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>800</td>
<td>663</td>
<td>807</td>
</tr>
<tr>
<td>AIC (df)</td>
<td>358.586 (31)</td>
<td>257.4613 (28)</td>
<td>299.9367 (28)</td>
</tr>
<tr>
<td>marginal R2</td>
<td>0.8723</td>
<td>0.9240</td>
<td>0.9314</td>
</tr>
<tr>
<td>conditional R2</td>
<td>0.9013</td>
<td>0.9564</td>
<td>0.9543</td>
</tr>
</tbody>
</table>

Notes: control = control verb, e = imperative singular of strong verb with e/i-gradation with stem vowel e, i = imperative singular of strong verb with e/i-gradation with stem vowel i
The results of the analysis are listed in Table 33 and illustrated in Figure 31. The regression models reported in Table 33 are corrected for outliers (manually; cf. Appendix B), whereby 0.25% of the data points were removed from the analysis of recall accuracies for participants from Saxony, 1.19% from the analysis for older participants from Baden-Württemberg, and 0.37% for the younger participants from Baden-Württemberg.

![Figure 31: Interactive effect of base verb lemma token frequency and imperative form on recall accuracies of participants from Saxony and effect of imperative form on recall accuracies of participants from Baden-Württemberg](image)

In Figure 31, the y-axis represents the dependent variable, i.e. the probability for a stimulus imperative to be repeated as presented (lower end) or replaced by an imperative with a different stem vowel (e replaced by i, or vice versa; upper end). Such a replacement is not possible in the case of the control verbs, which only ever have one stem vowel; therefore, all data points for this category are found at the 0 percent replacement mark in all participant groups. Participants from Saxony generally repeat stimulus imperatives with the stem vowel i; if they replace the stem vowel of these imperatives by the analogical stem vowel e in recall, this only happens in very low frequent verbs. When this group is shown imperatives were with the stem vowel e, however, they replace them with a very low probability in low frequency verbs, but this probability increases with base verb lemma token frequency to a maximum of 58.42%. In contrast, in the panels for participants from Baden-Württemberg, the factor IMP.FORM has been moved to the x-axis instead of being represented by different linetypes. This reflects the fact that their recall accuracies are not significantly affected by the token frequency of the verb whose imperative occurred in the stimulus sentence. Recall accuracies for imperative stimuli with the stem vowel i obtained from participants from this region
are almost 100 percent, but more importantly, both groups also recall stimulus imperatives with the stem vowel e with a very high probability (older: 90.96 %, younger: 89.69 %). The stark contrast between recall accuracies of imperative singular of strong verbs with e/i-gradation obtained from participants from Saxony and Baden-Württemberg suggests more strongly that the two regions in Germany instantiate different stages of the analogical leveling process affecting this verb group.

6.4.3.2 Discussion of the recall results

Regression models fitted on subsets of the data for participant groups have shown that, whereas base verb lemma token frequency (BVLTF) and imperative form (IMP.FORM) and an interaction of both factors (BVLTF×IMP.FORM) have a significant impact on recall accuracies of participants from Saxony, only the factor imperative form (IMP.FORM) has a significant influence on whether a presented imperative singular form is repeated as presented by participants from Baden-Württemberg.

The fact that base verb lemma token frequency does or does not significantly influence participants’ recall accuracies does not preclude the possibility that verb lemma token frequency (VLTF) might not affect them as well. Variants of the regression analyses reported above were fitted, in which the fixed factors BVLTF and WORD.LENGTH were replaced by VLTF and LEMMA.LENGTH. Random intercepts were included for (1|VERB) instead of (1|BASE.VERB) and a random slope for (1+LEMA.LENGTH|PART), participants’ sensitivity to lemma length, instead of (1+WORD.LENGTH|PART). Log-likelihood comparisons confirm the expectation that verb lemma token frequency is a worse predictor of recall accuracies than base verb lemma token frequency: when BVLTF has a significant effect, as in the analysis of data from participants from Saxony, VLTF does not feature significantly, and when the influence of neither BVLTF nor VLTF reaches significance, the values for measures of model quality in the VLTF versions of regression models are constantly worse than those for the BVLTF versions.

It may be worth mentioning that the analysis also tested whether recall accuracies were affected by whether the respective participants were (self-proclaimed) dialect speakers (variable DIALECT) or not. However, the variable does not have a significant effect on recall accuracies of imperatives of any of the participant groups; thus, dialect use or knowledge does not seem to affect the ability to recall traditional or analogical imperative singular forms of strong verbs with e/i-gradation or the tendency to replace some of these variants.
No apparent-time change in the processing of the traditional and analogical imperative singular variants can be deduced from the data obtained from the two generations of participants from Baden-Württemberg. Analogical imperative stimuli with the stem vowel e are recalled equally well by both groups; this implies that even though these two groups differ in their online processing of the forms (cf. Section 6.4.2), they seem to be accustomed to the analogical forms to a sufficient and similar extent and, therefore, have little problems memorising them. Unsurprisingly, imperative forms with the stem vowel i are recalled with almost 100 percent probability throughout the frequency continuum by participants from Saxony as well. However, in stark contrast to the groups from Baden-Württemberg, recall accuracies for analogical imperatives with the stem vowel e obtained from participants from Saxony crucially depend on the token frequency of the base verb in which they are presented: the probability for exact repetition of these forms decreases from 97.41 % in the stimulus verbs with the lowest frequency to only 41.58 % in those stimulus verbs with the highest frequency. This finding seems to be another instance of a conserving effect of high token frequency: Analogically formed imperatives can enter the memory of participants from Saxony in low frequency verbs; as base verb lemma token frequency increases, they are (wittingly or unwittingly) rejected and replaced by imperative variants with the stem vowel i.

The arrangement of panels in Figure 31 might again represent a chronological order: At all times are the traditional imperative singular variants of strong verbs with e/i-gradation memorised with nearly 100 percent probability by speakers. Analogical imperatives with the stem vowel e, by contrast, have a higher potential of being repeated among lower frequency verbs at first (represented by the data from speakers from Saxony); some time later, they will show higher recall accuracies also in higher frequency base verbs (a hypothetical stage not represented in the data from the present study) until at last, they show probabilities for repetition almost equalling those of imperatives with the stem vowel i, independent of the token frequency of the presented stimulus verb (represented by the data from participants from Baden-Württemberg).

6.5 Results for participants who detected the aim of the study

In other experiments, the data from participants who detect the aim of the respective study might be considered “wasted data” because the participants behave very differently from what was intended by the experimenter. In the present study, however, these data are not excluded from the analysis altogether: since these participants do not (as intended) subconsciously process and perceive the presented imperative singular variants of strong verbs with e/i-
gradation, their reactions can be considered tokens of “monitored” change and thus may reveal some additional interesting insights about the change-in-progress under investigation. This group of participants consists of 13 younger participants from Baden-Württemberg, 2 older participants from Baden-Württemberg and 18 younger participants from Saxony (cf. Section 6.3.2).

6.5.1 Who detected the cover story?
Logistic regression models (function glm() in R base installation) tested why some of the participants detected the aim of the experimental study and others did not, in particular which of the sociolinguistic variables and possibly factors related to the experiment design had an influence on this question. The analysis is based on the dataframe with meta data from the background questionnaires and the debriefing sessions (cf. Sections 6.3.4 and 6.3.5). The dependent variable in the regression is DETECTED with the reference level “no” (participant did not detect the cover story) and the response level “yes” (participant detected the aim of the study). The impact of a number of potentially relevant predictor variables was tested by adding them to a baseline regression model, which contained only the dependent variable and the model intercept, and comparing the baseline and the subsequent regression in log-likelihood tests.

Three of the potential predictors did not have a significant effect on the question whether or not participants detected the aim of the study. The first of these has to do with the experiment design. As explained in 6.3.1, a Latin Square distribution was used in the present study: Participants are assigned to lists which determine which of the stimulus sentences they are presented in which condition. The variable LIST (6 factor levels) was added as a fixed factor to the regression model, chiefly as a test for a defective experiment design. It was found not to significantly influence variation in the dependent variable; thus, the experiment design itself, or more precisely a specific participant list, does not lower or raise chances for any of the participants to detect the aim of the study. The regression model for the second predictor PLACE tested whether participants from Baden-Württemberg or Saxony are generally more likely to be deceived by the cover story or not. This could be expected from the observation that 39 percent of the Saxonian participants (18 out of 46), but only 22 percent of participants from Baden-Württemberg (13 out of 44), detected the cover story. The analysis has shown, however, that this difference is not statistically significant. On the one hand, this is a result for one of the sociolinguistic variables; on the other hand, in line with the result for LIST, it disproves that there may have been problems in the way the experiment was designed and conducted. The fact that participants from different regions do not exhibit significantly different
chances for detecting the aim of the study serves to show that they did not adopt a fundamentally different strategy while completing the test. The third predictor with no significant effect in the regression models is the variable GENDER: The question whether participants are deceived by the cover story or not does not depend on whether they are female or male.

Two of the sociolinguistic variables had a significant effect on the dependent variable DETECTED: participants’ AGE and the question whether they are DIALECT speakers. The probability with which participants detect the aim of the experiment decreases significantly with their AGE: the youngest participants at 18 years of age show a probability of 37.4 percent for detecting the cover story, which decreases to 5.6 percent for the oldest participant at 75 years of age (p < 0.05).

![Figure 32: Effects of participants’ AGE and GENDER on dependent variable DETECTED](image)

The reason for this finding can be found in the self-reports of some of the younger participants that they wanted to find out what was tested in the experiment, whereas many of the older participants reported that they were busy with the tasks themselves and therefore did not concentrate on detecting the aim of the test.

Participants who identified as dialect speakers in the background questionnaire show a significantly lower probability for detecting the cover story (15.8 %) than participants who do not consider themselves dialect speakers (37.8 %, p < 0.01). This finding might be a reflection of the generally higher tendency towards regularisation in non-standard varieties (cf. Chambers 2003; 2004); dialect speakers are therefore more accustomed to non-standard forms. There is no significant interaction effect of the predictors AGE and DIALECT on the dependent variable DETECTED, but the effects of both add up (illustrated in Figure 32): the highest probability for detecting the aim of the study is thus found for the youngest participants who
are not dialect speakers (51.8 percent), and the lowest probability is found in the oldest participants who identified as dialect speakers (3.2 percent).

### 6.5.2 Reading times

Mixed-effects regression models for the dependent variable reading time (RT) were fitted to the data of the participants who detected the aim of the study in much the same way as explained in Section 6.4.2 for participants who did not detect the cover story: log-transformed reading times (LOG.RT) were residualised on word length; the residuals (RESID.LOG.RT) were used as the dependent variable in the subsequent regression analyses. Random intercepts and slopes are the same, and a baseline regression model was fitted, which contains the dependent variable, the random variables, the model intercept, and the fixed factors POS.IMP and POS.ITEM. The fixed factors whose influence was tested are base verb lemma token frequency (BVLTF), imperative form (IMP.FORM), participant group (GROUP), the question whether participants had detected the aim of the study by the time the particular stimulus was presented (WHEN.DETECTED) and the sociolinguistic variables AGE and GENDER. They were added to the baseline regression one at a time, and the resulting regression models were compared to the baseline in log-likelihood tests.

Comparisons showed that none of the sociolinguistic variables had a significant effect on the variation of the dependent variable RESID.LOG.RT, and reading times do no differ significantly between the three participant GROUPs. Apart from these factors, an effect on reading times was also expected for the variable WHEN.DETECTED: as explained in Section 6.3.5, participants may (wittingly or unwittingly) change their “strategy” during the experiment once they have detected the aim of the study and read analogical imperatives faster or slower than before, when they read the stimulus sentences more subconsciously. However, participants’ reading times were not significantly different before and after having recognised that reactions to imperative singular forms were tested in the experiment. Perhaps even more surprising is the finding that the variable imperative form (IMP.FORM) did not significantly affect variation in the dependent variable, either. Thus, even though some of the participants who detected the aim of the study stated that after having recognised the aim of the study, they paid more attention to imperative forms in general, and to imperative forms with the stem vowel e in particular, this is not reflected in their reading times.
Table 34: Results of regression analysis of residual reading times of participants who detected the aim of the study

<table>
<thead>
<tr>
<th>variable</th>
<th>factor level</th>
<th>estimate</th>
<th>standard error</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>- (numeric)</td>
<td>-0.042</td>
<td>0.087</td>
<td>-0.480</td>
<td>0.6328</td>
</tr>
<tr>
<td>POS.IMP</td>
<td>- (numeric)</td>
<td>0.025</td>
<td>0.011</td>
<td>2.338</td>
<td>0.0220 *</td>
</tr>
<tr>
<td>POS.ITEM</td>
<td>- (numeric)</td>
<td>-0.004</td>
<td>0.001</td>
<td>-4.326</td>
<td>0.0000 ***</td>
</tr>
<tr>
<td>BVLTF</td>
<td>- (numeric)</td>
<td>0.013</td>
<td>0.005</td>
<td>2.680</td>
<td>0.0091 **</td>
</tr>
</tbody>
</table>

random variables

<table>
<thead>
<tr>
<th>variance standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1</td>
</tr>
<tr>
<td>(1</td>
</tr>
<tr>
<td>(1</td>
</tr>
<tr>
<td>(1+POS.ITEM</td>
</tr>
<tr>
<td>(1+POS.IMP</td>
</tr>
<tr>
<td>(1+WORD.LENGTH</td>
</tr>
<tr>
<td>(1+BVLTF</td>
</tr>
<tr>
<td>(1+IMP.FORM</td>
</tr>
<tr>
<td>(1+IMP.FORM</td>
</tr>
</tbody>
</table>

measures of model quality

| N        | 958      |
| AIC (df) | 887.20 (28) |
| R² (marginal) | 0.0712 |
| R² (conditional) | 0.3026 |

Note: The regression model reported in this table is corrected for outliers; this procedure removed 5.32% of the original 1020 data points.

In fact, apart from a familiarisation effect across the experiment in the same magnitude as exhibited by participants who did not detect the aim of the study (captured by the predictor POS.ITEM) and a significant increase in reading times, the later the imperative form occurred in the stimulus sentence (POS.IMP), the only focus predictor which has a significant effect on the dependent variable reading time is the fixed factor base verb lemma token frequency (BVLTF). As displayed in Table 34, imperative forms of base verbs in the highest DeReWo frequency classes are read fastest; as the base verb lemma token frequency class in the underlying DeReWo list increases, which means that base verb lemma token frequency decreases, reading times increase.

6.5.3 Recall accuracy

Mixed-effects logistic regression models were fitted on the recall accuracy data of participants who detected the aim of the study as explained in Section 6.4.3 for participants who did not detect the cover story, i.e. using only the factor levels “yes” (exact repetition of the stimulus imperative, reference level) and “no” (replacement of the stem vowel in the stimulus impera-

---

37 Kenward-Roger estimates (see Appendix A)
ative, response level). Random effects were entered for the same variables as in regression models for the dependent variable reading time (6.5.2).

Table 35: Results of regression analysis of recall accuracies of participants who detected the aim of the study

<table>
<thead>
<tr>
<th>variable</th>
<th>factor level</th>
<th>estimate</th>
<th>standard error</th>
<th>z-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>- (numeric)</td>
<td>0.470</td>
<td>1.055</td>
<td>0.445</td>
<td>0.6562</td>
</tr>
<tr>
<td>POS.IMP</td>
<td>- (numeric)</td>
<td>0.032</td>
<td>0.097</td>
<td>0.332</td>
<td>0.7401</td>
</tr>
<tr>
<td>POS.ITEM</td>
<td>- (numeric)</td>
<td>-0.013</td>
<td>0.014</td>
<td>-0.941</td>
<td>0.3466</td>
</tr>
<tr>
<td>WORD.LENGTH</td>
<td>- (numeric)</td>
<td>-0.027</td>
<td>0.126</td>
<td>-0.213</td>
<td>0.8313</td>
</tr>
<tr>
<td>IMP.FORM e</td>
<td>reference level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i</td>
<td></td>
<td>-6.075</td>
<td>1.360</td>
<td>-4.466</td>
<td>0.0000 ***</td>
</tr>
<tr>
<td>control</td>
<td>reference level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sax</td>
<td></td>
<td>-20.201</td>
<td>357.941</td>
<td>-0.056</td>
<td>0.9550</td>
</tr>
<tr>
<td>bawü.old</td>
<td>reference level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bawü.young</td>
<td></td>
<td>-21.930</td>
<td>407.785</td>
<td>-0.054</td>
<td>0.9571</td>
</tr>
<tr>
<td>GENDER female</td>
<td>reference level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>male</td>
<td></td>
<td>-1.611</td>
<td>0.709</td>
<td>-2.271</td>
<td>0.0232 *</td>
</tr>
</tbody>
</table>

random variables

<table>
<thead>
<tr>
<th>variance</th>
<th>standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1</td>
<td>BASE.VERB)</td>
</tr>
<tr>
<td>(1</td>
<td>ITEM.ID)</td>
</tr>
<tr>
<td>(1</td>
<td>PART)</td>
</tr>
<tr>
<td>(1+POS.ITEM</td>
<td>PART)</td>
</tr>
<tr>
<td>(1+POS.IMP</td>
<td>PART)</td>
</tr>
<tr>
<td>(1+WORD.LENGTH</td>
<td>PART)</td>
</tr>
<tr>
<td>(1+BVLTF</td>
<td>PART)</td>
</tr>
<tr>
<td>(1+IMP.FORM</td>
<td>PART) i</td>
</tr>
<tr>
<td>(1+IMP.FORM</td>
<td>PART) control</td>
</tr>
</tbody>
</table>

measures of model quality

| N | 909 |
| AIC (df) | 394.8028 (31) |
| R^2 marginal | 0.9183 |
| R^2 conditional | 0.9701 |

Note: sax = Saxony (18-30 years old), bawü.old = Baden-Württemberg (43-75 years old), bawü.young = Baden-Württemberg (18-30 years old)

The baseline regression model contains the dependent variable, the random variables, the model intercept, and the fixed factors POS.IMP, POS.ITEM and WORD.LENGTH. The fixed factors whose influence was tested are base verb lemma token frequency (BVLTF), imperative form (IMP.FORM), participant group (GROUP), the question whether participants had detected the aim of the study by the time the particular stimulus was presented (WHEN.DETECTED) and the sociolinguistic variables AGE and GENDER. They were added to the baseline regression one at a time, and the resulting regression models were compared to the baseline in log-likelihood tests.
Similarly as in the data obtained from participants who detected the cover story of the experiment, recall accuracies of participants who did detect the aim of the study are significantly influenced by the imperative form (IMP.FORM) in the stimulus sentence. Imperative singular forms of control verbs are recalled to 100 percent because a replacement of the imperative stem vowel is not possible, and stimulus imperatives with the stem vowel i are also recalled with 99.8 % probability. Imperative forms with the stem vowel e, on the other hand, are recalled with a probability of only 19.8 % on average, i.e. they are replaced by forms with the stem vowel i with a probability of 80.2 %. In addition, the predictor GENDER was found to have a significant effect on the dependent variable: male participants generally show a higher probability of repeating the presented imperative forms with the stem vowel e (98.8 %) than female participants (94.3 %; p < 0.001). The last predictor with a significant effect on recall accuracies identified in the regression model is the factor participant group (GROUP). Participants from Saxony tend to replace more stimulus imperative stems than participants in the same age group (i.e. the younger participants) from Baden-Württemberg (6.2 % and 2.2 %, respectively). Older participants from Baden-Württemberg recall stimulus imperatives with (almost) 100 percent probability in all three forms; note, however, that there are only two participants in this group. Due to this very small number of participants, their behaviour cannot necessarily be generalised to all older speakers from Baden-Württemberg. The three predictors do not have an interactive effect on the dependent variable; their effects add up as illustrated in Figure 33:

![Figure 33](image-url)

**Figure 33:** Effects of imperative form, participant gender and group on recall accuracies of participants who detected the aim of the study

Note: e = stem vowel e, i = stem vowel i, control = control verb
The regression analysis revealed no significant effect of the predictors WHEN.DETECTED, BVLTF and AGE on recall accuracies of participants who detected the aim of the study. Unexpectedly thus, participants do not show higher or lower recall accuracy rates after having recognised the aim of the study, i.e. they do not change their strategy and simply repeat or replace all analogical imperative stimuli from then on. Recall accuracies of participants from Baden-Württemberg who did not detect the cover story were not affected by the factor base verb lemma token frequency (BVLTF). Likewise, participants who did detect the aim of the study recall imperative forms equally well, regardless of the stimulus verb’s token frequency. Furthermore, older participants in the latter group do not behave significantly differently in the recall task from younger participants. This finding can be explained by the small number of older participants from Baden-Württemberg in this group, as significance of contrasts hinges critically on sample size.

