

# Wertschätzung sprachlicher Variation

Eine Untersuchung zur  
Veränderbarkeit von  
Spracheinstellungen im  
Deutschunterricht

Eugen Unterberger

Thema Sprache –  
Wissenschaft für den Unterricht



| wbv

# **Wertschätzung sprachlicher Variation**

Eine Untersuchung zur Veränderbarkeit von  
Spracheinstellungen im Deutschunterricht

Eugen Unterberger

## **Reihe „Thema Sprache – Wissenschaft für den Unterricht“**

Ziel dieser Reihe ist die Auslotung des wissenschaftlichen Potenzials, das eine Beschäftigung mit Sprache in Bezug auf schulische Kontexte hat. Dabei wird zum einen gefragt, wie Anwendungsfelder und Erkenntnisse der wissenschaftlichen Disziplinen Sprachwissenschaft, Sprachlehrforschung und den Sprachdidaktiken gewinnbringend für schulische Ziele in die Unterrichtspraxis übertragen werden können, zum anderen, welche Impulse aus dem Unterricht für die wissenschaftlichen Disziplinen ausgehen könnten.

Die Reihe unterliegt dem doppelt blinden Peer-Reviewverfahren.

Die Reihe wird herausgegeben von **Anja Binanzer** (Technische Universität Dresden), **Miriam Langlotz** (Universität Kassel) und **Verena Wecker** (Westfälische Wilhelms-Universität Münster).

Wissenschaftlicher Beirat:

Ursula Bredel (Hildesheim), Doreen Bryant (Tübingen), Nicole Marx (Köln), Anja Müller (Mainz), Iris Rautenberg (Ludwigsburg), Claudia Riemer (Bielefeld), Michael Rödel (München), Björn Rothstein (Bochum), Rosemarie Tracy (Mannheim), Constanze Weth (Luxembourg)

Eugen Unterberger

# Wertschätzung sprachlicher Variation

Eine Untersuchung zur Veränderbarkeit von  
Spracheinstellungen im Deutschunterricht



| wbv

Publikationskosten gefördert durch die Paris Lodron Universität Salzburg

Originaltitel der Dissertation: „Lassen sich Spracheinstellungen im Zuge des Deutschunterrichts verändern?  
Eine Interventionsstudie im Land Salzburg“

Die Dissertation, die dieser Publikation zu Grunde liegt, ist im Zuge des vom Land Salzburg geförderten Projekts ‚Sprachliche Vielfalt, verstehen, wertschätzen und ausbauen (Zahl: 20204-WISS/244/3–2019)‘ entstanden.

Ein Schneider-Titel bei  
wbv Publikation  
2024 wbv Publikation  
ein Geschäftsbereich der  
wbv Media GmbH & Co. KG, Bielefeld

Gesamtherstellung:  
wbv Media GmbH & Co. KG  
Auf dem Esch 4, 33619 Bielefeld,  
service@wbv.de **wbv.de**

Umschlaggestaltung:  
Christiane Zay, Passau

ISBN Print: 978-3-7639-7757-4  
ISBN E-Book: 978-3-7639-7758-1  
DOI: 10.3278/9783763977581

Printed in Germany

Diese Publikation ist frei verfügbar zum Download unter  
[wbv-open-access.de](http://wbv-open-access.de)

Diese Publikation ist unter folgender Creative-Commons-  
Lizenz veröffentlicht:  
[creativecommons.org/licenses/by-sa/4.0/deed.de](http://creativecommons.org/licenses/by-sa/4.0/deed.de)



Für alle in diesem Werk verwendeten Warennamen  
sowie Firmen- und Markenbezeichnungen können  
Schutzrechte bestehen, auch wenn diese nicht als solche  
gekennzeichnet sind. Deren Verwendung in diesem Werk  
berechtigt nicht zu der Annahme, dass diese frei verfü-  
gbar seien.

---

#### **Bibliografische Information der Deutschen Nationalbibliothek**

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie;  
detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

---

Die freie Verfügbarkeit der E-Book-Ausgabe dieser Publikation wurde ermöglicht durch ein Netzwerk wissenschaftlicher Bibliotheken und Institutionen zur Förderung von Open Access in den Sozial- und Geisteswissenschaften im Rahmen der *wbv OpenLibrary 2024*.

Die Publikation beachtet unsere Qualitätsstandards für Open-Access-Publikationen, die an folgender Stelle nachzulesen sind:

[https://www.wbv.de/fileadmin/importiert/wbv/PDF\\_Website/Qualitaetsstandards\\_wbvOpenAccess.pdf](https://www.wbv.de/fileadmin/importiert/wbv/PDF_Website/Qualitaetsstandards_wbvOpenAccess.pdf)

Großer Dank gebührt den Förderern der *wbv OpenLibrary 2024* im Fachbereich *Lehramt*:

Humboldt-Universität zu **Berlin** | Universitätsbibliothek **Bielefeld** | Pädagogische Hochschule **Freiburg** | Justus-Liebig-Universität **Gießen** | TIB **Hannover** | Universitätsbibliothek **Kaiserslautern-Landau** | **Karlsruhe** Institute of Technology (KIT) | Universitätsbibliothek **Kassel** | Universitäts- und Stadtbibliothek **Köln** | Zentral- und Hochschulbibliothek (ZHB, **Luzern**) | Fachhochschule **Münster** | Universitäts- und Landesbibliothek **Münster** | Carl von Ossietzky Universität **Oldenburg** | Universitätsbibliothek **Osnabrück** | Universität **Potsdam** | Universität **Regensburg**



# Danksagung

Eine Dissertation zu schreiben ist nicht einfach. In meinem Fall hat es bedeutet, mich mit Statistik auseinanderzusetzen, Programmieren zu lernen, eine wissenschaftliche Untersuchung zu planen und durchzuführen (und da es sich um ein Schulprojekt handelte, auch alles behördlich anmelden zu lassen), Unterrichtsmaterialien zu erstellen, Lehrkräfte zu schulen, Veranstaltungen zu planen und durchzuführen und kreative Lösungen zu finden, wenn etwas vollkommen Unvorhergesehenes passiert (wie eine Coronapandemie).