6.5.4 Discussion of results for participants who were not deceived by the cover story

Perhaps the most surprising finding from the analysis of the data from participants who detected the aim of the study is that they do not read or recall the stimulus imperatives significantly differently after having detected that the aim of the experiment was to measure reactions to imperative forms. Part of the reason for this may be found in the fact that most of them detected the cover story at an early point during the experiment (on average after having read 20 percent of the stimulus sentences); thus, the small sample size of stimulus sentences read before detection may have caused the high significance value of the variable WHEN.DETECTED in the regression models.

Apart from this, the regression analysis has revealed big differences in the set of variables which affect the reading times and recall accuracies of imperative singular variants of strong verbs with e/i-gradation of speakers who know the aim of the study and unsuspecting participants, who did not detect the cover story, in the same participant groups. In contrast to the reading times of participants who did not detect the aim of the study, which are dependent on the form of the stimulus imperative (variable IMP.FORM), the reading times of participants who detected the cover story are not significantly affected by the factor IMP.FORM. This finding can be explained as a consequence of the fact that these participants detected the aim of the study. They realise at an early point during the experiment that their reactions to imperatives are tested and that these imperatives can occur in two variants: one with the stem vowel i and one with the stem vowel e. Since both imperative variants occur with the same frequency in the experiment, for the participants who detected this, both variants have an equal probability of occurring in the stimulus sentences, instead of the frequency and prob-
ability constellations revealed by the Walkthrough Corpus study. Therefore, processing times
do not differ between traditional and analogical imperatives. Instead, the processing of the
stimulus imperatives by participants who know the aim of the experiment is only affected by
the time required to access the verb in the mental lexicon in which they are presented, which
is why only the effect of the factor base verb lemma token frequency (BVLTF) reaches sig-
ificance in the respective regression model (cf. 6.5.2).

The variables influencing recall accuracies are more similar for all participants of the
experiment. Regression analyses of both data sets revealed significant differences between
participant groups on the one hand; on the other hand, an important predictor of recall accu-
ricy in all groups is the form of the stimulus imperative (IMP.FORM). Unlike the participants
who did not detect the aim of the study, female and male participants who detected the aim
recall the stimulus imperative forms significantly differently. The higher tendency for female
speakers to replace the presented analogical stem vowel $e$ in imperative forms by the tradi-
tional variants with the stem vowel $i$ in comparison to male participants is reminiscent of find-
ings from studies conducted by William Labov and colleagues and students which are sum-
mamed in his “principles of sexual differentiation” in linguistic change (Labov 1990: 205; cf.
1.4.3). Participants who detected the cover story of the present experiment react to variants of
a linguistic variable in much the same way as speakers have been found to use them in the
sociolinguistic studies quoted by Labov (1990; 2001): in stable linguistic variation or change
from above the level of consciousness, i.e. when speakers are aware of the variation or
change, “women use more standard forms, responding to the overt prestige associated with
them” (Labov 1990: 210). The participants in the current study who are aware of the fact that
their reactions to imperatives are tested, are also aware that the variant of the imperative sin-
gular of strong verbs with e/i-gradation with the stem vowel $i$ is the traditional standard vari-
ant. In the light of Labov’s principles, it is thus not surprising to see female participants re-
place more of the presented analogical imperative forms by the standard variant than the male
participants do. As concerns the influence of the variable participant group (GROUP) on
recall accuracies, it was explained above (cf. 6.5.3) that the results for older participants are
negligible because of the small number of participants in this group. The significant differ-
ence between recall accuracies of younger participants from Saxony and Baden-Württemberg
is in line with the results for participants who did not detect the aim of the study: Analogical
levelling in the imperative singular of strong verbs with e/i-gradation seems to be at work in
Baden-Württemberg but not to have reached Saxony, yet. Therefore, participants from

---
38 A discussion of possible reasons for this “careful behavior of women” can be found in Labov (2001: 275-279).
Saxony are less accustomed to the analogical formation of these imperatives and thus generally replace more of the forms with the stem vowel \( e \) presented in the experiment by variants with the stem vowel \( i \) during recall.

6.6 Discussion of the results of the experiment

Previous usage-based research has found a conserving effect of high token frequency in analogical change which is corroborated in the analysis of the Walkthrough Corpus; this conserving effect is generally assumed to be a reflection of the mental entrenchment of forms. The present experiment endeavoured to find evidence for this explanation in terms of reactions to traditional and analogical imperative singular variants of higher and lower frequency strong verbs with \( e/i \)-gradation by native speakers of German.

The conserving effect holds that an irregular inflection is preserved in high frequency verbs because the respective high frequency irregular forms are highly entrenched in the mental lexicons of speakers, owing to their high frequency of use. Assuming that strongly entrenched linguistic items are more easily accessed in the mental lexicon, high frequency irregular forms like \textit{nimm} should be processed faster by native speakers of German than all other frequency and irregularity constellations, i.e. analogical variants of high frequency verbs (e.g. \textit{nehm(e)}) and analogical and irregular variants of low frequency verbs (e.g. \textit{zertrett(e)} and \textit{zertritt}). This hypothesis was borne out by reading time data obtained from participants who did not detect the aim of the study (cf. Figure 30): the traditional irregular imperative variants of higher frequency verbs are the imperative stimuli which are read fastest, which shows that they are most strongly entrenched in the speakers’ minds. In terms of the second measure of entrenchment, recall accuracies, the data obtained from participants from Saxony provide further evidence in this direction (cf. Figure 31): these participants only recall the irregular imperative singular forms to 100 percent; in all other conditions, the probability for replacement of the presented stem vowel of the stimulus imperative is higher.

The conserving effect also entails that irregular imperative variants of lower frequency verbs and analogical variants of high frequency verbs, which occur with a lower frequency of occurrence in the Walkthrough Corpus, should be less strongly entrenched than the irregular variants of the high frequency verbs; otherwise, the former would resist analogical levelling as strongly as the latter variants. Again, this hypothesis is confirmed by the reading time data obtained from participants from Baden-Württemberg. The analogical imperative singular variants of high frequency verbs like \textit{nehm(e)} as well as the irregular variants of lower frequency verbs like \textit{zertritt} are read slower than the traditional irregular imperative singular
forms of high frequency verbs (e.g. *nimmt*); that means, they are less strongly entrenched in speakers’ minds (cf. Figure 30). As mentioned above, recall accuracies of analogical imperatives of high frequency verbs and irregular imperatives of low frequency verbs obtained from participants from Saxony are (much) lower than those for the traditional irregular imperative forms of high frequency verbs and thus are indicative of the validity of the hypothesis as well.

Importantly, however, on the basis of their relatively high frequency of occurrence in the Walkthrough Corpus, also the analogical imperative singular variants of lower frequency strong verbs with *e/i*-gradation such as *zertret(e)* would be expected to be to some extent entrenched in speakers’ minds. Therefore, they were assumed to be processed faster in the experiment by native speakers of German than the less frequently occurring analogical imperative variants of higher frequency verbs like *nehmen* and irregular imperative variants of lower frequency verbs with the stem vowel *i* like *zerritt*. In fact, participants from Baden-Württemberg read the analogical imperative variants of lower frequency verbs faster than the irregular variants of the same verbs and not significantly slower than the analogical imperative variants of higher frequency verbs (cf. Figure 30). And the recall accuracies of analogical imperative singular variants of low frequency verbs obtained from participants from Saxony are much higher than those of analogical variants of high frequency verbs and not significantly different from those of the standard irregular imperative variants of low frequency verbs. This provides further evidence that the analogical imperative variants of low frequency verbs like *zertret(e)* are indeed more strongly entrenched than the analogical variants of high frequency verbs and irregular variants of low frequency verbs but still less strongly than the traditional irregular imperative variants of high frequency verbs like *nimmt*. This latter finding does not provide evidence for the explanation of the conserving effect on the basis of entrenchment, but it highlights the validity of measuring processing latencies as a reflection of entrenchment induced by higher or lower frequency of occurrence.

It must be noted that the discussion so far is a simplification of the results: entrenchment is assumed to be a gradual phenomenon, just as the frequency of occurrence of variants in the Walkthrough Corpus is gradual. Thus, as indicated in the sketches of expected results in Figures 23 and 25, processing differences as measured by reading times and recall accuracies should also be gradual. Since data points obtained from the experiment were annotated for the original DeReWo frequency class of the stimulus verb, this makes it possible to observe that in the majority of cases, reading times and recall accuracies are in fact decreasing or increasing gradually (exceptions are recall accuracies from participants who detected the aim of the
study and those from participants from Baden-Württemberg who did not detect the cover story).

A few words must also be said on the reading times of participants from Saxony and recall accuracies of participants from Baden-Württemberg because they do not match the expected patterns. Instead of displaying the expected cross-over effect of imperative singular variants with the stem vowels e and i of higher and lower frequency strong verbs with e/i-gradation (Figure 26), participants from Saxony consistently read the analogical imperative singular variants with the stem vowel e faster than the irregular variants with the stem vowel i. As mentioned in Section 6.4.2.3, this finding can be explained as a sign that analogical leveling in the imperative singular of strong verbs with e/i-gradation has not yet reached the stage of change in Saxony; therefore, the analogical variants of these verbs are consistently less strongly entrenched in the mental lexicons of this participant group than the traditional imperative variants with the stem vowel e. By contrast, the change is in effect in Baden-Württemberg, and it is ongoing: reading times show that younger participants react less conservatively towards analogical imperative singular forms of strong verbs with e/i-gradation than speakers from an older generation (cf. 6.4.2.3). Both groups from this region, however, seem to have grown sufficiently accustomed to the analogical variants to recall them irrespective of the frequency of the verb in which they occur.

When assessing whether the results of the experimental study corroborate or speak against the hypotheses and expectations outlined in the introductory sections of this chapter, a few more points need to be kept in mind. Firstly, the sketch of expected reading times in Figure 20 tried to take into account that verbs with a higher token frequency are in general read faster than verbs with a low token frequency; this effect could not, self-evidently, be estimated precisely before the experiment was conducted. The actual residual reading times obtained during the experiment show that the main effect of verb frequency was overestimated; the difference between reading times for higher and lower frequency is smaller than sketched in the figure; this may partly explain why the cross-over between imperative singular variants with the stem vowels i and e does not stand out as clearly as expected. Secondly, with regard to recall accuracies, the expected results are largely in line with the findings for participants from Saxony; however, particularly the recall of analogical imperative singular variants is most probably higher in the artificial situation created by the experiment setting than it would be in natural language use. In contrast to the on-line task of reading stimulus sentences, the off-line task of repeating them is most certainly driven by desirability considerations, inflicted or at least reinforced by the use of the verb ‘repeat’ in the instructions for the experimental
test ("Bitte wiederholen Sie den Satz"). The reason for using this instruction, as opposed to others like "Geben Sie den Satz wieder" ‘reproduce the sentence’, was to prevent participants from paraphrasing the content of the stimulus sentence in the form of infinitive, indicative or modal constructions. Finally, although the results for participants who detected the cover story do not match the expected outcome, this does not call into question the validity of the explanation of frequency effects in analogical change on the basis of entrenchment. These participants were aware of the fact that imperative singular variants of strong verbs with e/i-gradation occur with an equal probability with the two stem vowels \(e \) and \(i \). This explains why both variants are processed very similarly; therefore, reading times obtained from speakers who knew the aim of the study were only significantly influenced by the base verb lemma token frequency of the verb in which they are presented because higher frequency are accessed faster in the mental lexicon than lower frequency verbs. Their awareness of the aim of the experimental study also explains why they repeat or replace stimulus imperatives regardless of the frequency of the verb in which they are presented.

6.7 Summary

Perhaps because the explanation of the conserving effect of high token frequency in analogical change is generally accepted by the majority of linguists, the present study constitutes the first systematic test of the psychological or cognitive reality of the link between corpus frequencies and mental entrenchment in this context. To this end, processing latencies are applied as a measure of entrenchment in a “self-paced reading with recall” experiment. This means that participants are presented with analogical and irregular imperative singular variants of strong verbs with e/i-gradation in self-paced reading, after which they are asked to reproduce the respective stimulus sentence.

The newly-introduced off-line measure of entrenchment in terms of recall accuracies is found to be less conservative because it is more strongly affected by conscious reflection by speakers than the on-line measure of reading times. Nevertheless, the data obtained from the experiment support the explanation of the resistance of high frequency irregular forms to analogical levelling on the basis of their high mental entrenchment: the traditional irregular imperative singular variants of high frequency strong verbs with e/i-gradation are most strongly entrenched in speakers’ mental lexicons, reflected by very low reading times and 100 percent recall by all participants. As is expected from the findings of previous usage-based studies, the less frequently occurring analogical imperative variants of high frequency verbs and traditional irregular variants of low frequency verbs are found to cause higher reading times and
lower recall accuracies and are thus less strongly entrenched in the mental lexicons of speakers than the traditional irregular variants of high frequency verbs. The analysis of the experiment data results in a final observation which was expected from frequency distributions in the Walkthrough Corpus but not necessarily from previous usage-based research: the analogical imperative singular variants of low frequency strong verbs with e/i-gradation are processed faster and recalled more often than the irregular variants of the same verbs and the analogical variants of high frequency verbs; this shows that they are in the process of becoming entrenched in the mental lexicons of contemporary speakers.

The present study also incorporated an apparent-time design, a method which is best known from sociolinguistic studies of variation (6.1.3). The comparison of processing and perception data from two generations of speakers provides evidence for the classification of the analogical levelling process in the imperative singular of strong verbs with e/i-gradation as a change-in-progress across a longer stretch of time than was possible in the Walkthrough Corpus study (Chapter 4). Moreover, data are compared which were obtained from participants from two rather distant regions within Germany (Freiburg im Breisgau, Baden-Württemberg and Leipzig in Saxony); this method proves that the collection of data from speakers of different regional backgrounds can work as a window on the past also in behavioural experiments: analogical change in the imperative singular of strong verbs with e/i-gradation is in effect in the Low Alemannic region of Baden-Württemberg, but it does not seem to have reached Saxony yet.

In summary, the results of the present experiment show how analogical formations can spread through a community: in the first stage, the traditional irregular formation is more entrenched in the mental lexicon of speakers than analogical variants of the same variable, as reflected by the reading times of imperative singular forms of strong verbs with e/i-gradation obtained from participants from Saxony (Figure 30). Nevertheless, occasional uses of analogical variants in low frequency verbs tend to go unnoticed (they are not “corrected”) or may even be repeated by speakers in these verbs, while they are replaced by their irregular counterparts in high frequency verbs. This is instantiated in recall accuracies obtained from participants from Saxony in Figure 31. At a later point in time, once the analogical variants of low frequency verbs have spread to some extent, speakers have become accustomed to these forms; this stage is captured in the reading times obtained from speakers from Baden-Württemberg who process analogical imperative singular variants of low frequency verbs faster than the supposedly correct irregular variants of these verbs (Figure 30). At this stage or some time later, analogical variants have become sufficiently entrenched in low frequency
verbs to qualify as equally eligible realisations of the construction; therefore, participants from Baden-Württemberg recall both imperative singular variants to a similar extent irrespective of the frequency of the verb in which they are presented (cf. Figure 31). Thus, the way is paved for them to spread in use also in higher frequency verbs. The analysis of data from participants who detected the aim of the study shows that this spread probably does not proceed in an unrestrained manner; women and speakers from regions in which the change is not in effect may work as a “buffer” to the development (cf. the quote by Tschirch (1969) in Section 1.2.1)
Conservation of *i* in the second and third person singular

The corpus studies in Chapters 4 and 5 have examined frequency effects involved in the direction and the trajectory of analogical levelling in the imperative singular of strong verbs with *e/i*-gradation. A question which remains to be answered is why only the irregular form of the imperative singular is changing towards an analogical formation with the stem vowel *e* at the time of this study, when the second and third person singular present indicative of these verbs exhibit the same stem vowel alternation. As an illustration, Table 36 presents the mean proportions of analogical variants of the three paradigm forms in the corpora used in the present study (see 7.3). In two of the corpora, analogical variants account for two thirds of the imperative singular forms of strong verbs with *e/i*-gradation, whereas the proportions of analogical variants of the second and third person singular present indicative do not exceed five percent. In the third corpus, the proportion of analogical imperative singular variants is lower; at the same time, only one instance of an analogical variant of the second person singular present indicative of the verb *sehen* ‘see’ is found, accounting for 0.06% of instances of this paradigm form.

<table>
<thead>
<tr>
<th>corpus</th>
<th>mean proportion of analogical variants (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>imp sg</td>
</tr>
<tr>
<td>moca database</td>
<td>33.27</td>
</tr>
<tr>
<td>FOLK corpus</td>
<td>33.45</td>
</tr>
<tr>
<td>Wortschatz Universität Leipzig</td>
<td>7.01 (0.06)*</td>
</tr>
</tbody>
</table>

Notes: imp sg = imperative singular, 2\textsuperscript{nd} sg = second person singular present indicative, 3\textsuperscript{rd} sg = third person singular present indicative form of the verb *sehen* ‘see’; see Appendix F for token frequencies and proportions of individual verbs

Most explanations of partial change in morphological paradigms in previous research are unable to account for this sequence of the ongoing change in the German strong verbs with *e/i*-gradation. Therefore, the present study assesses the explanatory potential of the frequency-based approach with regard to the conservation of the traditional *i*-ablaut in the second and third person singular forms.

7.1 Previous accounts of partial change in morphological paradigms

The fact that morphological paradigms tend to change gradually, i.e. one form at a time, has been explained by a number of theories which start out from similar assumptions. Some of these attempts will be outlined briefly in Sections 7.1.1 to 7.1.3, along with their predictions.
for the trajectory of the change-in-progress in imperative formation of German strong verbs with e/i-gradation. Section 7.2 will present the hypotheses of the frequency-based approach against which the other approaches are tested in the remainder of the chapter.

7.1.1 Natural morphology

Natural morphology is a theory of language change most prominently represented by Wurzel (1984), Mayerthaler (1981) and Dressler (2003; cf. Dressler et al. 1987). In this view, features of a language that are unnatural or marked are changed, so that the language becomes more natural or unmarked. The naturalness of a feature can be assessed on several parameters.

(i) Iconicity describes the extent to which the form of a linguistic unit is iconic. Dressler gives several examples, among them the diagrammatic (= most iconic) form *oafs* in which the morphotactic addition of the plural -s mirrors the morphosemantic addition of plurality, the less iconic form *feet* in which plurality is only recognisable metaphorically by vowel modification, and the form *sheep* which is categorised as non-iconic because the morphosemantic feature of plurality is not morphologically realised (2003: 263). (ii) The second universal parameter is that of (morphosemantic and morphotactic) transparency: linguistic units are fully transparent and most natural when they have a fully compositional meaning, as in the example plural form *oafs*. An example of a more opaque form cited by Dressler (2003: 464) is that of morphophonological palatalisation in the plural form *Polacy* of the singular word *Polak* ‘Pole’- the link of the plural form to its singular base is not as transparent as it would be without palatalisation. The remaining major parameters are (iii) binarity, i.e. “the preferred patterning consists in concatenating one element to one base” (Dressler 2003: 465), (iv) indexicality - adjacency is preferred to distance - and (v) (bi-)uniqueness (as opposed to ambiguity) which means that a link between form and meaning in at least one of two directions holds: one and the same form always has the same meaning and/or the same meaning is always expressed by one and the same form.

According to this theory, “the more natural a phenomenon is on a given parameter, the more stable, that is, the more resistant it should be to morphological change” (Dressler 2003: 463); thus, the replacement of the traditional irregular stem vowel *i* in the formation of the imperative singular of strong verbs with e/i-gradation and its conservation in the second and third person singular present indicative may be explained on the basis of the degree of naturalness of these paradigm forms. In terms of iconicity (i), all three forms are not fully natural: like the example form *loaves* cited by Dressler (2003: 463) they range between diagrammatic and metaphoric because they exhibit both stem alternation, e.g. *gib > geb*, and the addition of inflectional morphemes for mood or person, e.g. *gib-ø*, *gib-st* and *gib-t*. Consequently, all
three are examples of hypercharacterization (Dressler 2003: 464), which means that the morphological category is doubly marked, and thus highly transparent (ii). They are all binary forms (iii) - one inflectional element is added to one base - and this inflectional morpheme is added at the same place in the imperative singular and the second and third person singular present indicative; thus, they are equally indexical (iv). Finally, as long as the traditional irregular stem vowel \(i\) is retained, the degree of naturalness on the parameter bi-uniqueness (v) is identical for all three paradigm forms: The forms \(gib\), \(gibst\) and \(gibt\) have only one meaning each which can only be expressed by these forms. In a very small group of strong verbs with \(e/i\)-gradation, the forms with the stem vowel \(i\) are not unambiguous; for example, \(tritt\) (infinitive \(treten\) ‘step/kick’) expresses both imperative singular and third person singular present indicative and the form \(liest\) (infinitive \(lesen\) ‘read’) refers to both the second and third person singular present indicative (likewise \(vergessen\), \(essen\), \(fressen\), \(messen\)). Thus, according to natural morphology, all three paradigm forms should resist change because they exhibit identical degrees of naturalness on all parameters, or all three should change because they exhibit the unnatural feature of stem vowel alternation (cf. Nübling et al. 2013: 63).

If the stem vowel alternation to \(i\) was removed in German strong verbs with \(e/i\)-gradation by a formation with the stem vowel \(e\), this would increase iconicity and thus naturalness of all three paradigm forms. Likewise, the analogical variants of all three are still highly transparent because their meaning is compositional and the link to the infinitive base is even more transparent: for example, the base meaning \(geb\)- is combined with the meaning of imperative singular (\(-\varphi\)), second person singular present indicative (\(-st\)) or third person singular present indicative (\(-t\)). Naturalness values of all three paradigm forms on the two parameters of binarity and indexicality do not change, either, when the stem vowel \(i\) is replaced by \(e\). It is only on the parameter of (bi-)uniqueness that the forms start to differ in naturalness. If the stem vowel alternation in the imperative singular and the second and third person singular present indicative was removed, this would not improve naturalness in the group of verbs comprising \(lesen\), \(vergessen\), \(essen\), \(fressen\) and \(messen\): the form \(lest\) would still be ambiguous and refer to the second and third person singular present indicative alike. In all other strong verbs with \(e/i\)-gradation, a formation with the stem vowel \(e\) would not mean a difference in terms of the parameter (bi-)uniqueness for the second person singular present indicative: the forms \(gebst\) and \(tretest\) would be the only forms referring exclusively to the second person singular present indicative of the respective verb. However, as mentioned before, the analogical imperative singular variants of strong verbs with \(e/i\)-gradation, e.g. \(geb(e)\), are identical with the first person singular present indicative forms of the same verbs (contracted
first person singular forms are not uncommonly used), and analogical variants of the third person singular present indicative are also identical with other forms in the paradigm, viz. the forms of the second person plural present indicative and imperative (e.g. gebt). In summary, this means that the replacement of the traditional irregular stem vowel e in the imperative singular of strong verbs with e/i-gradation results in a decline of naturalness, and the same happens in the third person singular present indicative; only the degree of naturalness of the second person singular present indicative does not decrease when the stem vowel alternation is removed. Thus, even when the resulting analogical forms are concerned, the theory of natural morphology predicts a different sequence for analogical levelling in strong verbs with e/i-gradation from the sequence which can be observed in reality: the analogical second person singular present indicative form is more natural and therefore should be less resistant to change than the forms of the imperative singular and the third person singular present indicative.

Andreas Bittner (1996) developed an *implicational hierarchy* of New High German verb inflection, likewise on the basis of the tenets of naturalness theory; this hierarchy matches the observed sequence of change in the strong verbs with e/i-gradation; as Figure 34 illustrates, the presence of a stem vowel alternation in the imperative singular of a verb class implies the presence of a concurrent stem vowel alternation in the second and third person singular present indicative.

As Bittner explains, the implicational pattern “allows for a meaningful classification of verbs on the basis of functional relations, constitutes the basis for learning and handling different inflectional paradigms and documents the sequence of the removal of non-weak characteristics” (1996: 80, emphasis added). Thus, the hierarchy implies that the stem vowel alternation in the imperative singular will be removed prior to that in the second and third person singular. However, Bittner’s figure is nothing more than a documentation of the current state of affairs and a classification tool: even if his implicational patterns holds true, it does not give any reasons or explain a mechanism for the depicted chronology. This problem reflects the
main point of criticism which has been voiced against natural morphology or naturalness theory: it is teleological (goal-oriented) and its tenets are not falsifiable (e.g. Bybee 2015). It may offer good post-hoc explanations and documentations of different phenomena of language change, but the existence of conflicting parameters and principles makes it hard to predict which one of them will win out in a particular case of change (e.g. iconicity or indexicality, cf. Dressler 2003: 465, 469-470).

7.1.2 Relevance theory

Relevance theory has adopted the concept of markedness from natural morphology; it assumes that cross-linguistically, linguistic categories are marked in different ways as a function of their semantic relevance, i.e. “the extent to which the meaning of the grammatical morpheme affects the meaning of the stem” (Bybee 1994: 2559; cf. Bybee 1985: 13-24). In nominal morphology, for example, the morphological category of number affects the meaning of the nominal stem more strongly than definiteness or case marking. Nübling et al. give examples from German (15-17) which illustrate this idea very clearly: it makes a big conceptual difference whether a dog is chased by one (19) or several cats (20); the presence of a plural suffix duplicates the semantic content of the stem. However, if the dog is chasing (18) or being chased (19) does not affect the meaning of the stem ‘dog’ - “Hund bleibt Hund” ‘a dog remains a dog’ (Nübling et al. 2013: 69).

(18) Der Hund jagt die Katze.
‘The dog chases the cat.’

‘The cat chases the dog.’

(20) Die Katzen jagen den Hund.
‘The cats chase the dog.’

As concerns verbs, the category of tense affects the meaning of the verb stem more strongly than the category person. Whereas the former transports the action to a different time, so that it is either already past or has not yet happened, the question whether a first or third person is performing the action is assumed to be less important. The category of mood ranges between tense and aspect on the relevance continuum: it “refers to the way the speaker presents the truth of the proposition in the discourse and real-world context [by expressing] probability, possibility and certainty” (Bybee 1985: 28). Conceptually, the category number ranges between mood and person, but due to diachronic developments grammatical morphemes in many languages express both number and person; therefore, Bybee places person on one rank with number (1994: 2559). Both are subsumed under the category of agreement.
which is argued to be less relevant to the verb stem because it does not affect the situation described by the verb stem but (only) the participants in discourse (Bybee 1985: 22-23).

The degree of relevance of a particular category is assumed to have two effects on its formal realisation. First, the relevance of a grammatical morpheme has an influence on its position with respect to the stem. In a study of 50 different languages, Joan Bybee found that more relevant grammatical categories tend to be placed closer to the stem, while grammatical morphemes which are semantically least relevant to the stem are placed farthest from it (1985: 33-35). Thus, suffixes follow on verbal stems in the order aspect-tense-mood-person/number, and prefixes precede the stem in the reverse order (Bybee 1994: 2559). Secondly, relevance has an influence on the degree of fusion of grammatical morphemes with the stem. In the context of verb morphology, Bybee has observed that “if the meaning of an inflectional morpheme is highly relevant to the verb, then it will often be the case that their surface expression units will be tightly fused, while the less relevant morphemes will have a looser association with the verb stem” (1985: 36). Thus, the positioning of grammatical morphemes with respect to the stem is replicated in the degree to which they fuse with it. For example, in German verb morphology, Nübling et al. observe that stem alternation and allomorphy were diachronically first removed for person and later for number; the category mood is signalled by umlaut which, however, depends on the ablaut by which the most relevant category tense is marked, e.g. third person singular preterite indicative sang and subjunctive sänge of the verb singen (2013: 71).

At first glance, the relevance concept may seem to offer a more appropriate explanation of the sequence of change in German strong verbs with e/i-gradation. However, the category mood is deemed to be semantically more relevant to the verb stem than person and number; therefore, a stem vowel alternation (highest degree of fusion with the stem) would be expected to be preserved longer in the imperative singular form than in the second and third person singular present indicative. Thus, the predictions of relevance theory are at odds with the observed chronology of change in the verb class, where the imperative singular starts to be formed analogically without stem vowel alternation, while vowel gradation in the supposedly less relevant second and third person singular present indicative is conserved. The inflectional suffixes marking the respective grammatical category are positioned in the same place with respect to the stem in all three paradigm forms; this potential effect of relevance does not provide evidence in support of or against the explanatory power of the relevance theory with regard to the change in the strong verbs with e/i-gradation.
Admittedly, relevance is defined mainly but not exclusively on the basis of semantics; two further factors named by Bybee are irregularity and frequency (1985: 57). Undoubtedly, stem vowel alternation in the imperative singular is equally irregular as it is in the second and third person singular present indicative; thus, this factor does not add to the explanatory value of the relevance concept. The present study aims to assess the power of the frequency-based approach in explaining the conservation of the traditional irregular stem vowel *i* in the second and third person singular present indicative of strong verbs with *e/i*-gradation; therefore, the discussion of the results will revisit the relevance concept and determine whether additional frequency information improves the adequacy of its predictions in the present change phenomenon.

### 7.1.3 Prototypicality and the Paradigm Structure Hypothesis

The last account of partial change in morphological paradigms to be discussed here, the *Paradigm Structure Hypothesis* (Janda et al. 2010; Nesset & Janda 2010), is an extension of the findings of relevance theory and insights from other linguistic fields. Instead of degrees of relevance, *prototypicality* is assumed to be the organizing principle of morphological paradigms. This concept is adopted from prototype theory (Rosch 1973; 1975) which describes the semantic content of members of a linguistic category in terms of their resemblance to an ideal representative of the category, the *prototype*. Members which bear more resemblance to the prototype are located closer to it; members with little resemblance to the prototype are located in the periphery of the category. According to Nesset and Janda, in the case of verbs, prototypicality is determined by finiteness, mood, person and number. In this context, the authors draw on findings from work by Bybee (1985) which suggests that finite forms are psycholinguistically more prototypical for speakers than non-finite forms (2010: 707). As concerns mood, they explain that the indicative is the most prototypical: “it presents the simplest relationship of a situation to reality. Indicative is about real things that are being observed” (Janda 2018: 308). The conditional/subjunctive and the imperative are more marked because they “depict a situation as outside reality [or] the speaker wants something that is outside actual reality to become part of actual reality” (Nesset & Janda 2010: 708). Similarly as in the case of relevance theory, evidence for the higher prototypicality of, for example, the indicative mood is provided on the basis of cross-linguistic comparison: “Typologically, if we look at verbs in indicative versus imperative forms, we find that it’s very rare for verbs to lack imperative forms” (Janda 2018: 308). The third person is held to be more prototypical than the first and second person, and the singular is more prototypical than the plural; this classification is again based on findings by Bybee (1985) and supported by evidence from Corbett...
(2000) and Lyashevskaya (2004): zero expression is more typical for third than first and second person and more typical for singular than for plural. In summary, Nesset and Janda present the following prototypicality scale of verb inflection forms, with the third person singular as the prototype and the non-finite gerund and participle forms farthest in the periphery (2010: 709; Janda et al. 2010: 31):

![Prototypical vs. peripheral forms](image)

**Figure 35**: Prototypicality within the paradigm (Janda 2018: 311)

Under the assumptions of the Paradigm Structure Hypothesis, change in verbal paradigms is expected to proceed from the periphery to the prototype (Nesset & Janda 2010: 712). The authors test the hypothesis on the basis of the Russian suffix shift, an ongoing analogical change in the morphology of a group of 37 Russian verbs from the traditional suffix -a to the more productive suffix -aj. By means of a statistical analysis of a dataset of both variants of all paradigms forms of these verbs extracted from the Russian National Corpus, they are able to demonstrate an almost perfect match between the above prototypicality scale and the propensity of the respective paradigm forms of the Russian verbs to occur with the analogical -aj suffix (Nesset & Janda 2010: 713-719).

As a case of ongoing change in a verbal paradigm, analogical levelling in German strong verbs with e/i-gradation presents another possible test case for the proposed prototypicality scale. According to the hypothesis set up by Nesset and Janda, the most peripheral imperative form of these verbs should change first, followed by the second person singular pre-
sent indicative form which is more prototypical. The third person singular present indicative as the prototype would be expected to resist analogical levelling most strongly. This sequence of change is borne out by the development of the imperative singular, which is currently changing from the traditional irregular formation with the stem vowel $i$ to the analogical stem vowel $e$. The analogical change in the strong verbs with $e/i$-gradation is not progressed far enough to provide evidence for the sequencing of the second and third person singular present indicative; the conservation of the traditional irregular stem vowel alternation in these paradigms is an indication of the validity of the Paradigm Structure Hypothesis.

However convincing this may sound, the suggested prototypicality scale seems to fall prey to a few contradictions. On the one hand, although this prototypicality is explained with reference to a lot of evidence provided by Bybee (1985) for her relevance theory account of inflectional marking, the two approaches make contradictory predictions as to the development of the imperative and second and third person singular of German strong verbs with $e/i$-gradation. Whereas the grammatical category of mood is held to be more relevant than number and person (Bybee 1985) and therefore more resistant to change than the second and third person singular, the imperative mood is assumed to be less prototypical (Nesset & Janda 2010; Janda 2018) and therefore more prone to change than the second and third person singular. Moreover, assuming that the imperative mood is more peripheral and therefore, cross-linguistically, more typically marked by inflection, it would seem all the more logical for the imperative singular of German strong verbs with $e/i$-gradation to retain its unambiguous inflection, as analogical levelling of the majority of these forms would result in ambiguity between the imperative singular and the first person singular present indicative. Finally, the explanation of the order of elements in the prototypicality scale suggested by Janda and Nesset seems to rely heavily on verbal inflection and on the Russian language in particular, which may explain why conditional forms and tense distinctions are not included in the sketch in Figure 35. Time may tell whether the model holds across languages and different instances of analogical change.

Interestingly, Nesset and Janda also test whether the observed Russian suffix shift hierarchy is correlated with the frequency of forms in the paradigms of the Russian verbs subject to this analogical change. They find that “frequency correctly predicts the order of four of the [six] forms” (Nesset & Janda 2010: 720); the incorrect order of the remaining two paradigm forms, gerunds and participles, is explained by the authors on the basis of a bias in the corpus data: the Russian National Corpus consists of written language data in which both paradigm forms occur more frequently than in spoken language. On the basis of frequency
data of the imperative singular and the second and third person singular present indicative of strong verbs with e/i-gradation from corpora of written and spoken German, the present study will test whether a purely frequency-based account of partial change in morphological paradigms may be sufficient to explain the conservation of the traditional stem vowel alternation in the second and third person singular present indicative of German strong verbs with e/i-gradation.

7.2 Hypotheses

In the Walkthrough Corpus study (Chapter 4) and other studies of analogical change (Bybee & Thomson 1997; Poplack 2001; Smith 2001), verb token frequency has been found to exert a conserving effect in analogical change: lower frequency verbs change earlier and faster than higher frequency verbs. A similar conserving effect has also been observed to work within morphological paradigms in that “the more frequent members of paradigms are less likely to change” (Bybee 1994: 2560). Such a frequency effect may explain the observation that the imperative singular of strong verbs with e/i-gradation is increasingly used in analogically formed variants with the stem vowel e, while the second and third person singular present indicative conserve the traditional irregular formation with the stem vowel i. Thus, the latter paradigm forms are expected to occur more frequently in different corpora than the imperative singular.

Like the conserving effect exerted by verb lemma token frequency, a potential intraparadigmatic conserving effect would be explained on the basis of entrenchment. The second and third person singular indicative forms can be expected to be more strongly entrenched in speakers’ mental lexicons on the basis of their higher frequency; they are therefore easily accessible and less susceptible to analogical levelling. By contrast, the less frequent imperative singular form is expected to be weakly entrenched; instead of its traditional irregular form, the productive imperative singular formation without stem vowel alternation is accessed more readily.

7.3 Data

For the purpose of testing these hypotheses, token frequencies for the imperative singular and the second and third person singular present indicative are extracted from three corpora of German which represent different types of language. The corpus Wortschatz Universität Leipzig (Leipzig University 1998-2019) contains texts from online newspapers and thus
represents written language. The two other corpora used in the study, moca (University of Freiburg 2019) and FOLK (Institut für deutsche Sprache 2018), contain searchable transcripts of spoken German. The corpora have been selected so as to approximate the time spanned by the Walkthrough Corpus data (2001-2013). Further details of the corpora are provided in Chapter 3.1.