Daneben bedarf es aber auch guter Organisation und viel Durchhaltevermögen. Diese Arbeit wird immens dadurch erleichtert, wenn man Personen hat, an die man sich wenden kann, und zwar egal, ob es sich um eine fachliche Meinung handelt oder ein Gespräch, um die Seele zu erleichtern. Ich hatte das Glück, dass ich viele Personen in meiner Dissertationsphase um mich hatte, die mich unterstützten. Zum einen sind das die Kolleginnen aus dem Projekt, in dem die Dissertation entstanden ist – Barbara Hauser, Simone Kalchgruber und Magdalena Schobersberger. Besonders danken möchte ich Cordula Pribyl-Resch, mit der ich im Projekt eng zusammengearbeitet habe und die immer ein offenes Ohr für mich hatte.

Daneben möchte ich meinem Betreuerteam, Andrea Ender, Irmtraud Kaiser und Stephan Elspaß, danken. Auch wenn im Unileben oft viel zu tun und nur wenig Zeit ist, konnte ich mich immer an sie wenden. Andrea Ender und Irmtraud Kaiser möchte ich noch zusätzlich für ihren Einsatz danken, dass das Projekt zustande kam, dass sie viel Zeit zur Überwindung bürokratischer Hürden aufwendeten, vor allem aber dafür, dass die Zusammenarbeit im Projekt immer angenehm und konstruktiv war – ich habe diese Zeit sehr genossen!

Auch meine Kinder, Elias, David und Belena, sollen nicht unerwähnt bleiben. Für sie war es sicherlich spannend zu sehen, dass ihr Papa eine Dissertation schreibt, und mitzubekommen, was damit alles in Verbindung steht. Ich möchte ganz besonders meiner Frau Katrin danken, die in dieser Phase mit mir mitgefiebert hat und mit der ich viele Ideen (zum Teil sehr oft) diskutieren konnte – manche Bereiche der Dissertation kennt sie sicherlich genauso gut wie ich.

Zu guter Letzt möchte ich allen Lehrkräften und Partnerschulen danken, ohne die das Projekt nicht durchführbar gewesen wäre, sowie dem Land Salzburg für die Finanzierung des Projekts „Sprachliche Vielfalt verstehen, wertschätzen und ausbauen (Zahl: 20204-WISS/244/3–2019)“, in dem die Dissertation entstanden ist.



# Inhalt

|   |           |
|---|-----------|
| Zusammenfassung .....   | 13        |
| Abstract .....  | 14        |
| <b>1 Veränderung von Spracheinstellungen im Zuge des Deutschunterrichts – warum die Mühe? .....</b> | <b>15</b> |
| <b>2 Regionale und soziale Sprachvariation im Untersuchungsraum .....</b>                           | <b>19</b> |
| 2.1 Regionale Variation .....   | 19        |
| 2.2 Historische Hintergründe sozialer Sprachvariation .....   | 20        |
| 2.3 Sprachgebrauch heute .....  | 22        |
| <b>3 Dimensionen von Spracheinstellungen .....</b>  | <b>25</b> |
| 3.1 Grundlegende Eigenschaften von Einstellungen .....  | 25        |
| 3.2 Spezifika von Spracheinstellungen .....   | 27        |
| 3.3 Spracheinstellungen, soziale Gruppen und Identität .....  | 28        |
| 3.3.1 Social Identity Theory .....  | 28        |
| 3.3.2 Kontextualisierungstheorie .....  | 31        |
| 3.3.3 Akkommodationstheorie .....   | 33        |
| 3.3.4 Audience Design .....   | 35        |
| 3.3.5 Ethnolinguistic Identity Theory .....   | 36        |
| 3.3.6 Selbstkategorisierungstheorie .....   | 38        |
| 3.3.7 Sprachideologie .....   | 40        |
| 3.3.8 Zwischenfazit .....   | 42        |
| 3.4 Methoden der (Sprach)-Einstellungsforschung .....   | 43        |
| <b>4 Erkenntnisse aus der Spracheinstellungsforschung .....</b>                                     | <b>45</b> |
| 4.1 Dialekt und Standardsprache .....   | 48        |
| 4.2 Spracheinstellungen und Gender .....  | 50        |
| 4.3 Spracheinstellungen im Migrationskontext .....  | 52        |
| 4.4 Spracheinstellungen in städtischen und ländlichen Regionen .....                                | 52        |
| <b>5 Spracheinstellungen und Sprachvariation im schulischen Kontext .....</b>                       | <b>55</b> |
| 5.1 Die Sprachbarrierediskussion .....  | 55        |
| 5.2 Spracheinstellungen bei Lehrkräften in Österreich .....   | 57        |
| 5.3 Die Rolle der Schule bei der Weitertradierung stereotyper<br>Spracheinstellungen .....          | 59        |
| 5.4 Dialekte im Deutschunterricht: Potenzial oder vergebene Liebesmühe? ..                          | 60        |
| 5.4.1 Dialekt als Barriere für Bildungserfolg? .....  | 60        |