7.4 Method

All corpora were searched for forms of the imperative singular and the second and third person singular present indicative of all strong verbs with e/i-gradation. The majority of counts could be transferred directly; only in the case of ambiguous forms, such as liest ‘you read/ he, she, it reads”, prior manual inspection of the query hits was necessary. The number of query hits for a particular form was entered into one of three contingency tables for the three corpora; within the tables, counts were entered separately for each verb. The token frequencies for each of the paradigm forms were added up for all verbs in each of the corpora, as illustrated in Table 37.

Table 37: Sample token frequencies for paradigm forms in three corpora of German

<table>
<thead>
<tr>
<th>verb</th>
<th>moca database</th>
<th></th>
<th>FOLK corpus</th>
<th></th>
<th>Wortschatz Universität Leipzig</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>imp sg 2nd sg 3rd sg</td>
<td>imp sg 2nd sg 3rd sg</td>
<td>imp sg 2nd sg 3rd sg</td>
<td>imp sg 2nd sg 3rd sg</td>
<td>imp sg 2nd sg 3rd sg</td>
<td>imp sg 2nd sg 3rd sg</td>
</tr>
<tr>
<td>sprechen</td>
<td>8 231 1,480</td>
<td>4 26 175</td>
<td>327 460 48,401</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sehen</td>
<td>0 185 651</td>
<td>18 293 780</td>
<td>384 1,652 140,980</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>135 615 6,706</td>
<td>443 522 4,094</td>
<td>5,332 4,286 956,252</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: imp sg = imperative singular, 2nd sg = second person singular present indicative, 3rd sg = third person singular present indicative

7.5 Results

Token frequency counts for forms of the imperative singular and the second and third persons singular present indicative were extracted for 22 strong verbs with e/i-gradation from the moca database, for 28 verbs from the FOLK corpus, and for 39 verbs from the corpus Wortschatz Universität Leipzig. Figure 36 displays the distribution of the three paradigm forms in the three corpora. The third person singular form is by far the most frequent, while the ranking is not as clear for the two remaining paradigm forms. In the moca database and the FOLK corpus, the second person singular form is slightly more frequent than the impera-
tive singular form; in the Wortschatz Universität Leipzig the difference in token frequency between the two forms is hardly noticeable.

Figure 36: Token frequencies of the imperative singular and the second and third person singular present indicative of strong verbs with e/i-gradation in three corpora of German

7.6 Discussion

The generally higher proportion of imperative singular and second person singular present indicative forms in the moca database and the FOLK corpus can be explained as a result of the data basis: both corpora contain transcripts of spontaneous conversations between speakers who are more or less familiar with each other. The instances of these paradigm forms in the corpus Wortschatz Universität Leipzig occur predominantly in readers’ comments to news articles, which make up the smallest share of the corpus data.

In the case of the third person singular present indicative form, a conserving effect of high token frequency as the reason for its resistance to analogical levelling seems to be a plausible explanation. These forms are used so frequently by speakers in different types of discourse that they are highly entrenched in their mental lexicons and quickly retrieved when needed. The small difference in token frequency between the imperative singular and the second person singular present indicative, however, indicate that an entirely frequency-based explanation cannot account for the observed sequence of analogical levelling in strong verbs with e/i-gradation (cf. Table 36): on the basis of the quantitative data consulted here, both paradigm forms should be expected to show similar proportions of analogical variants. Even when the token frequency of paradigm forms is taken into consideration, the predictions from natural morphology (NM) and relevance theory (RT) do not match linguistic reality, either.
According to both theories, the second person singular present indicative should be least resistant to levelling because its analogical variant is more natural than that of the other two paradigm forms (NM) or because it is assumed to be as little relevant to the meaning of the verb stem as the third person singular present indicative which is preserved on the basis of its high token frequency (RT).

The only account whose predictions match the observed sequence of analogical leveling in strong verbs with e/i-gradation is the Paradigm Structure Hypothesis (Janda et al. 2010; Nesset & Janda 2010): Out of the three paradigm forms potentially affected by the change, the imperative singular is assumed to be the least prototypical and therefore least resistant to levelling. In contrast to a purely frequency-based account which would predict similar proportions of analogical variants in the imperative singular and the second person singular present indicative, the Paradigm Structure Hypothesis assumes that the second person singular form is more prototypical and therefore resists change more strongly than the imperative singular. This is borne out by data from the FOLK corpus presented in Table 36: in terms of the mean proportion of analogical variants, the second person singular form ranks second after the imperative singular. The probability of the conservation of the third person singular present indicative form on the basis of its prototype status would be supported by its very high token frequency in all the three corpora consulted here as well. However, given the reservations voiced in Chapter 7.1.3, the prototypicality scale does not appear as a more reliable explanation of the sequence of analogical levelling in German strong verbs with e/i-gradation than an explanation based solely on the frequencies of the paradigm forms.

Perhaps, change in the imperative singular of the strong verbs with e/i-gradation and a concurrent conservation of the second and third person singular forms is a case of a “therapeutic change”. Aitchison (2001: 169) uses this term as a heading for changes which repair inconsistencies in the phonology, morphology and syntax of languages which were caused by previous changes. Chapter 1.2.2 had explained that the first person singular of strong verbs with e/i-gradation had undergone analogical levelling already during Early New High German. The inflection system for the present indicative and imperative of German strong verbs which was created by this change has only one exception: the imperative singular of strong verbs with e/i-gradation. Apart from the latter form, the system is divided between verbs which show the same stem vowel throughout the present indicative and imperative (ablaut classes I, II, IIIa, VIIc, VIIf) and verbs with a stem vowel alternation in the second and third person singular present indicative (IIIb, IV, V, VI, VIIa,b,d,e). Only strong verbs with e/i-gradation usually exhibit the additional stem vowel alternation which is the topic of the pre-
sent dissertation. If this alternation was removed, the German verb inflection system would be repaired to such an extent that i) strong verbs would only show stem vowel gradation according to tense and in the second and third person singular indicative and ii) the imperative singular of all verbs in German could be formed from the infinitive stem.

7.7 Summary

The present study started out on the observation that the imperative singular of German strong verbs with e/i-gradation is subject to analogical levelling, while the second and third person singular present indicative forms with the same stem vowel alternations are preserved in their traditional irregular formation to a much higher extent. On the basis of data from three corpora of spoken and written German, it has contrasted the explanatory potential of a purely frequency-based account with that of previous explanations of partial change in morphological paradigms with regard to analogical levelling in the strong verbs with e/i-gradation.

The predictions from natural morphology and relevance theory do not match the observed sequence of change in the verb class; both would predict that analogical levelling starts out in the second person singular present indicative. Token frequencies of the paradigm forms in the three corpora could explain the observed change in the imperative singular and the conservation of the third person singular form: the latter has a very high token frequency and can thus be assumed to be so highly entrenched in speakers’ minds that it resists change. By contrast, the imperative singular has a very low token frequency; its resulting weak mental entrenchment makes it succumb to analogical replacement more readily. However, the second person singular has a similar frequency to that of the imperative singular, that is, in the corpora used for the present study; thus, the frequency-based account would incorrectly predict both paradigm forms to exhibit equal proportions of analogical variants. The Paradigm Structure Hypothesis is the only account whose predictions on the basis of a prototypicality hierarchy of forms in morphological paradigms match the observed sequence of change in strong verbs with e/i-gradation so far, but the underlying prototypicality scale does not seem to be fully reliable. It is possible that change in the imperative singular simply serves to “neaten” the system of German verb inflection, as was the case for the previous change in the first person singular of the same verb class (cf. 1.2.2).
8 Conclusions and outlook

The present investigation set out to explore the explanatory potential, but also the limitations, of the frequency-based approach with regard to the ongoing analogical change in the imperative singular of German strong verbs with e/i-gradation from the traditional formation with the stem vowel i to an analogical formation with the stem vowel e. In line with previous usage-based research, corpus studies examined whether type and token frequencies can explain the trajectory and direction of this change-in-progress and the fact that analogical levelling in this verb class only affects the imperative singular although the second and third person singular present indicative exhibit the same stem vowel alternation. In addition, an experimental study was conducted in search of converging evidence for the prevalent explanation of frequency effects in analogical levelling on the basis of the cognitive entrenchment of forms.

The first corpus study endeavoured to explain the direction of analogical levelling in strong verbs with e/i-gradation on the basis of the type and token frequency and schematicity of other verb inflection classes of German. One of the merits of the investigation is that, instead of drawing on existing (arguably outdated) counts of the type frequency of these inflection classes, recent corpus-based material in the form of the DeReWo list of lemma token frequency (Institut für deutsche Sprache 2012b) was consulted, so that the study provides contemporary figures of the distribution of German verb inflection classes. More importantly, it provides evidence that the productivity of these inflectional patterns is first and foremost determined by their type frequency and schematicity, i.e. the number of linguistic items which they already apply to and their structural applicability to new items. The token frequency of members of the different inflection classes does not reinforce the productivity of any of the verb classes; also in this regard, the study confirms findings from previous research (Bybee 1985; 1995).

The starting point of the second corpus study is the observation that analogical levelling in the imperative singular does not seem to progress through the class of strong verbs with e/i-gradation uniformly. A dataset of observations of the traditional irregular and analogical variants of these verbs was extracted from a corpus which was compiled specifically for the present investigation from a website of walkthroughs. The statistical analysis of this dataset confirms that the token frequency of a strong verb with e/i-gradation has an effect on its tendency to succumb to or resist analogical levelling, also known as the conserving effect of high token frequency (Bybee & Thompson 1997: 380). Thus, verbs with a low token frequency are used with a much higher probability with the analogical stem vowel e than verbs
with a higher token frequency, which are preserved in the traditional formation with the stem vowel \( i \). Apart from this finding, the study led to further insights which may be of relevance to other studies of morphological change. In the case of the imperative singular of strong verbs with e/i-gradation, the conserving effect is not exerted by verb lemma token frequency, the frequency measure which is considered in most previous studies in this context, but by the token frequency of the base verb lemma, i.e. German prefixed verbs and the simplex base verbs from which they are derived behave alike during analogical levelling. Further frequency variables which were taken into account in the analysis of the walkthrough dataset, such as intraparadigmatic type and token frequencies, did not have a significant effect on the formation of the imperative singular. Depending on the language under investigation, the concept of frequency may thus have to be reconsidered, as in the present analysis; in languages with even richer morphological paradigms than German, intraparadigmatic frequencies may have an influence on analogical change as well. Finally, the study revealed that investigations of ongoing language change need to take into account the potential influence of variables which have been shown to influence phenomena of linguistic variation. In the present case, persistence is found to have an effect on the suffixation of imperative singular forms; thus, although analogical levelling in the imperative singular of strong verbs with e/i-gradation is argued to have passed a stage of pure linguistic variation on the basis of the results of this corpus study, it is still affected by some of the relevant processes.

The conserving effect of high token frequency in analogical change is traditionally explained on the basis of entrenchment (Bybee & Thompson 1997: 380; Langacker 1987): irregular formations are retained in high frequency items of a language because their frequent repetition makes them strongly entrenched in the mental lexicon of speakers. Although this explanation has been taken up in many studies of analogical change, twenty years after its formulation no converging evidence for its validity exist in the form of behavioural data contrasting reactions to traditional and analogical variants of the same verb obtained from psycholinguistic experiments. The ongoing analogical levelling process in the imperative singular of strong verbs with e/i-gradation presents a prime test case in this regard. The traditional irregular imperative singular variant of higher frequency strong verbs with e/i-gradation with the stem vowel \( i \) is assumed to be conserved because it is more strongly entrenched in the minds of speakers of German than analogical variants of the same verbs and the variants of lower frequency verbs. In order to test this and related hypotheses, an experiment was conducted with 120 participants who were presented with stimulus sentences containing the imperative singular variants of strong verbs with e/i-gradation of different frequency (matching
the frequency distributions of the variants in the Walkthrough Corpus) and asked to repeat the stimulus sentences. Reading times and recall accuracies constitute measures of the strength of entrenchment of the stimulus forms in participants’ mental lexicons. The results of the experiment provide converging evidence for the explanation of the conserving effect of high token frequency on the basis of mental entrenchment: the traditional imperative singular variant of high frequency strong verbs with e/i-gradation is the form which causes the lowest reading times and the highest recall accuracies. In addition, the analysis of the experiment data confirms that base verb lemma token frequency is more relevant to the conservation of the imperative singular of strong verbs with e/i-gradation than verb lemma token frequency.

Not only does analogical levelling in strong verbs with e/i-gradation affect the imperative singular of verbs to a different extent; it is also observed that the second and third person singular present indicative of these verbs with the same traditional stem vowel alternation do not show real signs of change. The third corpus study therefore examines whether the token frequency of these paradigms exerts a conserving effect similar to the one which was identified in the Walkthrough Corpus. On the basis of data from three corpora of spoken and written German, a purely frequency-based explanation of the conservation of the traditional stem vowel alternation in the second and third person singular present indicative must be rejected. The only account whose predictions match the sequence of change in strong verbs with e/i-gradation is the Paradigm Structure Hypothesis (Janda et al. 2010; Nesset & Janda 2010) which explains partial change in morphological paradigms on the basis of the prototypicality of the forms contained in them. Since the explanation of the underlying prototypicality scale is not entirely convincing, the respective chapter suggests that change in the imperative singular only serves to repair an inconsistency in German verb inflection, so that change in the second and third person singular forms is not necessary.

In summary, the present investigation has shown that the frequency-based account can offer explanations of many aspects of analogical change. It also provides converging evidence from experimental data for one of its cornerstones, the explanation of the conserving effect of high token frequency on the basis of the mental entrenchment of forms. However, the summary above also illustrates that frequency is not a methodological cure-all. Intraparadigmatic frequencies do not have an influence on a verb’s tendency to succumb to analogical levelling; neither can they explain why the imperative singular of strong verbs with e/i-gradation is currently changing from the traditional irregular formation with the stem vowel i to an analogue formation with the stem vowel e, while the same stem vowel alternation is preserved in the second and third person singular of the same verbs.
Analogical levelling in strong verbs with e/i-gradation should be monitored in future work in linguistics: the replacement of the traditional stem vowel alternation in the imperative by variants formed in analogy to more productive verb inflection classes may still revert. The more plausible expectation, however, is that it extends to the imperative singular of higher frequency verbs and possibly to the second and third person singular present indicative of these verbs. Nevertheless, the third person form can be expected to be preserved for a very long time, due to its extremely high token frequency.
References

Corpora and frequency dictionaries


Software


Bibliography


### Appendices

#### A Eligible verbs for stimulus sentences in the experimental study

Verbs in boldface occur in the Walkthrough Corpus dataset and in the stimulus sentences for the experimental study. Verbs in regular print occur only in the stimulus set for the experiment. Verbs in grey do not occur in either of the two studies.

<table>
<thead>
<tr>
<th>Base verbs with a token frequency class (DeReWo) above the median</th>
<th>Base verbs with a token frequency class (DeReWo) below the median</th>
</tr>
</thead>
<tbody>
<tr>
<td>geben (6)</td>
<td>fressen (13)</td>
</tr>
<tr>
<td>nehmen (7)</td>
<td>stechen (13)</td>
</tr>
<tr>
<td>sehen (7)</td>
<td>bestechen (13)</td>
</tr>
<tr>
<td>treffen (8)</td>
<td>erschrecken (13)</td>
</tr>
<tr>
<td>sprechen (8)</td>
<td>unterwerfen (13)</td>
</tr>
<tr>
<td>helfen (9)</td>
<td>verhelfen (13)</td>
</tr>
<tr>
<td>lesen (9)</td>
<td>durchbrechen (13)</td>
</tr>
<tr>
<td>sterben (9)</td>
<td>zerbrechen (13)</td>
</tr>
<tr>
<td>übernehmen (9)</td>
<td>befehlen (14)</td>
</tr>
<tr>
<td>treten (9)</td>
<td>fechten (14)</td>
</tr>
<tr>
<td>brechen (10)</td>
<td>verderben (14)</td>
</tr>
<tr>
<td>werfen (10)</td>
<td>vermessen (14)</td>
</tr>
<tr>
<td>vergessen (10)</td>
<td>verwerten (14)</td>
</tr>
<tr>
<td>ergeben (10)</td>
<td>benennen (15)</td>
</tr>
<tr>
<td>versprechen (10)</td>
<td>bewerfen (15)</td>
</tr>
<tr>
<td>essen (11)</td>
<td>erstechen (15)</td>
</tr>
<tr>
<td>stehlen (11)</td>
<td>dressen (16)</td>
</tr>
<tr>
<td>werben (11)</td>
<td>flechten (16)</td>
</tr>
<tr>
<td>empfehlen (11)</td>
<td>behelfen (16)</td>
</tr>
<tr>
<td>vergeben (11)</td>
<td>bersten (16)</td>
</tr>
<tr>
<td>erwerben (11)</td>
<td>beobachten (16)</td>
</tr>
<tr>
<td>übergeben (11)</td>
<td>bestehlen (16)</td>
</tr>
<tr>
<td>unternehmen (11)</td>
<td>melken (16)</td>
</tr>
<tr>
<td>messen (12)</td>
<td>übertreten (16)</td>
</tr>
<tr>
<td>begeben (12)</td>
<td>ermessern (17)</td>
</tr>
<tr>
<td>entwerfen (12)</td>
<td>überlesen (17)</td>
</tr>
<tr>
<td>besprechen (12)</td>
<td>zerstechen (17)</td>
</tr>
<tr>
<td>bergen (12)</td>
<td>zertreten (17)</td>
</tr>
<tr>
<td>betreten (12)</td>
<td>durchstechen (18)</td>
</tr>
<tr>
<td>bewerben (12)</td>
<td>erlechtern (19)</td>
</tr>
<tr>
<td>entnehmen (12)</td>
<td>verdreschen (19)</td>
</tr>
<tr>
<td>übersprechen (12)</td>
<td>zerstechen (23)</td>
</tr>
<tr>
<td>übertreffen (12)</td>
<td>zerwerfen (23)</td>
</tr>
<tr>
<td>umgeben (12)</td>
<td></td>
</tr>
<tr>
<td>unterbrechen (12)</td>
<td></td>
</tr>
<tr>
<td>verbergen (12)</td>
<td></td>
</tr>
<tr>
<td>widersprechen (12)</td>
<td></td>
</tr>
</tbody>
</table>
All statistical analyses were carried out using the programme R, “a free software environment for statistical computing and graphics” (R Core Team 2019; version 3.5.3) and the environment RStudio (RStudio Team 2016; version 1.1.463). Mixed-effects regression models were fitted using the functions `lmer()` for linear mixed-effects regressions and `glmer()` for logistic mixed-effects regressions in the package `lme4` (Bates et al. 2015; version 1.1-21, March 2019). Significance levels (p-values) for predictors are automatically provided in `glmer` for mixed-effects logistic regression models but not in `lmer` models because it is not trivial to calculate the appropriate degrees of freedom (cf. Baayen 2008: 269); the function `get_Lb_ddf()` in the package `pbkrtest` (Halekoh & Højsgaard 2014; version 0.4-7, March 2017) was used in order to get adjusted degrees of freedom (by a Kenward-Roger approximation) for the linear mixed-effects regression models (`lmer`) on the basis of which significance values can be calculated. Significance values below the threshold of 0.05 are marked with asterisks, following the usual conventions: \( p < 0.05 \ast, p < 0.01 \ast\ast, p < 0.001 \ast\ast\ast \).

The R default of using treatment contrast coding for regressions was not changed. “Treatment contrast” means that for categorical (nominal and ordinal) predictors, the alphabetically first or numerically lowest level of the factor is selected as a reference level; the means for all remaining levels of the variable are compared to the mean of this reference level. In sum contrast coding, the means of all levels of the categorical variable would be compared to the grand mean, i.e. the “mean of the means of the dependent variable, when it is grouped by the independent variable” (Gries 2013: 268). In Harald Baayen’s example presented in Section 2.3, one mean would be calculated for all reading times obtained in an experiment under the condition “white noise” and one of all reading times obtained under the condition “no white noise”; the mean of these two means is the “grand mean”. In sum contrast coding, the means of the levels “white noise” and “no white noise” would be compared to this grand mean, whereas in treatment contrast coding the mean of “white noise” is compared to that of the reference level “no white noise. Treatment contrast coding will be retained in the presentation of regression models as well. When, in the regression tables presented in this investigation, the significance values of some levels of a categorical predictor are above 0.05, this does not indicate that the predictor as such does not have a significant effect on the dependent variable; it simply means that the means for the respective levels are not significantly different from that of the reference level.

Section 2.3 hinted at the fact that in linear regression the original values of the dependent variable are retained, e.g. reading times, whereas in logistic regressions with a binary
dependent variable chances, or more precisely odds, are computed for the occurrence of the response level as compared to the reference level, e.g. the odds for correct recall of a stimulus in the experiment vs. incorrect recall. In all logistic regression models used in the present paper, the original log-odds estimates provided by the glmer() function will be reported. They describe both the direction and strength of influence: positive signs indicate an increase in the probability of the occurrence of the response level, negative signs point towards a higher probability of the reference level; the value itself quantifies the change on the y-axis, for numeric predictors with regard to the y-axis intercept, for categorical variables with regard to the intercept of the reference level. These log-odds values between -∞ and +∞ can be translated into probabilities between 0 and 1, where a value of 0 corresponds to a 0 percent probability and a value of 1 corresponds to 100 percent probability for the response level to be used (cf. Gries 2013: 300). In the present paper, effects from both the linear and logistic regression models were extracted using the package effects (Fox 2003; version 4.1-0, November 2018), thereby transforming log-odds to probabilities, which were illustrated in graphs created with the help of the graphics package ggplot2 (Wickham 2009; version 3.1.0, October 2018). Some parts of figures were combined with the help of the function ggarrange in the package egg (Auguie 2018; version 0.4.2, November 2018).

The final regression models reported in the individual chapters are corrected for outliers, i.e. “data points with values that are surprisingly large or small given all data points considered jointly” (Baayen 2008: 29). In the experimental study, for example, a reading time of 58,460 msec for one stimulus word was measured from one participant who must have been distracted; this data point is an obvious outlier at an average reading time of 663.8 msec. In regression analyses of the Walkthrough Corpus data (Chapter 5.3.2) and those of reading times obtained from the experimental study (Chapters 6.4.2.2 and 6.5.2), such outlier data points were removed by creating a boxplot of the residuals of the respective regression model and identifying data points that are located outside (above or below) the whiskers in the boxplot, i.e. they have a value higher or lower than 1.5 times the interquartile range (cf. Baayen 2008: 32-33). Thus, instead of “blindly remov[ing] data points with extreme values [...] a priori” (Baayen 2008: 266), outliers are removed only after variation in the dependent variable caused by the fixed factors has been accounted for. In regression analyses of recall data obtained from the experimental study (Chapters 6.4.3.1 and 6.5.3), automatic outlier correction on the basis of residuals would have removed a large part of the data points for recall of imperative singular stimulus forms with the stem vowel e (amounting to 21.68 % of all data points) and thereby the effect of the variable IMP.FORM. Therefore, outlier data points in
these analyses were defined as “unusual values on the dependent variable, given the values of the independent variables” (Menard 2002: 67) and identified on the basis of the distribution of fitted values of the dependent variable instead of using the automatic procedure. These outlier data points are illustrated in the circles in Figures 37 to 39 for the individual participant groups.

Figure 37: Outlier data points in analyses of recall accuracy from participants from Saxony

Figure 38: Outlier data points in analyses of recall accuracy from older participants from Baden-Württemberg
Figure 39: Outlier data points in analyses of recall accuracy from younger participants from Baden-Württemberg

Measures of the quality of a regression model in accounting for the variation in the respective dependent variable are provided for the final outlier-corrected regression models. For convenience, the function summ() in the R package jtools (Long 2019; version 2.0.1, April 2019) is used to extract values for the number of data points analysed, AIC, degrees of freedom (df) in the model, and $R^2$ marginal and $R^2$ conditional. The AIC (Akaike Information Criterion) is computed on the basis of the degrees of freedom in the regression model (determined by the number of parameters in the model) and its maximum likelihood-ratio; the lower its values, the better does the regression model account for the variation in the data. This measure is penalising in the sense that out of two regression models with the same likelihood-ratio, the one with the higher number of degrees of freedom has a higher AIC value, i.e. it is evaluated as the worse model, because it does not account for more variation in the dependent variable even though it contains a higher number of predictors (or predictors with a higher number of levels). Therefore, the AIC value is always provided along with the degrees of freedom in the regression table. $R^2$ values are typically reported for linear regressions as a measure of the share of variation in the dependent variable accounted for by the predictors in the regression model. Pseudo-$R^2$ values are now often reported for mixed-effects regressions as well, as an indication of how much variation is explained by the fixed factors only ($R^2$ marginal) or by the fixed and random factors together ($R^2$ conditional.)
C  Extraction of data from the Walkthough Corpus

THE SEARCHING PROCEDURE

As explained in 3.1.2.2, a speech tagger programme was employed in the process of building up the Walkthrough Corpus in order to enable the extraction of relevant data from the corpus on the basis of POS annotations, hereby reducing the amount of manual word level searching. A comparison of the output for the query strings *vvimp2geben* (POS-level search) and *gib* (word-level search) revealed that the number of instances carrying the lemma tag “vvimp2geben” were identical in both documents and all instances were indeed uses of an imperative form of *geben*. Instances listed only in the output for the search string *gib* were correctly identified by the programme as uses of the string *gib* as a proper noun (a first or last name, tagged as “ne2gib”), or misspellings of the third person singular indicative, e.g. “es gib ...” instead of “es gibt ...” ‘there is/ are ...” or instances of dialectal variation in the first person singular indicative, e.g. “ich gib einen Tipp” ‘I give a hint’ (tagged as “vvfin2geben”, a finite form of the verb *geben*). Thus, while the POS-level query ”vvimp2[verb]” returned an exhaustive and exclusive list of instances of the use of imperative singular forms of strong verbs with *e/i*-gradation with the stem vowel *i*, uses of analogical variants with the stem vowel *e* needed to be searched on word level, resulting in two additional searches for the suffixed and the unsuffixed variants.

POS-TAG CORRECTION

Neither of the output documents for the queries *geb* and *gebe* in the Walkthrough Corpus contained instances tagged as “vvimp2geben”. Instead, the *geb* output (93 query hits) allegedly consisted of 27 finite forms of the verb *geben* and 66 proper nouns. However, manual correction showed that 13 items (≈14%) in fact turned out to be imperative singular forms. Similarly, in the output for the query *gebe* (413 hits), almost all instances were automatically recognised as “vvfin2geben”. Astonishingly, only 151 of these tokens are uses of finite forms of the verb (the first person singular present indicative or first or third person singular present subjunctive), while the remaining 262 query hits were found to be uses of the form *gebe* as a singular imperative. The erroneous POS-tagging is most probably caused by the fact that the TreeTagger was not trained on a text sample containing analogical imperative singular forms of strong verbs with *e/i*-gradation. After this first step of manual processing of the corpus hits, the dataset contained 2,446 instances of imperative singular forms of strong verbs with *e/i*-gradation.
DEDUPLICATION

However, two more steps of manual processing remained to be performed. The first of these is peculiar to the corpus: as the web crawling process did not involve a stage of duplicate removal, repeated instances needed to be removed manually. In the Walkthrough Corpus, repetitions of entire texts arise when a game has been published for several platforms and the walkthrough for it is copied from one platform entry to another. In these cases, only the “oldest” hit remained in the data set, i.e. the one with the earliest recorded game release date, and all succeeding editions were excluded from the analysis. When formulations were repeated within one and the same text, all instances remained in the data set. Following this procedure, 198 data points were excluded from the data set, reducing the number to 2,248 hits.

AUTHORSHIP CONFIRMATION

In much the same way as a high number of walkthroughs in German is written in the imperative singular, instructions included in the games themselves are also often found in this form. However, the present analysis did not aim at analysing the language of innominate publishers, programmers or translators but only the original language of the users of the forum. In order to obviate this risk, all 2,248 instances remaining at this point were checked for their context. In a number of texts, the forms of the imperative singular were found in quotation marks, suggesting that the user copied this material from within the game; such instances were excluded from the analysis. Similarly, instances of in-game terminology, such as *Brich aus-Modus* (‘bust out mode’, in the game *Die Sims brechen aus* for the Nintendo GameCube) were removed. The largest group of quoted material from the games, however, were enumerations or tables of trophies and achievements which can be won during the game play, which were excluded from the data set as well. Most of the irrelevant hits were easily identified on the basis of the walkthrough texts; cases of doubt were disambiguated by consulting YouTube Let’s Play videos for the respective games39. During this last step of manual processing, another 309 hits were excluded from the dataset, so that 1,939 instances finally remained.

---

D Results of the MCMCglmm regression model

Similarly as in the `multinom()` model (5.3.1), estimates for the effects of the predictor variables are provided for the analogical imperative singular variants in reference to the traditional variant with the stem vowel $i$ (for reasons of space, the reference level is omitted in the regression table).

Table 38: Results of MCMCglmm model

<table>
<thead>
<tr>
<th>predictor variable</th>
<th>factor level</th>
<th>dependent variable IMPERATIVE</th>
<th>estimate</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td>stem vowel e, unsuffixed</td>
<td>-1.7427</td>
<td>0.012 *</td>
</tr>
<tr>
<td>Intercept</td>
<td></td>
<td>stem vowel e, suffixed</td>
<td>-0.9967</td>
<td>0.010 *</td>
</tr>
<tr>
<td>PC1.STEM</td>
<td>- (numeric)</td>
<td>stem vowel e, unsuffixed</td>
<td>-0.3987</td>
<td>0.020 *</td>
</tr>
<tr>
<td>PC1.STEM</td>
<td>- (numeric)</td>
<td>stem vowel e, suffixed</td>
<td>-0.8285</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>POST.YEAR</td>
<td>- (numeric)</td>
<td>stem vowel e, unsuffixed</td>
<td>-0.1738</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>POST.YEAR</td>
<td>- (numeric)</td>
<td>stem vowel e, suffixed</td>
<td>-0.1350</td>
<td>0.004 **</td>
</tr>
<tr>
<td>LAST.SECOND.7</td>
<td>last suffixed</td>
<td>stem vowel e, unsuffixed</td>
<td>-1.58450</td>
<td>0.084</td>
</tr>
<tr>
<td></td>
<td>second unsuffixed</td>
<td>stem vowel e, unsuffixed</td>
<td>-0.93089</td>
<td>0.290</td>
</tr>
<tr>
<td></td>
<td>last suffixed, second unsuffixed</td>
<td>stem vowel e, unsuffixed</td>
<td>-2.89231</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td></td>
<td>last suffixed, second suffixed</td>
<td>stem vowel e, unsuffixed</td>
<td>0.44107</td>
<td>0.352</td>
</tr>
<tr>
<td></td>
<td>last unsuffixed, second suffixed</td>
<td>stem vowel e, unsuffixed</td>
<td>1.52504</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td></td>
<td>last unsuffixed, second unsuffixed</td>
<td>stem vowel e, unsuffixed</td>
<td>-3.10860</td>
<td>0.204</td>
</tr>
<tr>
<td></td>
<td>no pretarget</td>
<td>stem vowel e, unsuffixed</td>
<td>-1.20288</td>
<td>0.044 *</td>
</tr>
<tr>
<td></td>
<td>no pretarget</td>
<td>stem vowel e, suffixed</td>
<td>0.42754</td>
<td>0.398</td>
</tr>
<tr>
<td></td>
<td>suffixed</td>
<td>stem vowel e, unsuffixed</td>
<td>-8.06136</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td></td>
<td>suffixed</td>
<td>stem vowel e, suffixed</td>
<td>0.72510</td>
<td>0.466</td>
</tr>
<tr>
<td></td>
<td>suffixed SVe/iG</td>
<td>stem vowel e, unsuffixed</td>
<td>-1.20288</td>
<td>0.044 *</td>
</tr>
<tr>
<td></td>
<td>suffixed SVe/iG</td>
<td>stem vowel e, suffixed</td>
<td>0.42754</td>
<td>0.398</td>
</tr>
</tbody>
</table>

Please see section 5.3.2 for an explanation of the variable LAST.SECOND.7 and its factor levels.
INSTRUCTIONS
Willkommen.
In diesem Test werden Sie Sätze auf dem Monitor lesen und diese mündlich wiederholen. 
(control group: ... Fragen dazu beantworten.)

Gruppe B: zufällige Reihenfolge

Sie werden zunächst aufgefordert, einen Satz zu lesen:
"Bitte lesen Sie den folgenden Satz."
Wenn Sie die Leertaste drücken, sehen Sie ein Kreuz in der Mitte des Bildschirmes. Auf einen erneuten Druck der Leertaste wird an dieser Stelle das erste Wort eines Satzes erscheinen.
Mit jedem Druck der Leertaste erscheint ein weiteres Wort des Satzes. Wenn Sie nach dem letzten Wort die Leertaste drücken, sehen Sie wieder das Kreuz.

Um in Ihrem natürlichen Tempo zu lesen, müssen Sie die Leertaste eventuell in sehr kurzen Abständen drücken.

Wenn Sie das zweite Kreuz gesehen haben und wieder die Leertaste drücken, werden Sie aufgefordert, den gelesenen Satz zu wiederholen:
"Nach einem Tastendruck wiederholen Sie den Satz bitte mündlich."
Sie müssen die Leertaste noch einmal drücken, damit ein neuer Bildschirm mit der Meldung "Aufnahme läuft" (in rot) erscheint. Jetzt beginnt die Aufnahme und Sie wiederholen den gelesenen Satz. Mit einem weiteren Druck der Leertaste beenden Sie die Aufnahme.

(control group:)
Wenn Sie das zweite Kreuz gesehen haben, wird nach kurzer Zeit eine Frage zum eben gelesenen Satz auf dem Bildschirm eingeblendet. Diese ist mit einem Druck der Taste "Y" (für Ja) oder der Taste "N" (für Nein) zu beantworten.

Sie werden während des gesamten Tests nur die Leertaste benötigen. Sie können Ihre(n) Finger also dort belassen, um schnell reagieren zu können. Die Hand bzw. den/die Finger, mit denen Sie die Taste betätigen, können Sie frei wählen.
control group:
Sie werden während des Lesens der Sätze nur die Leertaste benötigen.
Sie können Ihre(n) Finger also dort belassen, um schnell reagieren zu können.
Die Hand bzw. den/die Finger, mit denen Sie die Taste betätigen, können Sie frei wählen.
Nur für die Beantwortung der Fragen müssen Sie zu den Tasten "Y" oder "N" wechseln.

Im Abstand von 5 Sätzen werden Sie aufgefordert, eine kurze Pause einzulegen:
"Atmen Sie kurz durch."
Wenn Sie jetzt wieder die Leertaste drücken, werden Sie aufgefordert, den nächsten Satz zu lesen.

Diese Pausen sollen Ihnen helfen, sich besser konzentrieren zu können.
Wenn Sie zusätzliche Pausen einlegen möchten, tun Sie dies bitte, bevor Sie einen neuen Satz lesen, also bei der Meldung
"Bitte lesen Sie den folgenden Satz."

Das war schon alles. Nur zur Erinnerung, eine Kurzzusammenfassung des Tests:

1. "Bitte lesen Sie den folgenden Satz"
Sie müssen die Leertaste drücken, um erst ein Kreuz, dann alle Wörter des Satzes, und zum Schluss wieder ein Kreuz zu sehen.

2. "Nach einem Tastendruck wiederholen Sie den Satz bitte mündlich."
Drücken Sie noch einmal die Leertaste. Sprechen Sie erst bei der Meldung "Aufnahme läuft"!

3. Nach dem Beenden der Aufnahme werden Sie vom Computer automatisch aufgefordert, den nächsten Satz zu lesen (und im Anschluss zu wiederholen usw.)

control group:
2. Nun wird eine Frage zum Satz eingeblendet, die Sie mit Ja (Taste Y) oder Nein (Taste N) beantworten.

3. Nach dem Beantworten der Frage werden Sie vom Computer automatisch aufgefordert, den nächsten Satz zu lesen (und die zugehörige Frage zu beantworten usw.)

Zusätzliche Pausen legen Sie bitte vor dem Lesen eines Satzes ein.

Das Ende des Tests wird anhand eines Bildschirms mit der Meldung "Fertig!" angezeigt.
Sie müssen nur noch der Versuchsleiterin Bescheid geben, dass sie fertig sind, und einen kurzen Fragebogen ausfüllen.

Wenn Sie Fragen zum Ablauf haben, stellen Sie sie jetzt oder nach dem Übungsteil.
Ein Übungssatz zur Gewöhnung...
(practice sentence #1 is presented and repeated)

Hier noch ein zweiter Übungssatz...
(practice sentence #2 is presented and repeated)

Die nächsten Sätze sind den eigentlichen Testsätzen schon sehr ähnlich.
(practice sentences #3 to 7 are presented and repeated)

Der Übungsteil ist hiermit abgeschlossen.

Wenn Sie noch Fragen haben, stellen Sie diese bitte jetzt.

Wenn Sie keine Fragen haben, können Sie mit dem Test beginnen.

Los geht's...
(practice sentences #8 to 10 are presented and repeated (questions in the control group), followed by the actual target and filler sentences)

Fertig!!
Danke für Ihre Teilnahme.

Geben Sie bitte Bescheid, dass Sie fertig sind.

STIMULUS SENTENCES
Comprehension questions for the control group (see 6.2) follow in brackets after each stimulus sentence (in the case of target sentences after the last realisation of the sentence). The correct answer is indicated by a Y for “yes” and an N for “no” immediately following the question.