|          |  |            |
|----------|--|------------|
| 5.4.2    | Vor- und Nachteile innerer Mehrsprachigkeit im Kontext der<br>Bildungslandschaft ..... | 62         |
| <b>6</b> | <b>Veränderung von Spracheinstellungen .....</b>                                       | <b>67</b>  |
| 6.1      | Persuasion .....   | 67         |
| 6.2      | Critical-Language-Awareness/Sprachbewusstheit .....                                    | 68         |
| <b>7</b> | <b>Materialien zur Bewusstmachung und Reflexion von Spracheinstellungen .....</b>      | <b>77</b>  |
| 7.1      | Meine, deine, unsere Sprachen .....  | 79         |
| 7.2      | Gemeinsam Dialekt erforschen .....   | 79         |
| 7.3      | Altes Deutsch – neues Deutsch .....  | 80         |
| 7.4      | Das sagt man (nicht)!? .....   | 81         |
| 7.5      | Vorhang auf: Das haben wir gelernt! .....  | 82         |
| <b>8</b> | <b>Erhebungsinstrumente und Auswertungsmethoden .....</b>                              | <b>83</b>  |
| 8.1      | Fragestellungen und Hypothesen .....   | 83         |
| 8.2      | Fragebogendesign .....   | 86         |
| 8.3      | Auswahl der Schulen .....  | 88         |
| 8.4      | Stichprobe .....   | 89         |
| 8.5      | Datenerhebung .....  | 90         |
| 8.6      | Quantitative Auswertungsmethoden .....   | 91         |
| 8.6.1    | Deskriptive Statistik .....  | 92         |
| 8.6.2    | Inferenzstatistik .....  | 98         |
| 8.6.3    | Einzelne Schritte der Auswertung .....   | 106        |
| 8.7      | Ordinal- oder Intervalldaten? Das Skalenproblem .....                                  | 108        |
| 8.8      | Open Science Framework – ein Plädoyer .....  | 110        |
| <b>9</b> | <b>Welche Spracheinstellungen bringen die Proband:innen mit? .....</b>                 | <b>113</b> |
| 9.1      | Dialekt und Standarddeutsch .....  | 115        |
| 9.1.1    | Detaillierte Ergebnisse .....  | 115        |
| 9.1.2    | Generelle Ergebnisse .....   | 118        |
| 9.1.3    | Zwischenfazit Dialekt und Standarddeutsch .....  | 120        |
| 9.2      | Sprecherinnen und Sprecher .....   | 121        |
| 9.2.1    | Detaillierte Ergebnisse .....  | 121        |
| 9.2.2    | Generelle Ergebnisse .....   | 124        |
| 9.3      | Raterinnen und Rater .....   | 127        |
| 9.3.1    | Detaillierte Ergebnisse .....  | 127        |
| 9.3.2    | Zwischenfazit Raterinnen und Rater .....   | 133        |
| 9.4      | Städtische und ländliche Regionen .....  | 134        |
| 9.4.1    | Detaillierte Ergebnisse .....  | 134        |
| 9.4.2    | Generelle Ergebnisse .....   | 140        |
| 9.4.3    | Zwischenfazit städtische und ländliche Regionen .....                                  | 142        |
| 9.5      | Erstsprache .....  | 143        |
| 9.5.1    | Detaillierte Ergebnisse .....  | 143        |

---

|        |   |     |
|--------|---|-----|
| 9.5.2  | Generelle Ergebnisse                                      | 146 |
| 9.5.3  | Zwischenfazit Erstsprache                                 | 147 |
| 9.6    | Bildungshintergrund                                       | 148 |
| 9.6.1  | Detaillierte Ergebnisse                                   | 148 |
| 9.6.2  | Generelle Ergebnisse                                      | 153 |
| 9.6.3  | Zwischenfazit Bildungshintergrund                         | 155 |
| 10     | <b>Diskussion der erhobenen Spracheinstellungen</b>       | 157 |
| 11     | <b>Lassen sich die Spracheinstellungen verändern?</b>     | 161 |
| 11.1   | Allgemeine Prä-Posttest-Unterschiede                      | 161 |
| 11.1.1 | Detaillierte Ergebnisse                                   | 163 |
| 11.1.2 | Generelle Ergebnisse                                      | 177 |
| 11.1.3 | Zusammenfassung der allgemeinen Prä-Posttest-Unterschiede | 188 |
| 11.2   | Dialekt und Standarddeutsch                               | 191 |
| 11.2.1 | Detaillierte Ergebnisse                                   | 191 |
| 11.2.2 | Generelle Ergebnisse                                      | 195 |
| 11.2.3 | Zusammenfassung Dialekt und Standarddeutsch               | 199 |
| 11.3   | Sprecherinnen und Sprecher                                | 201 |
| 11.4   | Raterinnen und Rater                                      | 205 |
| 11.4.1 | Detaillierte Ergebnisse                                   | 205 |
| 11.4.2 | Generelle Ergebnisse                                      | 208 |
| 11.4.3 | Zusammenfassung Sprecherinnen und Sprecher                | 212 |
| 11.5   | Städtische und ländliche Regionen                         | 213 |
| 11.5.1 | Generelle Ergebnisse                                      | 216 |
| 11.5.2 | Zusammenfassung städtische und ländliche Regionen         | 219 |
| 11.6   | Erstsprache   | 220 |
| 11.6.1 | Detaillierte Ergebnisse                                   | 221 |
| 11.6.2 | Generelle Ergebnisse                                      | 227 |
| 11.6.3 | Zusammenfassung Erstsprache                               | 231 |
| 11.7   | Bildungshintergrund                                       | 232 |
| 11.7.1 | Detaillierte Ergebnisse                                   | 232 |
| 11.7.2 | Generelle Ergebnisse                                      | 235 |
| 11.7.3 | Zusammenfassung Bildungshintergrund                       | 236 |
| 12     | <b>Diskussion der Spracheinstellungsveränderung</b>       | 237 |
| 13     | <b>Beantwortung der Hypothesen und Ausblick</b>           | 241 |
| 14     | <b>Schluss</b>  | 251 |
|        | <b>Literaturverzeichnis</b>                               | 253 |



# Zusammenfassung

Im bairischsprachigen Teil Österreichs findet man eine lebendige Dialektlandschaft vor. So geben fast 80 % der Bevölkerung an, mit ihren Arbeitskolleg:innen Dialekt oder Umgangssprache zu sprechen (Ender/Kaiser 2009: 284). Dabei unterliegen die im Alltag verwendeten Sprachvarietäten bestimmten Einstellungsmustern (z. B. Vergeiner et al. 2019): Dialektsprechende werden als gemütlich, aber weniger gebildet, Standardsprechende als kompetenter, aber unfreundlicher eingeschätzt. Auch im schulischen Kontext werden diese Spracheinstellungen nach wie vor transportiert. So geben Lehrkräfte in Österreich zwar an, dass dialektale Sprachverwendung ein wichtiger Faktor für die Identität, jedoch wenig förderlich für den schulischen Erfolg ist (Pribyl-Resch 2021). Dabei hätte es viele Vorteile, Sprachvariation als Chance für den Sprachunterricht zu sehen (Kaiser/Ender 2020). So zeigen Programme, die eine *Critical Language Awareness* im Unterricht implementieren, verbesserte Schulleistungen bei dialekt-sprechenden Schüler:innen (Yiakoumetti 2007; Wheeler/Swords 2004).

Im Zuge des Projekts „Sprachliche Vielfalt verstehen, wertschätzen und ausbauen“ wurden daher erstmalig zu diesem Thema Materialien für den Deutschunterricht entwickelt und in sechs Partnerschulen von den regulären Deutschlehrkräften im Unterricht eingesetzt. Mit einem Prä-Posttest-Design mit Interventions- und Kontrollgruppen (n = 214) wurden die Spracheinstellungen der Schüler:innen mit der Matched-Guise-Methode erhoben und sowohl deskriptiv als auch mit linearen gemischten Modellen ausgewertet.

Dabei zeigten sich die im österreichischen Raum bekannten Spracheinstellungen. So wurden im Allgemeinen standardsprechende Personen kompetenter und dialekt-sprechende sympathischer eingeschätzt. In detaillierter Betrachtung traten deutliche Geschlechterunterschiede ans Licht. So wurden speziell Dialektsprecherinnen signifikant inkompetenter eingeschätzt als Dialektsprecher oder standardsprechende Personen. Die Betrachtung der Stadt-Land-Unterschiede zeigte, dass dialekt-sprechende Personen speziell in den ländlichen Regionen am wenigsten kompetent eingeschätzt werden, zugleich fand man bei den Personen, die Deutsch und eine weitere Erstsprache bzw. eine andere Erstsprache als Deutsch angaben, niedrige Kompetenzeinschätzungen. Auch die Gruppe mit dem höchsten Bildungshintergrund zeigte diese negative Einschätzung. In allen diesen Gruppen zeigten sich teils deutliche Veränderungen durch die Intervention hin zu einer positiveren Kompetenzeinschätzung. Zusätzlich wurden auch moderate Veränderungen auf Sympathieebene gefunden, speziell in der Einschätzung von Dialektsprecherinnen und standardsprechenden Personen.

# Abstract

In the Bavarian-speaking part of Austria, one finds a lively dialect landscape. Thus, almost 80 % of the population state that they speak dialect or colloquial language with their work colleagues (Ender/Kaiser 2009: 284). In this context, the language varieties used in everyday life are subject to certain attitudinal patterns (e. g. Vergeiner et al. 2019): Dialect speakers are seen as cozy but less educated, standard speakers as more competent but less friendly. These language attitudes continue to be conveyed in the school context as well. For example, teachers in Austria state that dialectal language use is an important factor for identity, but not very conducive to academic success (Pribyl-Resch 2021). Yet there would be many advantages to seeing language variation as an opportunity for language teaching (Kaiser/Ender 2020). For example, programs that implement critical language awareness in the classroom show improved school performance among dialect-speaking students (Yiakoumetti 2007; Wheeler/Swords 2004).

Therefore, in the course of the project „Sprachliche Vielfalt verstehen, wertschätzen und ausbauen“ materials for teaching German were developed for the first time on this topic and used in the classroom by regular German teachers in six partner schools. Using a pre-posttest design with intervention and control groups (n = 214), the language attitudes of the students were assessed using the matched-guise method and evaluated both descriptively and using inferential statistical methods, mainly linear mixed models. In part, this revealed language attitudes that are familiar in Austria. In general, standard speakers were rated as more competent and dialect speakers as more likeable. A detailed analysis revealed clear gender differences. In particular, female dialect speakers were rated significantly more incompetent than male dialect speakers or standard speakers of both sexes. An examination of the urban-rural differences showed that in particular in the rural regions dialect speakers were assessed as least competent; at the same time it was found that persons who indicated German and another first language or a first language other than German assessed dialect speakers as having a low level of competence. The group with the highest educational background also showed this negative assessment. All these groups displayed partly significant changes through the intervention towards a more positive competence assessment. In addition, moderate changes were also found at the sympathy level, especially in the assessment of female dialect speakers and male standard speakers.