PRACTICE SENTENCES
1 Dieser Satz ist ein kurzer und leichter Übungssatz.
   (Ist der Satz lang? N)
2 Dieser Satz dient auch zur Übung, und er ist leicht zu wiederholen.
   (Dient der Satz zur Übung? Y)
3 Wenn du mit deinem Partner zusammen spielst, kann dir eigentlich nichts passieren.
   (Solltest du mit deinem Partner zusammen spielen? Y)
4 Links siehst du einen Baum; hinter ihm liegt ein wichtiger Gegenstand.
   (Steht der Baum rechts? N)
5 Im Raum mit den Kisten öffne nur die dritte und du bekommst Pfeile.
   (Ist im Raum nur eine Kiste? N)
6 Deine Begleiter haben verschiedene Fähigkeiten, die du dir im Menü anschauen kannst.
(Siehst du die Eigenschaften deiner Begleiter im Menü? Y)
7 Mach dir keine Sorgen; normalerweise reicht dein Gold bis zum Ende aus.
(Musst du Angst haben, pleite zu gehen? N)
8 Von hinter der Mauer hast du eine gute Sicht, ohne selbst getroffen zu werden.
(Bist du in deinem Versteck sicher? Y)
9 Merk dir den angezeigten Weg gut, damit du dich unterwegs nicht verläufst.
(Ist es egal, wo du langläufst? N)
10 Wenn du diese Anleitung gelesen hast, solltest du das Spiel problemlos gewinnen.
(Hilft die Anleitung beim Gewinnen des Spiels? Y)

TARGET SENTENCES
1 a Dann nimm die zwei roten Magnete aus dem Regal in deine Hand.
1 b Dann nehm die zwei roten Magnete aus dem Regal in deine Hand.
1 c Dann halt die zwei roten Magnete aus dem Regal aneinander.
1 d Dann verbirg die zwei roten Magnete aus dem Regal in deiner Hand.
1 e Dann verberg die zwei roten Magnete aus dem Regal in deiner Hand.
1 f Dann kleb die zwei roten Magnete aus dem Regal zusammen.
(Sind die Magnete rot? Y)

2 a Du findest eine Kuh im Stall; damit du nicht verhungerst, iss sie schnell.
2 b Du findest eine Kuh im Stall; damit du nicht verhungerst, ess sie schnell.
2 c Du findest eine Kuh im Stall; damit du nicht verhungerst, fang sie schnell.
2 d Du findest eine Kuh im Stall; damit du nicht verdurstest, milk sie schnell.
2 e Du findest eine Kuh im Stall; damit du nicht verdurstest, melk sie schnell.
2 f Du findest Weizen auf dem Feld; damit du nicht verhungerst, mahl ihn schnell.
(Überlebt deine Figur ohne Nahrung? N)

3 a Dann triff als nächstes den Arm des Gegners, um ihn zu verwunden.
3 b Dann treff als nächstes den Arm des Gegners, um ihn zu verwunden.
3 c Dann bekom als nächstes den Arm des Gegners zu fassen, um ihn zu verwunden.
3 d Dann stich als nächstes in den Arm des Gegners, um ihn zu verwunden.
3 e Dann stech als nächstes in den Arm des Gegners, um ihn zu verwunden.
3 f Dann attackier als nächstes den Arm des Gegners, um ihn zu verwunden.
(Ist dein Ziel der Arm des Gegners? Y)

4 a Wenn du genug vom Kämpfen hast, dann wirf deinen Gegner in die Schlucht.
4 b Wenn du genug vom Kämpfen hast, dann werf deinen Gegner in die Schlucht.
4 c Wenn du genug vom Kämpfen hast, dann stoß deinen Gegner in die Schlucht.
4 d Wenn du genug vom Kämpfen hast, dann erstich deinen Gegner mit dem Schwert.
4 e Wenn du genug vom Kämpfen hast, dann erstech deinen Gegner mit dem Schwert.
4 f Wenn du genug vom Kämpfen hast, dann eliminier deinen Gegner mit dem Schwert.
(Willst du ewig weiter kämpfen? N)

5 a Bist du auf dem Friedhof angekommen, dann birg jetzt die Schriftrolle aus dem Sarg.
5 b Bist du auf dem Friedhof angekommen, dann berg jetzt die Schriftrolle aus dem Sarg.
5 c Bist du auf dem Friedhof angekommen, dann berühr jetzt den Schriftzug auf dem Sarg.
5 d Bist du auf dem Friedhof angekommen, dann überlies nicht den Schriftzug auf dem Sarg.
5 e Bist du auf dem Friedhof angekommen, dann überles nicht den Schriftzug auf dem Sarg.
5 f Bist du auf dem Friedhof angekommen, dann entroll jetzt die Schriftrolle aus dem Sarg.
(Befindest du dich auf dem Friedhof? Y)

6 a Wenn du an der Reihe bist, übertriff alle Feinde in dieser Runde.
6 b Wenn du an der Reihe bist, übertreff alle Feinde in dieser Runde.
6 c Wenn du an der Reihe bist, besieg alle Feinde in dieser Runde.
6 d Wenn du an der Reihe bist, verdrisch alle Feinde in dieser Runde.
6 e Wenn du an der Reihe bist, verdresch alle Feinde in dieser Runde.
6 f Wenn du an der Reihe bist, vermöbel alle Feinde in dieser Runde.
(Kämpfst du gegen einen einzigen Feind? N)

7 a Aber versprich deiner Truppe nur etwas, was du auch garantiert umsetzen kannst.
7 b Aber versprech deiner Truppe nur etwas, was du auch garantiert umsetzen kannst.
7 c Aber kündige deiner Truppe nur etwas an, was du auch garantiert umsetzen kannst.
7 d Aber befiehl deiner Truppe nur etwas, was du auch garantiert umsetzen kannst.
7 e Aber befehl deiner Truppe nur etwas, was du auch garantiert umsetzen kannst.
7 f Aber unterbreite deiner Truppe nur etwas, was du auch garantiert umsetzen kannst.
(Musst du eine bestimmte Art von Aktionen planen? Y)

8 a Während deine Gruppe eine Pause macht, sprich alle Fremden auf dem Platz an.
8 b Während deine Gruppe eine Pause macht, sprech alle Fremden auf dem Platz an.
8 c Während deine Gruppe eine Pause macht, fördere alle Fremden auf dem Platz heraus.
8 d Während deine Gruppe eine Pause macht, bestiehl alle Fremden auf dem Platz.
8 e Während deine Gruppe eine Pause macht, bestehl alle Fremden auf dem Platz.
8 f Bevor deine Gruppe eine Pause macht, verjage alle Fremden von dem Platz.
(Kennst du alle Leute auf dem Platz? N)

9 a Bist du an der Karte angekommen, dann sieh dir den richtigen Weg an.
9 b Bist du an der Karte angekommen, dann seh dir den richtigen Weg an.
9 c Bist du an der Karte angekommen, dann lass dir den richtigen Weg zeigen.
9 d Bist du an der Karte angekommen, dann besieh dir den richtigen Weg.
9 e Bist du an der Karte angekommen, dann beseh dir den richtigen Weg.
9 f Bist du an der Karte angekommen, dann entsperr dir den richtigen Weg.
(Wird der Weg auf einer Karte angezeigt? Y)

10 a Sobald dein Partner Mark bedrängt wird, hilf ihm schnell mit einem starken Angriff.
10 b Sobald dein Partner Mark bedrängt wird, helf ihm schnell mit einem starken Angriff.
10 c Sobald dein Partner Mark bedrängt wird, unterstütz ihn schnell mit einem starken Angriff.
10 d Sobald dein Partner Mark bedrängt wird, befehl ihm schnell einen starken Angriff.
10 e Sobald dein Partner Mark bedrängt wird, befehl ihm schnell einen starken Angriff.
10 f Sobald dein Partner Mark bedrängt wird, erlös ihn schnell mit einem starkem Angriff.
(Kommt dein Partner alleine zurecht? N)

11 a Auch wenn du nicht weiter kommst, gib die Aktion auf gar keinen Fall auf.
11 b Auch wenn du nicht weiter kommst, gib die Aktion auf gar keinen Fall auf.
11 c Auch wenn du nicht weiter kommst, lass die Aktion auf gar keinen Fall sein.
11 d Auch wenn du nicht weiter kommst, verwirf die Aktion auf gar keinen Fall.
11 e Auch wenn du nicht weiter kommst, verwerf die Aktion auf gar keinen Fall.
11 f Auch wenn du nicht weiter kommst, hak die Aktion auf gar keinen Fall ab.
(Sollst du bei der Aufgabe hartnäckig bleiben? Y)

12 a Dann tritt die Tür zur Kammer ein, um deinen Partner zu befreien.
12 b Dann tret die Tür zur Kammer ein, um deinen Partner zu befreien.
12 c Dann schlag die Tür zur Kammer ein, um deinen Partner zu befreien.
12 d Dann durchbrich die Tür zur Kammer, um deinen Partner zu befreien.
12 e Dann durchbrech die Tür zur Kammer, um deinen Partner zu befreien.
12 f Dann renn die Tür zur Kammer ein, um deinen Partner zu befreien.
(Führt die Tür zum Kerker? N)

13 a Sobald du die schwarze Truhe siehst, brich sie mit der Axt auf.
13 b Sobald du die schwarze Truhe siehst, brech sie mit der Axt auf.
13 c Sobald du die schwarze Truhe siehst, öffne sie mit der Axt.
13 d Sobald du die schwarze Truhe siehst, zerdrisch sie mit der Axt.
13 e Sobald du die schwarze Truhe siehst, zerdresch sie mit der Axt.
13 f Sobald du die schwarze Truhe siehst, zerkloppe sie mit der Axt.
(Ist die Truhe schwarz? Y)

14 a Wenn du deine Armee schnell vergrößern willst, übernimm einfach das nächste Dorf.
14 b Wenn du deine Armee schnell vergrößern willst, übernehm einfach das nächste Dorf.
14 c Wenn du deine Armee schnell vergrößern willst, besetz einfach das nächste Dorf.
14 d Wenn du deine Armee schnell vergrößern willst, unterwirf einfach das nächste Dorf.
14 e Wenn du deine Armee schnell vergrößern willst, unterwerf einfach das nächste Dorf.
14 f Wenn du deine Armee schnell vergrößern willst, erober einfach das nächste Dorf.
(Ist deine Armee groß genug? N)

15 a Um sein Vertrauen zu gewinnen, empfiehl dich dem König als treuer Berater.
15 b Um sein Vertrauen zu gewinnen, empfehl dich dem König als treuer Berater.
15 c Um sein Vertrauen zu gewinnen, erweise dich dem König als treuer Berater.
15 d Um sein Vertrauen zu gewinnen, benimm dich dem König gegenüber als treuer Berater.
15 e Um sein Vertrauen zu gewinnen, benehm dich dem König gegenüber als treuer Berater.
15 f Um sein Vertrauen zu gewinnen, mühe dich dem König gegenüber als treuer Berater ab.
(Soll der König dir vertrauen? Y)
16 a Wenn du nahe genug rankommst, stiehl einen wertvollen Gegenstand von deinem Gegner.
16 b Wenn du nahe genug rankommst, stehl einen wertvollen Gegenstand von deinem Gegner.
16 c Wenn du nahe genug rankommst, fang einen wertvollen Gegenstand von deinem Gegner.
16 d Wenn du nahe genug rankommst, zertritt einen wertvollen Gegenstand von deinem Gegner.
16 e Wenn du nahe genug rankommst, zertret einen wertvollen Gegenstand von deinem Gegner.
16 f Wenn du nahe genug rankommst, ergauner einen wertvollen Gegenstand von deinem Gegner.

(Ist der Gegenstand wertlos? N)

17 a Um dir Zutritt zum Verlies zu verschaffen, gib den Wachen am Tor Gold.
17 b Um dir Zutritt zum Verlies zu verschaffen, geb den Wachen am Tor Gold.
17 c Um dir Zutritt zum Verlies zu verschaffen, lass den Wachen Gold da.
17 d Um dir Zutritt zum Verlies zu verschaffen, bestich die Wachen am Tor.
17 e Um dir Zutritt zum Verlies zu verschaffen, bestech die Wachen am Tor.
17 f Um dir Zutritt zum Verlies zu verschaffen, täusch die Wachen am Tor.

(Willst du ins Verlies gelangen? Y)

18 a Und stirb bloß nicht, wenn der Riese angreift; sonst ist das Spiel gelaufen.
18 b Und sterb bloß nicht, wenn der Riese angreift; sonst ist das Spiel gelaufen.
18 c Und fall bloß nicht hin, wenn der Riese angreift; sonst ist das Spiel gelaufen.
18 d Und erschrick bloß nicht, wenn der Riese angreift; sonst ist das Spiel gelaufen.
18 e Und erschreck bloß nicht, wenn der Riese angreift; sonst ist das Spiel gelaufen.
18 f Und zöger bloß nicht, wenn der Riese angreift; sonst ist das Spiel gelaufen.

(Greift dich ein Zwerg an? N)

19 a Aber wenn der dritte Versuch auch erfolglos bleibt, vergiss den Auftrag lieber.
19 b Aber wenn der dritte Versuch auch erfolglos bleibt, vergess den Auftrag lieber.
19 c Aber wenn der dritte Versuch auch erfolglos bleibt, beende den Auftrag lieber.
19 d Aber wenn der dritte Versuch auch erfolglos bleibt, verwirf den Auftrag lieber.
19 e Aber wenn der dritte Versuch auch erfolglos bleibt, verwerf den Auftrag lieber.
19 f Aber wenn der dritte Versuch auch erfolglos bleibt, überspringe den Auftrag lieber.

(Kannst du nach dem dritten Versuch aufgeben? Y)

20 a Willst du von hinten ins Haus einsteigen, dann wirf die Fenster ein.
20 b Willst du von hinten ins Haus einsteigen, dann werf die Fenster ein.
20 c Willst du von hinten ins Haus einsteigen, dann zerstör die Fenster.
20 d Willst du von hinten ins Haus einsteigen, dann zerwirf die Fenster.
20 e Willst du von hinten ins Haus einsteigen, dann zerwerf die Fenster.
20 f Willst du von hinten ins Haus einsteigen, dann zerklopp die Fenster.

(Willst du das Haus durch die Tür betreten? N)
21 a Im nächsten Duell erwirb den goldenen Gürtel und du wirst endlich unverwundbar.
21 b Im nächsten Duell erwerb den goldenen Gürtel und du wirst endlich unverwundbar.
21 c Nach dem nächsten Duell empfang den goldenen Gürtel und du wirst endlich unverwundbar.
21 d Im nächsten Duell erfecht den goldenen Gürtel und du wirst endlich unverwundbar.
21 e Im nächsten Duell erfecht den goldenen Gürtel und du wirst endlich unverwundbar.
21 f Im nächsten Duell erheisch den goldenen Gürtel und du wirst endlich unverwundbar.
(Macht der Gürtel dich unverwundbar? Y)

22 a Wenn dein Team versammelt ist, unternimm den Angriff auf die feindliche Gruppe.
22 b Wenn dein Team versammelt ist, unternehm den Angriff auf die feindliche Gruppe.
22 c Wenn dein Team versammelt ist, bereite den Angriff auf die feindliche Gruppe vor.
22 d Wenn dein Team versammelt ist, ermiss den Schaden durch die feindliche Gruppe.
22 e Wenn dein Team versammelt ist, ermess den Schaden durch die feindliche Gruppe.
22 f Wenn dein Team versammelt ist, berede den Angriff auf die feindliche Gruppe.
(Bist in dieser Szene alleine? N)

23 a Als letztes entnimm der Kiste, die du siehst, die tausend kleinen Teile.
23 b Als letztes entnehm der Kiste, die du siehst, die tausend kleinen Teile.
23 c Als letztes beseitige die Kiste, die du siehst, und die tausend kleinen Teile.
23 d Als letztes zerdrisch die Kiste, die du siehst, in tausend kleine Teile.
23 e Als letztes zerdrisch die Kiste, die du siehst, in tausend kleine Teile.
23 f Als letztes zerkloppe die Kiste, die du siehst, in tausend kleine Teile.
(Führst du diese Aktion als letztes aus? Y)

24 a Danach wirf dem Gegner ohne Ende Feuerbälle zu; irgendwann muss er aufgeben.
24 b Danach werf dem Gegner ohne Ende Feuerbälle zu; irgendwann muss er aufgeben.
24 c Danach verletz den Gegner ohne Ende mit Feuerbällen; irgendwann muss er aufgeben.
24 d Danach bewirf den Gegner ohne Ende mit Feuerbällen; irgendwann muss er aufgeben.
24 e Danach bewerf den Gegner ohne Ende mit Feuerbällen; irgendwann muss er aufgeben.
24 f Danach belager den Gegner ohne Ende mit Feuerbällen; irgendwann muss er aufgeben.
(Attackierst du den Gegner mit Eisbällen? N)

25 a Stattdessen nimm ihm seinen Triumph, indem du deinen Partner zu Hilfe rufst.
25 b Stattdessen nehm ihm seinen Triumph, indem du deinen Partner zu Hilfe rufst.
25 c Stattdessen zeig ihm deinen Triumph, indem du deinen Partner zu Hilfe rufst.
25 d Stattdessen verdirb ihm seinen Triumph, indem du deinen Partner zu Hilfe rufst.
25 e Stattdessen verderb ihm seinen Triumph, indem du deinen Partner zu Hilfe rufst.
25 f Stattdessen entreiß ihm seinen Triumph, indem du deinen Partner zu Hilfe rufst.
(Lässt du dir von deinem Partner helfen? Y)

41 In order to ensure an equal plausibility of all realisations of the sentence, the word “Angriff” needed to be exchanged. The word “Schaden” matches it in token frequency (DeReWo), number of syllables and number of letters.
26 a Um nicht so viel Krach zu machen, lies die Glaskugeln ganz vorsichtig auf.
26 b Um nicht so viel Krach zu machen, lies les\textsuperscript{42} die Glaskugeln ganz vorsichtig auf.
26 c Um nicht so viel Krach zu machen, öffne die Glaskugeln ganz vorsichtig.
26 d Um nicht so viel Krach zu machen, zerbrich die Glaskugeln ganz vorsichtig.
26 e Um nicht so viel Krach zu machen, zerbrich/zerbrech die Glaskugeln ganz vorsichtig.
26 f Um nicht so viel Krach zu machen, spalte die Glaskugeln ganz vorsichtig.
(Darfst du die Kugeln unsanft behandeln? N)

27 a Damit du deine Chance nicht verpasst, unterbrich ihn gleich am Anfang der Szene.
27 b Damit du deine Chance nicht verpasst, unterbrech ihn gleich am Anfang der Szene.
27 c Damit du deine Chance nicht verpasst, beseitige ihn gleich am Anfang der Szene.
27 d Damit du deine Chance nicht verpasst, bestiehl ihn gleich am Anfang der Szene.
27 e Damit du deine Chance nicht verpasst, bestehl ihn gleich am Anfang der Szene.
27 f Damit du deine Chance nicht verpasst, entwaffne ihn gleich am Anfang der Szene.
(Musst du in dieser Szene schnell handeln? Y)