# 1 Veränderung von Spracheinstellungen im Zuge des Deutschunterrichts – warum die Mühe?

Wenn man in einem Land wie Salzburg, einem der neun Bundesländer der Bundesrepublik Österreich, aufwächst, dann gehört Sprachvariation in der deutschen Sprache, aber auch im mehrsprachigen Kontext, zum Alltag. In ländlichen Gegenden wird zumeist Dialekt oder Umgangssprache gesprochen, im urbaneren Bereich und auch im Schulsetting wird zusätzlich von gewissen Personen die Standardsprache, häufig mit österreichischer Färbung, verwendet (daneben gibt es die Gruppe der deutschen Einwanderer, die zumeist einen bundesdeutschen Standard sprechen). Die Varietäten werden aber nicht als gleichwertig angesehen (vgl. z. B. Vergeiner et al. 2019; Bellamy 2012) und auch wenn Dialekten in Österreich eine wichtige Funktion bei der Bildung regionaler Identität zugesprochen wird (Pribyl-Resch 2021), wird deren Sprecher:innen häufig niedrigere Kompetenz attestiert.

Kinder kommen spätestens in der Sekundarstufe I in Berührung mit diesem Thema, da es im Bildungssystem genug Lehrkräfte gibt, die sich primär als Normvermittler:innen sehen und Normtoleranz dialektaler Varietäten gegenüber als hinderlich empfinden (Davies 2007). Problematisch wird es dann, wenn dialektsprechende Kinder durch Handlungen, wie ständiges Korrigieren der Sprechweise, keine Wertschätzung erfahren. Dadurch können sie eine Abneigung gegen das bestimmte Schulfach, die Standardsprache oder schlimmer, die Schule als Institution selbst entwickeln und sich von (höherer) Bildung abwenden. So etwas sollte in der Schule vermieden werden – jedes Kind, egal welche Herkunft und welchen sprachlichen Hintergrund es besitzt, sollte im Schulsystem die gleichen Chancen bekommen. Eine Möglichkeit aufzuzeigen und zu evaluieren, wie man dies im Zuge des Deutschunterrichts erreichen kann, ist das primäre Ziel dieser Arbeit.

Diese Arbeit gliedert sich in mehrere Teile: In den ersten sechs Kapiteln wird der theoretische Hintergrund der Arbeit besprochen. In Kapitel 2 wird in die regionale und soziale Sprachvariation im Untersuchungsgebiet eingeführt. Abschnitt 2.1 referiert kurz die historische Entwicklung der Sprachlandschaft im Untersuchungsgebiet, in Abschnitt 2.2 wird die historische Entwicklung der Bewertung von Dialekt und Standardsprache aufgezeigt und Abschnitt 2.3 stellt die heutige soziolinguistische Situation sowie den Sprachgebrauch im Untersuchungsgebiet dar.

Kapitel 3 widmet sich Spracheinstellungen aus theoretischer Perspektive. So werden in Abschnitt 3.1 die grundlegenden Eigenschaften von Einstellungen besprochen und in Abschnitt 3.2 die Spezifika von Spracheinstellungen erarbeitet. Der Abschnitt 3.3 widmet sich in mehreren Unterkapiteln unterschiedlichen Theorien, die im Bereich der Spracheinstellungsforschung relevant sind. Es werden die *Social Identity Theory*, die

*Kontextualisierungstheorie*, die *Akkommodationstheorie*, die Theorie zum *Audience Design*, die *Ethnolinguistic Identity Theory*, die *Selbstkategorisierungstheorie* und die Theorie zu *Sprachideologien* besprochen.

Kapitel 4 widmet sich den empirischen Erkenntnissen aus der Spracheinstellungsforschung. Spezieller Fokus wird auf die Forschungsergebnisse zu Varietätenunterschieden (Abschnitt 4.1), zu Genderunterschieden (Abschnitt 4.2), zu unterschiedlichen Herkunftsgruppen und Erstsprachen (Abschnitt 4.3) und Stadt-Land-Unterschieden (Abschnitt 4.4) gelegt, da diese Bereiche in der empirischen Studie wieder aufgegriffen werden.

In Kapitel 5 werden Spracheinstellungen und Sprachvariation im schulischen Kontext besprochen. Großen Einfluss auf den Umgang mit Sprachvariation hatte die Sprachbarrierendebatte, die in Abschnitt 5.1 behandelt wird. Welche Spracheinstellungen Lehrkräfte in Österreich mitbringen, zeigt Abschnitt 5.2. In Abschnitt 5.3 wird die Rolle, die die Schule bei der Weitertradierung stereotyper Spracheinstellungen hat, besprochen, und Abschnitt 5.4 geht der Frage nach, ob die Behandlung von Dialekten im Deutschunterricht Potenzial besitzt.

Der Frage, wie sich (Sprach)Einstellungen verändern lassen, wird in Kapitel 6 nachgegangen. Für die Einstellungsveränderung ist Persuasion ein wichtiger Mechanismus, dessen theoretische Modellierung in Abschnitt 6.1 besprochen wird. Für die Bewusstmachung von Spracheinstellungen mit einhergehender (potenzieller) Einstellungsveränderung sind verschiedene *Critical-Language-Awareness*-Ansätze entwickelt worden, die in Abschnitt 6.2 besprochen werden. Diese Ansätze bildeten auch die Grundlage für die Entwicklung der Unterrichtsmaterialien, die in dieser Studie verwendet wurden und die in Kapitel 7 referiert werden. Die Materialien bestehen aus fünf Modulen, die unterschiedliche Bereiche anvisieren. So wird in Modul 1 die individuelle Bedeutung von Sprachvariation fokussiert (Abschnitt 7.1), Modul 2 (Abschnitt 7.2) widmet sich der Wissensvermittlung über areale Variation und den regionalen Dialekt (speziell die Regelhaftigkeit in dialektaler Lautung und Grammatik). Die Sprachgeschichte und Sprachwandelphänomene werden in Modul 3 (Abschnitt 7.3) durchgenommen, die situative Angemessenheit von Varietätenverwendung steht in Modul 4 (Abschnitt 7.4) im Mittelpunkt, und in Modul 5 (Abschnitt 7.5) werden die Inhalte wiederholt und in Form eines kreativen Abschlussprojekts gefestigt.

Kapitel 8 bereitet auf die empirische Studie vor. Die Fragestellungen und Hypothesen werden in Abschnitt 8.1 formuliert, in Abschnitt 8.2 wird das Fragebogendesign vorgestellt, die Auswahl der Schulen und welche Kriterien diesbezüglich relevant waren, kann in Abschnitt 8.3 nachgelesen werden, die Beschreibung der Stichprobe folgt in Abschnitt 8.4, und wie die Daten erhoben wurden, wird in Abschnitt 8.5 dargelegt. In Abschnitt 8.6 werden in zwei Unterabschnitten die verwendeten quantitativen Darstellungs- und Auswertungsmethoden beschrieben. Ein weiterer Unterabschnitt zeigt exemplarisch auf, wie in der Analyse vorgegangen wurde. Abschnitt 8.7 widmet sich dem statistischen Problem, ob Matched-Guise-Daten als ordinalskalierte oder intervallskalierte Daten kategorisiert werden müssen und welche Konsequenzen die Entscheidung für die eine oder die andere Skala mit sich bringt. Das Kapitel endet mit

Abschnitt 8.8, in dem ein Plädoyer für den transparenten Umgang mit Forschungsdaten formuliert wird.

Kapitel 9 zeigt die Studienergebnisse zu den Spracheinstellungen, die die Schüler:innen mitbringen. Folgende Themenbereiche werden ausführlich analysiert und beschrieben:

- Dialekt und Standarddeutsch (Abschnitt 9.1)
- Sprecherinnen und Sprecher (Abschnitt 9.2)
- Raterinnen und Rater (Abschnitt 9.3)
- Städtische und ländliche Regionen (Abschnitt 9.4)
- Erstsprache (Abschnitt 9.5)
- Bildungshintergrund (Abschnitt 9.6)

Die einzelnen Abschnitte gliedern sich wiederum in drei Unterabschnitte. So werden im ersten Unterabschnitt die detaillierten Ergebnisse beschrieben, im zweiten generalisierte Ergebnisse, indem mehrere Variablenebenen zusammengefasst werden, und in einem dritten findet man die Zusammenfassung der Ergebnisse anhand einer Tabelle. In Kapitel 10 werden die Ergebnisse aus Kapitel 9 diskutiert.

Kapitel 11 handelt von der potenziellen Veränderbarkeit von Spracheinstellungen. Dieses Kapitel ist gleich strukturiert wie Kapitel 9. So finden sich ebenfalls die zuvor beschriebenen inhaltlichen Abschnitte sowie die drei Unterabschnitte, in denen wiederum zuerst die detaillierten, dann die generellen Ergebnisse präsentiert werden, diese zum Abschluss noch zusammengefasst werden. Die Diskussion der Ergebnisse findet in Kapitel 12 statt. Kapitel 13 führt dann die Diskussionen der Kapitel 10 und 12 zusammen, indem die Forschungsfragen erneut aufgeworfen und beantwortet werden, und ordnet die Arbeit in den Forschungskontext ein. Mit einem kurzen Schluss in Kapitel 14 endet die Arbeit.



## 2 Regionale und soziale Sprachvariation im Untersuchungsraum

### 2.1 Regionale Variation

Die Besiedlungs- und damit einhergehend auch die Sprachgeschichte im Land Salzburg lässt sich sehr weit zurückverfolgen. So lassen sich manche der Gewässernamen auf eine präkeltische Sprache zurückführen (Reiffenstein 2020: 2829). Gut belegt ist die keltische Besiedelung, die am Dürrnberg bei Hallein durch den Salzhandel überregionale Bedeutung erlangte und sehr kunstvolle Erzeugnisse hinterließ. Viele Flur- und Gewässernamen zeugen von dieser Besiedlung. Im Zuge der Expansion des Römischen Imperiums wurde dieser Raum fast fünf Jahrhunderte lang ins Römische Reich eingegliedert – auch aus dieser Zeit haben viele Flur- und Gewässernamen überdauert. Nach dessen Zerfall besiedelten im Rahmen der Völkerwanderung slawische Stämme den Süden des Landes, die aber von den etwa zeitgleich kommenden germanischen Stämmen vom Norden her im Laufe der Zeit verdrängt wurden (Reiffenstein 2020: 2893). Die germanische und keltoromanische Bevölkerung schien dann auch über lange Zeit friedlich koexistiert zu haben, worauf Gräberfunde hindeuten (Dopsch 2004). Bis ins 9. Jh. lassen sich in Salzburger Güterverzeichnissen Personen mit romanischen Namen finden (vgl. Reiffenstein 2020: 2897). Danach konsolidierte sich das Stammesherzogtum Bayern, dessen geografische Ausdehnung mit kleinen Abweichungen auch heute noch mit dem bairischen Sprachraum übereinstimmt. In dieser Ausdehnung hielt es aber nicht lange und zerbrach kurz danach schon wieder in kleinere Reiche (vgl. Dopsch 2004). Über Jahrhunderte war dieser Sprachraum von zwei Reichen dominiert, dem Königreich Bayern auf der einen und dem Herzogtum Österreich auf der anderen Seite, das über Jahrhunderte die Kaiserkrone des Heiligen Römischen Reichs Deutscher Nation innehatte und dann nach den Wirren der napoleonischen Kriege selbst zum Kaiserreich wurde. Dazwischen lag das kleine Fürsterzbistum Salzburg, das durch die jahrhundertelange Unabhängigkeit – es war zentraler Bischofssitz des bairischsprachigen Raums – einen eigenen Weg beschritten hatte. Dies schlug sich auch sprachlich nieder (vgl. z. B. Scheutz 2009b).