28 a Wenn kein Ausweg bleibt, ergib dich wie ein Held und du beendest das Spiel.
28 b Wenn kein Ausweg bleibt, ergeb dich wie ein Held und du beendest das Spiel.
28 c Wenn kein Ausweg bleibt, entscheide dich wie ein Held und du beendest das Spiel.
28 d Wenn kein Ausweg bleibt, befinde dich wie ein Held und du beendest das Spiel.
28 e Wenn kein Ausweg bleibt, benehm dich wie ein Held und du beendest das Spiel.
28 f Wenn kein Ausweg bleibt, gebärde dich wie ein Held und du beendest das Spiel.
(Sollst du weiter spielen? N)

29 a Sobald du auf Stufe fünf bist, tritt nur noch in Duellen gegen starke Gegner an.
29 b Sobald du auf Stufe fünf bist, tret nur noch in Duellen gegen starke Gegner an.
29 c Sobald du auf Stufe fünf bist, kämpfe nur noch in Duellen gegen starke Gegner.
29 d Sobald du auf Stufe fünf bist, ficht nur noch in Duellen gegen starke Gegner.
29 e Sobald du auf Stufe fünf bist, fecht nur noch in Duellen gegen starke Gegner.
29 f Sobald du auf Stufe fünf bist, verwickle nur noch starke Gegner in Duelle.
(Kannst du dir irgendwann die Gegner aussuchen? Y)

30 a Um deine Goldreserven aufzufüllen, brich das Schloss am Tor zur Schatzkammer auf.
30 b Um deine Goldreserven aufzufüllen, brech das Schloss am Tor zur Schatzkammer auf.
30 c Um deine Goldreserven aufzufüllen, öffne das Schloss am Tor zur Schatzkammer.
30 d Um deine Goldreserven aufzufüllen, zerbrich das Schloss am Tor zur Schatzkammer.
30 e Um deine Goldreserven aufzufüllen, zerbrech das Schloss am Tor zur Schatzkammer.
30 f Um deine Goldreserven aufzufüllen, knacke das Schloss am Tor zur Schatzkammer.
(Ist die Schatzkammer ungesichert? N)

\textsuperscript{42}Unfortunately, conditions b and e were identical with a and d for the first participants. As soon as this mistake was detected, \textit{lies} and \textit{zerbrich} in these conditions were replaced by \textit{les} and \textit{zerbrech}.
Du musst vor allem die Falltüren und die vielen Hindernisse beachten.
(Befinden sich Hindernisse auf dem Weg? Y)
Wenn man einmal den richtigen Dreh raushat, ist es danach sehr einfach.
(Wird das Spiel immer schwerer? N)
Das ganze wird noch drei mal wiederholt, und schon ist der Gegner fertig.
(Reichen drei Angriffe aus? Y)
Den anderen Gegenstand kannst du selbst mit deinem Partner nicht mehr tragen.
(Wiegt der Gegenstand wenig? N)
Nachdem die Tür aufgebrochen ist, musst du dem Gang nach rechts folgen.
(Sollst du nach rechts gehen? Y)
Du kämpfst ohne Partner; er wartet schon im nächsten Raum auf dich.
(Hilft dir dein Partner an dieser Stelle? N)
Die übrigen Befehle muss man im Hauptmenü während einer Pause eingeben.
(Musst du die Befehle ins Hauptmenü eingeben? Y)
Wenn du das immer wiederholst, bist du nach wenigen Schlägen der Gewinner.
(Reicht es aus, die Aktion einmal auszuführen? N)
Rechts oben bewacht der eine Soldat zwei Truhen, der andere einen Gefangenen.
(Sind die Soldaten rechts oben? Y)
Am Anfang des Spiels bist du in Marks Zimmer; du verlässt es nach dem Kampf.
(Bleibst du in Marks Zimmer? N)
Viel Erfolg mit dieser Komplettlösung und viel Spaß beim Spielen.
(Verhilft die Anleitung zum Erfolg im Spiel? Y)
In diesem Fall wäre es besser, gar keinen anderen Kämpfer einzusetzen.
(Sollst du andere Kämpfer einsetzen? N)
Wenn du dein erstes Spiel startest, ist ein guter Charakter besser geeignet.
(Hilft ein guter Charakter beim ersten Mal? Y)
Als nächstes solltest du zum nördlichen Turm laufen und die Treppen hoch steigen.
(Führen die Treppen auf den westlichen Turm? N)
Hinter der großen Eiche kannst du dich vor ihnen verstecken und abspeichern.
(Bist du hinter der Eiche kurzzeitig in Sicherheit? Y)
Wenn du ihn mit deinen Fäusten schlägst, bringt das leider gar nichts.
(Schaden Faustschläge dem Gegner? N)
Ein kräftiger Hieb sollte genügen, und er ist erst einmal außer Gefecht gesetzt.
(Kannst du den Gegner in Schach halten? Y)
Dein Erfolg in manchen Gefechten hängt stark davon ab, welchen Charakter du wählst.
(Ist dein Charakter egal? N)
Du hast leider keinen Einfluss darauf, wen du als nächsten Begleiter bekommst.
(Ist dein Begleiter vorbestimmt? Y)
Willst du an dieser Stelle nicht verlieren, brauchst du deinen Partner Mark.
(Heißt dein Partner Mirko? N)
Solange du den goldenen Gürtel noch nicht hast, kann es nicht klappen.
(Brauchst du den goldenen Gürtel? Y)
Diese Szene kannst du ruhig überspringen; da gewinnst du sowieso keine Punkte.
(Musst du diese Szene durchspielen? N)

23 Nachdem du diesen Teil überstanden hast, ist der Rest wirklich ein Kinderspiel.
(Wird das Spiel leichter? Y)

24 Du darfst dich aber nicht zu früh freuen, weil dich gleich wieder einer angreift.
(Wirst du gleich von jemandem belohnt? N)

25 Die billigere Rüstung tut es auch, und du sparst jede Menge Gold.
(Reicht die billigere Variante aus? Y)

26 Zuallererst kaufst du beim Händler die Ausrüstung für dich und deinen Partner.
(Kaufst du Vorräte beim Händler? N)

27 An den blau leuchtenden Kreuzen kannst du das Spiel speichern und dich ausruhen.
(Leuchten die Speicher-Kreuze blau? Y)

28 Vor dem Kampf gegen diesen Endgegner solltest du unbedingt nochmal Pause machen.
(Solltest du direkt in diesen Kampf starten? N)

29 Irgendwann kommst du auch selbst auf die Lösung; ich geb sie dir trotzdem.
(Ist die Aufgabe lösbar? Y)

30 Sogar wenn man das Spiel zum dritten Mal durchspielt, ist diese Stelle schwer.
(Wird es beim dritten Versuch leichter? N)
BACKGROUND QUESTIONNAIRE

Anne Krause
Graduiertenkolleg GRK DFG 1624 „Frequenzeffekte in der Sprache“
Belfortstraße 18, Büro 03011
79098 Freiburg

Studie zur Sprachverarbeitung und -wahrnehmung
Hintergrundfragebogen zum Test

Bitte beantworten Sie die folgenden Fragen vollständig und wahrheitsgemäß.
Bei ja/nein-Fragen bitte Zutreffendes ankreuzen, bei Inhaltsfragen so ausführlich wie möglich antworten. (Sie können hierfür auch die Rückseite des Blattes nutzen.)

1. Allgemeine Fragen zur Person:
1.1 Alter: _________________________________________________________
1.2 Geschlecht: _________________________________________________________
1.3 Wohnort:  _________________________________________________________
1.4 Aktuelle berufliche Tätigkeit: _____________________________________________

2. Fragen zum Dialekt:
2.1 Sind Sie Dialektsprecher?  □ ja  □ nein
2.2 Wenn ja, welcher Dialekt? ______________________________________________
2.3 Geburts-Bundesland  ______________________________________________
2.4 Mit wie vielen Jahren haben Sie Ihr Geburts-Bundesland verlassen? __________

3. Fragen zur Einstellung zu Sprache:
3.1 Hören Sie gerne unterschiedliche Dialekte?  □ ja  □ nein
3.2 Korrigieren Sie die Sprache ihrer Mitmenschen?  □ ja  □ nein
3.3 Wenn ja, was korrigieren sie zum Beispiel? __________________________________
4. Fragen zur Leseaufgabe:
4.1 Ist es Ihnen schwer gefallen, die Sätze zu lesen? □ ja □ nein
4.2 Wenn ja, wie sehr? □ weniger schwer □ schwer □ sehr schwer
4.3 Abgesehen vom Inhalt des Textes (Videospielanleitung), sind Ihnen in der Sprache der Sätze ungewöhnliche Formen begegnet? Wenn ja, welche?

4.4 Können Sie sich an konkrete Formulierungen (Wörter) erinnern, die Sie selbst nicht oder anders gebrauchen? Wenn ja, welche?

4.5 Haben Sie Dialektsprache gelesen? Wenn ja, welche/r Dialekt/e? (Bitte Beispiele nennen)

5. Fragen zur Wiederholungsaufgabe:
5.1 Ist es Ihnen schwer gefallen, die Sätze zu wiederholen? □ ja □ nein
5.2 Wenn ja, wie sehr? □ weniger schwer □ schwer □ sehr schwer
5.3 Haben Sie in der Wiederholung bewusst Veränderungen der präsentierten Sätze vorgenommen? Warum?

5.4 Haben Sie Sätze so wiederholt, wie sie präsentiert waren, obwohl Sie selbst sie so nicht sagen würden? Wenn ja, können Sie sich an einen konkreten Satz erinnern?

6. Fragen zum Untersuchungsgegenstand:
6.1 Haben Sie während des Tests eine Ahnung dafür entwickelt, welche sprachlichen Einheiten konkret untersucht werden? □ ja □ nein
6.2 Wenn ja, können Sie sich erinnern, an welchem Punkt im Laufe des Tests sie diese Ahnung entwickelt haben? z.B. bei welchem Satz?

6.3 Haben Sie ab diesem Punkt das Gefühl gehabt, die folgenden Sätze anders zu lesen oder zu wiederholen? Wenn ja, inwiefern?
DEBRIEFING

Studie zur Sprachverarbeitung und -wahrnehmung
Abschließende Aufklärung

In der Anleitung zum Test ist eine notwendige Täuschung eingebaut, die abschließend aufgeklärt werden muss: Die Täuschung besteht darin, dass die Testsätze rein zufällig ausgewählte Sätze sind, die nicht einem zusammenhängenden Text entstammen. Es gibt demzufolge auch keine „Gruppe A“, die die richtige Reihenfolge präsentiert bekommt; alle Teilnehmer dieses Tests sehen eine zufällige Reihenfolge. Die Erklärung aus der Anleitung, dass untersucht wird, wie Leseverhalten und Wiederholungsgenauigkeit von dieser (zufälligen oder richtigen) Reihenfolge abhängen können, sollte Sie vom eigentlichen Untersuchungsgegenstand ablenken. Dies ist übliche Praxis bei psycholinguistischen Experimenten.


Ist Ihnen im Laufe des Tests die Täuschung („Gruppe A“) bewusst geworden? ☐ja ☐nein
Können Sie aus dem Test weitere Verben nennen (außer *geben* und *nehmen*), die zur untersuchten Verbgruppe gehören? __________________________________________

Haben Sie den eigentlichen Untersuchungsgegenstand des Tests auch ohne diese Erklärung „erraten“? ☐ja ☐nein
Begegnen Ihnen die e-Formen der Verben (z.B. *geb!*, *gebe!*?) im Alltag? ☐ja ☐nein
CONSENT FORM

Studie zur Sprachverarbeitung und -wahrnehmung

Einverständniserklärung


Für die Aufzeichnung und Speicherung der Daten sowie für deren wissenschaftliche Auswertung wird Ihr Einverständnis benötigt.

Ich erlaube hiermit, dass die am …………………………. aufgezeichneten Daten wie folgt verwendet werden:

a) Speicherung in einer projekteigenen, zugangsbeschränkten Datenbank
b) Auswertung im Rahmen des Forschungsprojekts
c) Verwendung der Tonaufnahmen bei wissenschaftlichen Veranstaltungen (Vorträgen)

Ich habe die Informationen über das Forschungsvorhaben gelesen und bin mit der vorgesehenen Verarbeitung meiner Daten einverstanden.

Name (in Blockschrift)

Ort, Datum Unterchrift

Anonymisierte Teilnehmernummer
F Materials for the pilot experiment

INSTRUCTIONS

Willkommen.


Gruppe B: zufällige Reihenfolge

Sie werden zunächst aufgefordert, einen Satz zu lesen:
"Bitte lesen Sie den folgenden Satz."
Wenn Sie die Leertaste drücken, sehen Sie ein Kreuz in der Mitte des Bildschirms.
Auf einen erneuten Druck der Leertaste wird an dieser Stelle das erste Wort eines Satzes erscheinen.
Mit jedem Druck der Leertaste erscheint ein weiteres Wort des Satzes.
Wenn Sie nach dem letzten Wort die Leertaste drücken, sehen Sie wieder das Kreuz.

Versuchen Sie, in Ihrem natürlichen Tempo zu lesen, und so, dass Sie verstehen, was Sie lesen.

Wenn Sie das zweite Kreuz gesehen haben und wieder die Leertaste drücken, werden Sie aufgefordert, den gelesenen Satz zu wiederholen:
"Nach einem Tastendruck wiederholen Sie den Satz bitte mündlich."
Sie müssen die Leertaste noch einmal drücken, damit ein neuer Bildschirm mit der Meldung "Aufnahme läuft" (in rot) erscheint. Jetzt beginnt die Aufnahme. Mit einem weiteren Druck der Leertaste beenden Sie die Aufnahme.

Es kommt hierbei nicht darauf an, dass Sie den Satz wortwörtlich wiederholen; versuchen Sie jedoch, sich so genau wie möglich an das Original zu halten. Bitte versuchen Sie, in das Mikrophon zu sprechen.

Sie werden während des gesamten Tests nur die Leertaste benötigen. Sie können Ihre(n) Finger also dort belassen, um schnell reagieren zu können. Die Hand bzw. den/die Finger, mit denen Sie die Taste betätigen, können Sie frei wählen.

In gewissen Abständen werden Sie aufgefordert, eine kurze Pause einzulegen:
"Atmen Sie kurz durch."
Auf einen Druck der Leertaste erscheint eine neue Meldung auf dem Bildschirm:
"Und weiter geht's...".
Wenn Sie jetzt wieder die Leertaste drücken, werden Sie aufgefordert, den nächsten Satz zu lesen.

Diese Pausen sollen Ihnen helfen, sich besser konzentrieren zu können.
Wenn Sie zusätzliche Pausen einlegen möchten, tun Sie dies bitte, bevor Sie einen neuen Satz lesen, also bei der Meldung
"Bitte lesen Sie den folgenden Satz."

Das war schon alles. Nur zur Erinnerung, eine Kurzzusammenfassung des Tests:

1. Sie werden aufgefordert, einen Satz zu lesen.
Sie müssen die Leertaste drücken, um erst ein Kreuz, dann alle Wörter des Satzes, und zum Schluss wieder ein Kreuz zu sehen.

2. Lesen Sie in Ihrem natürlichen Tempo.
Lesen Sie nur laut, wenn Sie das sonst auch immer tun.

3. Sie werden aufgefordert, den gelesenen Satz zu wiederholen - nicht unbedingt wortwörtlich, aber möglichst genau. Sprechen Sie erst bei der Meldung "Aufnahme läuft"!

4. Nach dem Beenden der Aufnahme werden Sie vom Computer automatisch aufgefordert, den nächsten Satz zu lesen (und im Anschluss zu wiederholen usw.)

5. In Abständen sind Pausen in den Test eingebaut.
Zusätzliche Pausen legen Sie bitte vor dem Lesen eines Satzes ein.

Das Ende des Tests wird anhand eines Bildschirms mit der Meldung "Fertig!" angezeigt. Sie müssen nun nur noch der Versuchsleiterin Bescheid geben, dass sie fertig sind, und einen kurzen Fragebogen ausfüllen.

Wenn Sie Fragen zum Ablauf haben, stellen Sie diese der Versuchsleiterin jetzt.
(Auch nach dem Übungsteil haben Sie nochmals die Möglichkeit, Fragen zu stellen.)

Ein Übungssatz zur Gewöhnung...
(practice sentence #1 is presented and repeated.)

Hier noch ein zweiter Übungssatz...
(practice sentence #2 is presented and repeated.)

Die nächsten Sätze sind den eigentlichen Testsätzen schon sehr ähnlich.
(practice sentences #3 to 5 are presented and repeated.)

Der Übungsteil ist hiermit abgeschlossen.
Wenn Sie noch Fragen haben, stellen Sie diese bitte jetzt der Versuchsleiterin.

Wenn Sie keine Fragen haben, können Sie jetzt mit dem Test beginnen.

Los geht's...

(actual target and filler sentences are presented and repeated)

Fertig!!
Danke für Ihre Teilnahme.

Geben Sie bitte der Versuchsleiterin Bescheid, dass Sie fertig sind.

STIMULUS SENTENCES

PRACTICE SENTENCES

1 Dieser Satz ist ein kurzer und leichter Übungssatz.
2 Dieser Satz dient auch zur Übung und er ist leicht zu wiederholen.
3 Wenn du mit Herkules als Held spielst, kann dir eigentlich nichts passieren.
4 Links siehst du einen Baum; in einem Astloch ist ein wichtiger Gegenstand.
5 Im Raum mit den Kisten öffne nur die dritte und du bekommst Pfeile.

TARGET SENTENCES

1 a Dann nimm die zwei Magnete in deiner Tasche und leg sie übereinander.
1 b Dann nehme die zwei Magnete in deiner Tasche und leg sie übereinander.
1 c Dann nehme die zwei Magnete in deiner Tasche und leg sie übereinander.
1 d Dann finde die zwei Magnete in deiner Tasche und leg sie übereinander.
1 e Dann verbirg die zwei Magnete in deiner Tasche und leg sie übereinander.
1 f Dann verberge die zwei Magnete in deiner Tasche und leg sie übereinander.
1 g Dann verberge die zwei Magnete in deiner Tasche und leg sie übereinander.
1 h Dann bedenke die zwei Magnete in deiner Tasche und leg sie übereinander.

2 a Du findest eine Kuh im Stall; um nicht zu verhungern, iss sie.
2 b Du findest eine Kuh im Stall; um nicht zu verhungern, esse sie.
2 c Du findest eine Kuh im Stall; um nicht zu verhungern, esse sie.
2 d Du findest eine Kuh im Stall; um nicht zu verhungern, fang sie.
2 e Du findest eine Kuh im Stall; um nicht zu verdursten, melk sie.
2 f Du findest eine Kuh im Stall; um nicht zu verdursten, melke sie.
2 g Du findest eine Kuh im Stall; um nicht zu verdursten, melke sie.
2 h Du findest Weizen auf dem Feld; um nicht zu verhungern, mahl ihn.