Nach den napoleonischen Kriegen wurde Salzburg im Zuge des Wiener Kongresses Österreich zugeschlagen – heute ist es eines der neun Bundesländer der Bundesrepublik Österreich. Da sich im Land Salzburg keine autochthonen Siedlungsgebiete von Bewohner:innen von Minderheitensprachen befinden, ist die Amtssprache ausschließlich Deutsch.

Die bairischen Dialekte im Land Salzburg sind in zweierlei Hinsicht von Sprachgrenzen durchzogen. So verläuft auf der West-Ost-Achse die Grenze zwischen mittelbairischen und mittelsüdbairischen Dialekten durch den Flachgau, den nördlichsten Bezirk des Landes, zusätzlich findet man auf der Nord-Süd-Achse mittlerweile eine

bunte Mischung von basisdialektalen historischen westmittelbairischen und in jüngerer Zeit eingewanderten ostmittelbairischen Varianten, die vermehrt in umgangssprachlichen Varietäten in urbaneren Gegenden und bei jüngeren Sprecher:innen auffindbar sind (vgl. Scheutz 2009a; Wiesinger 2014b). Überdies findet man auf der West-Ost-Achse in der Laienwahrnehmung eine Sprachgrenze, die entlang der Nordseite der Kalkalpen verläuft, wo massive Gebirgsstöcke, wie das Tennen- und Hagengebirge, den Tennengau von den südlicheren Gebirgsgauen (Pinzgau, Pongau und Lungau) trennen. Die Bewohner nördlich dieser Trennlinie werden in der Bevölkerung als „Außergebirgler“, die Bewohner südlich davon als „Innergebirgler“ bezeichnet. Diese Zuordnung ist sprachlich durch Variablen wie die im Süden erfolgende und im Norden ausbleibende Affrizierung vom stimmlosen Velarplosiv auch recht treffsicher möglich.

## 2.2 Historische Hintergründe sozialer Sprachvariation

Für die hier angelegte Untersuchung ist jedoch die vertikale Variation, also die Variation zwischen Dialekt und Standardsprache, die bedeutendere. Dafür lohnt es sich, zwei historische Entwicklungen herauszuheben, die hierfür eine besondere Rolle spielten: die Ausformung der gesprochenen Standardsprache als Symbol des Bildungsbürgertums und die Einführung und Durchsetzung der allgemeinen Schulpflicht.

In der zweiten Hälfte des 18. Jh. begann sich eine Form der gesprochenen Standardvarietät herauszubilden, die primär bei einer kleinen Gruppe, dem Bildungsbürgertum, Anwendung fand. „Dieses Bildungsbürgertum stellt im 18. Jh. als verstaatlichte akademische Intelligenz die Funktionselite des spätabolutistischen Staates dar“ (Mattheier 2000: 1962). Von dieser kleinen Gruppe breitete sich die Standardsprache auf andere städtische Schichten wie das Besitzbürgertum, den Adel und dann auch zu einem gewissen Grad auf die Arbeiterschaft aus. Das Besitzbürgertum übernahm früh die Werte des Bildungsbürgertums (Mattheier 1991: 46), der Adel jedoch wehrte sich zu Beginn dagegen. Im ausgehenden 18. Jh. gehörte es in dieser Schicht noch zum Normalfall, entweder im ortsansässigen Dialekt oder in der damals prestigetragenden Sprache Französisch zu kommunizieren. Der Erwerb humanistischer Bildung hatte in dieser Schicht keinen großen Wert. Jedoch gerät „[d]ie Bildungskonzeption des Bürgertums [...] schon in seiner Entstehungsphase in eine Konfrontation zu den überkommenen Modellen der adeligen Bildung. Denn beide gesellschaftlichen Formationen stehen in einem direkten Konkurrenzverhältnis bei der Bewerbung um staatlich-öffentliche Berufslaufbahnen“ (Mattheier 1991: 56). Die Angehörigen des Adels bekamen nun dahingehend ein Problem, dass sie aufgrund der fehlenden Bildung das in Verwaltungs- und Militärlaufbahnen wichtige Standarddeutsche in Schrift und in der gesprochenen Form nicht normgerecht beherrschten. „Der sich modernisierende Staat des 19. Jahrhunderts gerät insofern in ein Dilemma, als einerseits der Adel durch Privilegien und Ämterpatronage Ansprüche auf wichtigste Verwaltungspositionen macht, andererseits bestens ausgebildete Funktionseliten gebraucht werden, die das speziell dafür ausgebil-

dete Bildungsbürgertum anbietet“ (Mattheier 1991: 57). Das führte dazu, dass immer mehr Angehörige des Adels die Ideale des Bildungsbürgertums übernahmen, um bei der Bewerbung um Positionen intellektuell konkurrenzfähig zu bleiben. Als „Vierter Stand“ übernahm auch die Arbeiterschaft die „Standardsprache als Orientierungspunkt für angemessenes Verhalten in der Öffentlichkeit“ (Mattheier 1991: 59), obwohl diese Schicht zumeist nicht die Möglichkeiten hatte, eine den oben Genannten gleichwertige Bildung zu erhalten. Damit war die Grundlage für vertikale Variation im städtischen Umfeld geschaffen.