3 a Dann stich ein paar mal mit dem Messer rein, um es zu töten.
3 b Dann stech ein paar mal mit dem Messer rein, um es zu töten.
3 c Dann steche ein paar mal mit dem Messer rein, um es zu töten.
3 d Dann feuere ein paar mal noch das Messer rein, um es zu töten.
3 e Dann triff ein paar mal mit dem Messer rein, um es zu töten.
3 f Dann treffe ein paar mal mit dem Messer rein, um es zu töten.
3 g Dann treffe ein paar mal mit dem Messer rein, um es zu töten.
3 h Dann setz ein paar mal noch das Messer ein, um es zu töten.

4 a Wenn du genug vom Kämpfen hast, dann wirf deine Gegner in die Lava.
4 b Wenn du genug vom Kämpfen hast, dann werf deine Gegner in die Lava.
4 c Wenn du genug vom Kämpfen hast, dann werfe deine Gegner in die Lava.
4 d Wenn du genug vom Kämpfen hast, dann drohe deinen Gegnern mit der Lanze.
4 e Wenn du genug vom Kämpfen hast, dann erstich deine Gegner mit der Lanze.
4 f Wenn du genug vom Kämpfen hast, dann erstech deine Gegner mit der Lanze.
4 g Wenn du genug vom Kämpfen hast, dann ersteche deine Gegner mit der Lanze.
4 h Wenn du genug vom Kämpfen hast, dann drängel deine Gegner in die Lava.

5 a Bist du auf dem Friedhof angekommen, lies jetzt den Schriftzug auf dem Grabstein.
5 b Bist du auf dem Friedhof angekommen, les jetzt den Schriftzug auf dem Grabstein.
5 c Bist du auf dem Friedhof angekommen, lese jetzt den Schriftzug auf dem Sarg.
5 d Bist du auf dem Friedhof angekommen, hol jetzt die Schriftrolle aus dem Sarg.
5 e Bist du auf dem Friedhof angekommen, birg jetzt die Schriftrolle aus dem Sarg.
5 f Bist du auf dem Friedhof angekommen, berg jetzt die Schriftrolle aus dem Sarg.
5 g Bist du auf dem Friedhof angekommen, berge jetzt die Schriftrolle aus dem Sarg.
5 h Bist du auf dem Friedhof angekommen, berühr jetzt den Schriftzug auf dem Sarg.

6 a Wenn du an der Reihe bist, sieh dir den übrigen Haufen an.
6 b Wenn du an der Reihe bist, seh dir den übrigen Haufen an.
6 c Wenn du an der Reihe bist, sehe dir den übrigen Haufen an.
6 d Wenn du an der Reihe bist, geh gleich zu dem übrigen Haufen.
6 e Wenn du an der Reihe bist, begib dich zu dem übrigen Haufen.
6 f Wenn du an der Reihe bist, begeb dich zu dem übrigen Haufen.
6 g Wenn du an der Reihe bist, begebe dich zu dem übrigen Haufen.
6 h Wenn du an der Reihe bist, schau dir den übrigen Haufen an.

7 a Und vergiss nicht, immer wenn du eine Nachricht bekommst, darauf zu antworten.
7 b Und vergess nicht, immer wenn du eine Nachricht bekommst, darauf zu antworten.
7 c Und vergesse nicht, immer wenn du eine Nachricht bekommst, darauf zu antworten.
7 d Und diskutier nicht immer, wenn du eine Nachricht bekommst, statt darauf zu antworten.
7 e Und erschrick nicht immer, wenn du eine Nachricht bekommst, statt darauf zu antworten.
7 f Und erschreck nicht immer, wenn du eine Nachricht bekommst, statt darauf zu antworten.
7 g Und erschrecke nicht immer, wenn du eine Nachricht bekommst, statt darauf zu antworten.
7 h Und versäum nicht, immer wenn du eine Nachricht bekommst, darauf zu antworten.

8 a Auf dem Platz macht ihr eine Pause; nun sprich dort mit allen Leuten.
8 b Auf dem Platz macht ihr eine Pause; nun sprech dort mit allen Leuten.
8 c Auf dem Platz macht ihr eine Pause; nun spreche dort mit allen Leuten.
8 d Auf dem Platz macht ihr eine Pause; nun spiel dort mit allen Leuten.
8 e Auf dem Platz macht ihr eine Pause; nun bestiehlt dort alle Leute.
8 f Auf dem Platz macht ihr eine Pause; nun bestehl dort alle Leute.
8 g Auf dem Platz macht ihr eine Pause; nun bestehle dort alle Leute.
Auf dem Platz macht ihr eine Pause; nun zock dort alle Leute ab.

Nach der Kreuzung nimm den westlichen Weg; dort findest du einen Hammer.
Nach der Kreuzung nehmt den westlichen Weg; dort findest du einen Hammer.
Nach der Kreuzung nehme den westlichen Weg; dort findest du einen Hammer.
Nach der Kreuzung bleib auf dem westlichen Weg; dort findest du einen Hammer.
Nach der Kreuzung betrét den westlichen Weg; dort findest du einen Hammer.
Nach der Kreuzung betret den westlichen Weg; dort findest du einen Hammer.
Nach der Kreuzung betrete den westlichen Weg; dort findest du einen Hammer.
Nach der Kreuzung wandere auf dem westlichen Weg; dort findest du einen Hammer.

Sobald Mark vom Gegner bedrängt wird, hilf ihm mit einem starken Angriff.
Sobald Mark vom Gegner bedrängt wird, helf ihm mit einem starken Angriff.
Sobald Mark vom Gegner bedrängt wird, helfe ihm mit einem starken Angriff.
Sobald Mark vom Gegner bedrängt wird, schlage ihn mit einem starken Angriff.
Sobald Mark vom Gegner bedrängt wird, befiehl ihm einen starken Angriff.
Sobald Mark vom Gegner bedrängt wird, befehl ihm einen starken Angriff.
Sobald Mark vom Gegner bedrängt wird, befehle ihm einen starken Angriff.
Sobald Mark vom Gegner bedrängt wird, blende ihn mit einem starkem Angriff.

Auch wenn du nicht weiter kommst, gib das Spiel niemals an dieser Stelle auf.
Auch wenn du nicht weiter kommst, geb das Spiel niemals an dieser Stelle auf.
Auch wenn du nicht weiter kommst, gebe das Spiel niemals an dieser Stelle auf.
Auch wenn du nicht weiter kommst, mach das Spiel niemals an dieser Stelle aus.
Auch wenn du nicht weiter kommst, unterbrich das Spiel niemals an dieser Stelle.
Auch wenn du nicht weiter kommst, unterbreche das Spiel niemals an dieser Stelle.
Auch wenn du nicht weiter kommst, wiederhole das Spiel niemals an dieser Stelle.

Dann tritt durch die Tür zum Balkon, wo sie auf dich wartet.
Dann tret durch die Tür zum Balkon, wo sie auf dich wartet.
Dann trete durch die Tür zum Balkon, wo sie auf dich wartet.
Dann lauf durch die Tür zum Balkon, wo sie auf dich wartet.
Dann durchbrich die Tür zum Balkon, wo sie auf dich wartet.
Dann durchbrech die Tür zum Balkon, wo sie auf dich wartet.
Dann eile durch die Tür zum Balkon, wo sie auf dich wartet.

Solltest du eine Truhe sehen, brich sie auf und du bekommst neue Waffen.
Solltest du eine Truhe sehen, brech sie auf und du bekommst neue Waffen.
Solltest du eine Truhe sehen, breche sie auf und du bekommst neue Waffen.
Solltest du eine Truhe sehen, sammle sie auf und du bekommst neue Waffen.
Solltest du eine Truhe sehen, schmilz sie auf und du bekommst neue Waffen.
Solltest du eine Truhe sehen, schmelz sie auf und du bekommst neue Waffen.
Solltest du eine Truhe sehen, zerlege sie auch und du bekommst neue Waffen.
Wenn du deine Armee schnell vergrößern willst, übernimm einfach das nächste Dorf.
Wenn du deine Armee schnell vergrößern willst, übernehm einfach das nächste Dorf.
14 c Wenn du deine Armee schnell vergrößern willst, übernehme einfach das nächste Dorf.
14 d Wenn du deine Armee schnell vergrößern willst, schlage einfach das nächste Dorf.
14 e Wenn du deine Armee schnell vergrößern willst, unterwirf einfach das nächste Dorf.
14 f Wenn du deine Armee schnell vergrößern willst, unterwerf einfach das nächste Dorf.
14 g Wenn du deine Armee schnell vergrößern willst, unterwerfe einfach das nächste Dorf.
14 h Wenn du deine Armee schnell vergrößern willst, erobere einfach das nächste Dorf.

15 a Dann triff zwei Tische in einer Runde, um deinen Gegner zu beeindrucken.
15 b Dann treffe zwei Tische in einer Runde, um deinen Gegner zu beeindrucken.
15 c Dann treffe zwei Tische in einer Runde, um deinen Gegner zu beeindrucken.
15 d Dann gewinne zwei Tische in einer Runde, um deinen Gegner zu beeindrucken.
15 e Dann zerbrich zwei Tische in einer Runde, um deinen Gegner zu beeindrucken.
15 f Dann zerbrech zwei Tische in einer Runde, um deinen Gegner zu beeindrucken.
15 g Dann zerbreche zwei Tische in einer Runde, um deinen Gegner zu beeindrucken.
15 h Dann meistere zwei Tische in einer Runde, um deinen Gegner zu beeindrucken.

16 a Wenn du nahe genug rankommst, stiehl einen wertvollen Gegenstand von deinem Gegner.
16 b Wenn du nahe genug rankommst, stehl einen wertvollen Gegenstand von deinem Gegner.
16 c Wenn du nahe genug rankommst, stehle einen wertvollen Gegenstand von deinem Gegner.
16 d Wenn du nahe genug rankommst, fange einen wertvollen Gegenstand von deinem Gegner.
16 e Wenn du nahe genug rankommst, zerritt einen wertvollen Gegenstand von deinem Gegner.
16 f Wenn du nahe genug rankommst, zertret einen wertvollen Gegenstand von deinem Gegner.
16 g Wenn du nahe genug rankommst, zertrete einen wertvollen Gegenstand von deinem Gegner.
16 h Wenn du nahe genug rankommst, zermalme einen wertvollen Gegenstand von deinem Gegner.

FILLER SENTENCES

1 Du musst vor allem die Falltüren und die vielen Hindernisse beachten.
2 Wenn man weiß, wie es geht, ist es wie immer sehr einfach.
3 Das ganze wird noch drei mal wiederholt, und schon ist der Gegner fertig.
4 Wenn du dein erstes Spiel startest, ist ein kleines Gebiet besser geeignet.
5 Nachdem die Tresortür aufgebrochen ist, musst du dem Gang nach rechts folgen.
6 Du kämpfst ohne Partner; er wartet schon im nächsten Raum auf dich.
7 Die übrigen Befehle muss man im Hauptmenü eingeben, während man Pause macht.
8 Wenn du das immer wiederholst, bist du nach wenigen Schlägen der Gewinner.
9 Rechts oben bewacht der eine Soldat zwei Truhen, der andere einen Gefangenen.
10 Am Anfang des Spiels bist du in Marks Zimmer; du verlässt es nach dem Kampf.
11 Viel Erfolg mit dieser Komplettlösung und viel Spaß beim Spielen.
12 In diesem Fall wäre es besser, gar keinen anderen Kämpfer einzusetzen.
13 Den anderen Gegenstand kannst du selbst mit deinem Partner nicht mehr tragen.
14 Als nächstes solltest du zum nördlichen Turm laufen und die Treppen hoch steigen.
15 Hinter der großen Eiche kannst du dich vor ihnen verstecken und abspeichern.
16 Wenn du ihn mit deinen Fäusten schlägst, bringt das leider gar nichts.
G Corpus frequencies of variants of the imperative singular and the second and third person singular of strong verbs with e/i-gradation

Table 39: Token frequencies and proportions of traditional and analogical variants of three paradigm forms of strong verbs with e/i-gradation in the corpus Wortschatz Universität Leipzig

<table>
<thead>
<tr>
<th>verb</th>
<th>imp sg</th>
<th>2nd sg</th>
<th>3rd sg</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>i</td>
<td>e</td>
<td>i</td>
</tr>
<tr>
<td>essen</td>
<td>124 3</td>
<td>110 0</td>
<td>2,845 0</td>
</tr>
<tr>
<td>helfen</td>
<td>399 6</td>
<td>75 0</td>
<td>27,162 1</td>
</tr>
<tr>
<td>lesen</td>
<td>385 29*</td>
<td>108 0</td>
<td>10,208 5</td>
</tr>
<tr>
<td>nehmen</td>
<td>1,020 21*</td>
<td>594 0</td>
<td>71,997 1</td>
</tr>
<tr>
<td>sehen</td>
<td>376 8*</td>
<td>1,651 1 (0.06 %)</td>
<td>140,973 3 (0.002 %)</td>
</tr>
<tr>
<td>sprechen</td>
<td>322 5</td>
<td>460 0</td>
<td>48,399 2</td>
</tr>
<tr>
<td>stechen</td>
<td>1 1 (50.00 %)</td>
<td>5 0</td>
<td>1,499 0</td>
</tr>
<tr>
<td>stellen</td>
<td>3 0</td>
<td>6 0</td>
<td>627 3 (0.476 %)</td>
</tr>
<tr>
<td>sterben</td>
<td>347 3 (0.86 %)</td>
<td>83 0</td>
<td>6,244 2 (0.032 %)</td>
</tr>
<tr>
<td>treffen</td>
<td>50 6 (10.71 %)</td>
<td>158 0</td>
<td>36,829 1 (0.003 %)</td>
</tr>
<tr>
<td>tremen</td>
<td>240 10 (4.00 %)</td>
<td>67 0</td>
<td>26,366 6 (0.023 %)</td>
</tr>
<tr>
<td>werfen</td>
<td>93 2* (2.11 %)</td>
<td>68 0</td>
<td>15,670 0</td>
</tr>
</tbody>
</table>

Table 40: Token frequencies and proportions of traditional and analogical variants of three paradigm forms of strong verbs with e/i-gradation in the FOLK corpus

<table>
<thead>
<tr>
<th>verb</th>
<th>i</th>
<th>e</th>
<th>2nd sg</th>
<th>i</th>
<th>e</th>
<th>3rd sg</th>
</tr>
</thead>
<tbody>
<tr>
<td>essen</td>
<td>2 6 (75.00 %)</td>
<td>13 0</td>
<td>35 0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>geben</td>
<td>198 13 (6.16 %)</td>
<td>49 0</td>
<td>2,569 1 (0.04 %)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>helfen</td>
<td>8 3 (27.27 %)</td>
<td>2 1 (33.33 %)</td>
<td>51 1 (1.92 %)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lesen</td>
<td>10 9 (47.37 %)</td>
<td>15 0</td>
<td>40 1 (2.44 %)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nehmen</td>
<td>135 9 (6.25 %)</td>
<td>80 0</td>
<td>218 0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sehen</td>
<td>17 1 (5.56 %)</td>
<td>293 0</td>
<td>780 0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sprechen</td>
<td>4 0 (13.50 %)</td>
<td>25 1 (3.85 %)</td>
<td>175 0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tremen</td>
<td>0 18 (100.00 %)</td>
<td>24 0</td>
<td>31 1 (3.13 %)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 41: Token frequencies and proportions of traditional and analogical variants of three paradigm forms of strong verbs with e/i-gradation in the moca database

<table>
<thead>
<tr>
<th>verb</th>
<th>i</th>
<th>e</th>
<th>2nd sg</th>
<th>i</th>
<th>e</th>
<th>3rd sg</th>
</tr>
</thead>
<tbody>
<tr>
<td>essen</td>
<td>1 1 (50.00 %)</td>
<td>24 0</td>
<td>36 0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>geben</td>
<td>65 2 (2.99 %)</td>
<td>26 0</td>
<td>4,054 0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lesen</td>
<td>7 1 (12.50 %)</td>
<td>35 0</td>
<td>48 0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nehmen</td>
<td>31 1 (3.13 %)</td>
<td>72 0</td>
<td>169 0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tremen</td>
<td>0 1 (100.00 %)</td>
<td>0 0</td>
<td>8 0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vergessen</td>
<td>6 1 (14.29 %)</td>
<td>9 0</td>
<td>14 0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>werfen</td>
<td>2 2 (50.00 %)</td>
<td>1 0</td>
<td>4 0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note to Tables 39 to 41: imp sg = imperative singular, 2nd sg = second person singular present indicative, 3rd sg = third person singular present indicative
Zusammenfassung


In einer zweiten Korpusstudie wurde untersucht, welche Faktoren das Auftreten der unterschiedlichen Formen des Imperativ Singular von starken Verben mit e/i-Wechsel beeinflussen. Aufgrund der Ergebnisse aus früheren gebrauchsbasierten Studien zu analogischem Wandel (Hooper 1976; Janda et al. 2010) wurde vermutet, dass ein sogenannter Conserving Effect (Bybee & Thompson 1997: 380) im Wandel des Imperativs von starken Verben mit e/i-Wechsel beobachtet werden kann; diese Annahme wird durch die Auswertung der Korpusdaten bestätigt: Vor allem niedrigfrequente Verben treten mit dem analogisch gebildeten Imperativ auf, während die Imperative höherfrequenter Verben in der traditionellen Flexion konservert werden. Außerdem zeigen die Ergebnisse dieser Studie, dass es sich bei dem vorliegenden Phänomen nicht bloß um reine Variation in der Bildung der Form des Imperativ Singular handelt, sondern um einen „change-in-progress“, d.h. einen beginnenden Wandel von der etablierten Bildung mit dem Stammvokal i (stirb) hin zu Formen mit dem Stammvokal e (sterb(e)). Schließlich wird jedoch auch unterstrichen, dass Analysen eines solchen „change-
in-progress“ Faktoren miteinbeziehen sollten, die typischerweise Einfluss auf linguistische Variation haben. Im vorliegenden Fall erklären Persistenz-Effekte (Szmrecsanyi 2005; 2006), d.h. die Verwendung von suffigierten Formen des Imperativ Singular im vorausgehenden Kontext, z.B. laufe und verlasse, die Wahrscheinlichkeit für die Suffigierung einer folgenden Form des Imperativ Singular eines starken Verbs mit e/i-Wechsel, z.B. gebe.


In der dritten und letzten Korpusstudie wird anhand von Daten aus drei schriftlichen und mündlichen Korpora des Deutschen (moca, FOLK, Wortschatz Universität Leipzig) erklärt, warum die zweite und dritte Person Singular starker Verben mit e/i-Wechsel, die den

Zusammenfassend konnten in der vorliegenden Dissertation Frequenzeffekte im morphologischen Wandel für die Sprachproduktion mithilfe mehrerer Korpusstudien bestätigt (Conserving Effect von Verbtokenfrequenz) oder widerlegt (intraparadigmatische Frequenzeffekte) werden. Daten über die Verarbeitung und Wahrnehmung der unterschiedlichen Imperativvarianten aus einem psycholinguistischen Experiment liefern erstmals Evidenz für die vorherrschende kognitive Erklärung des Conserving Effects im Bereich des analogischen Ausgleichs.