Die Einführung und Durchsetzung der allgemeinen Schulpflicht ist der zweite Aspekt, der sukzessive die Standardsprache in die Bevölkerung hineintrug. Im 19. Jh. existierten zwei Schultypen, die Volksschulen, die der gesamten Bevölkerung zugänglich waren, und die Gymnasien, die nur von einem kleinen Teil, vornehmlich den Sprösslingen des Bildungsbürgertums (und in weiterer Folge dem Besitzbürgertum und dem Adel), besucht wurden. Beide Schultypen vermittelten Kenntnisse in der deutschen Standardsprache. In den Volksschulen wurden diese nur rudimentär vermittelt, in den Gymnasien legte man viel Wert darauf, die Absolvent:innen in Grammatik, Schrift und Aussprache, die stark an die Literatursprache der deutschen Klassiker angelehnt war, auszubilden (Mattheier 1991: 66). Die Lektüre dieser Klassiker selbst durfte ebenfalls nicht zu kurz kommen und für diejenigen, die es nicht so genau mit der Lektüre nahmen und trotzdem mit Zitaten glänzen wollten, gab es den Büchermann (den es auch heute noch in x-ter Auflage zu kaufen gibt). Dieser spezielle Fokus in den Gymnasien war nicht zuletzt dazu da, nur einer bestimmten sozialen Gruppe dieses sozialindexikalische Symbol – also eine gesprochene, mit Zitaten der klassischen Literatur verzierte Standardsprache – zugänglich zu machen.

Diese klare Schichtentrennung sollte im Bildungswesen noch lange dauern: „Erst durch Entwicklungen im 20. Jahrhundert und teilweise sogar erst durch die Bildungsreformen am Ende der 60er-Jahre wird die enge Verknüpfung zwischen dem Verfügen über die deutsche Standardsprache und der Zugehörigkeit zu bestimmten gesellschaftlichen Gruppen aufgebrochen. Im 19. Jahrhundert liegt darin jedoch das zentrale Faktum, durch das Standardsprache zum Sozialsymbol werden konnte“ (Mattheier 1991: 68).

Auch in den ländlich geprägten Gebieten hatte die Einführung der Schulpflicht Konsequenzen für die sprachliche Realität. Waren zuvor die Basisdialekte die zumeist einzige Kommunikationsform, kam nun das Standarddeutsche hinzu, zumindest in rezeptiver Form durch den Erwerb von Lese- und Schreibfertigkeiten. Ebenso war das Potenzial geschaffen, den Kindern implizit und explizit Einstellungen über die eigene Erstsprache zu vermitteln, die bis dahin in der ländlichen Bevölkerung wohl keine große Rolle gespielt haben dürften. Da sich nach 1850 schon fast 70 Prozent der bürgerlichen Autor:innen von Autobiographien negativ über den Dialekt äußerten (vgl. Mattheier 1991: 53), liegt es nahe, dass diese Einstellungen über die Lehrkräfte in die Dörfer wanderten. Die vertikale Variationssituation und die Einstellungen zu den Varietäten konnten damit im gesamten sozialen Raum voll zum Tragen kommen.

Diese Neuerung im Sprachsystem sorgte aber auch dafür, dass sich das Shifting-Kontinuum, das die bairischen Dialekte auszeichnet (Ammon 2003), über die Zeit immer ausgebauter wurde und ein zunehmender Teil der Bevölkerung nicht nur Standardkompetenzen im schriftsprachlichen, sondern auch im gesprochensprachlichen Kontext erwarb. Diese Form der inneren Mehrsprachigkeit scheint heutzutage eher die Norm als die Ausnahme im Untersuchungsgebiet zu sein (s. Kapitel 2.3).

## 2.3 Sprachgebrauch heute

Wie sieht es aber heutzutage mit dem Sprachgebrauch im Untersuchungsgebiet aus? Prinzipiell existieren zwei Methoden, um Sprachgebrauch zu erheben: die objektive Messung des Sprachgebrauchs, die häufig in Form der Aufnahme von Sprachsituationen und einer anschließenden variablenanalytischen Untersuchung unternommen wird, und die Erhebung der subjektiven Selbsteinschätzung von Sprecher:innen. Im Bereich der objektiven Erfassung des Sprachgebrauchs innerhalb des Dialekt-Standard-Kontinuums in Österreich sind in den letzten Jahrzehnten eine Vielzahl von systemlinguistischen Forschungsarbeiten und -projekten entstanden. Exemplarisch sei der Spezialforschungsbereich „Deutsch in Österreich“ genannt, der in elf Teilprojekten unterschiedliche Ebenen der Sprachlandschaft Österreichs erforscht. Von unterschiedlichen Variationsphänomenen (z. B. Bülow/Wittibschlager/Lenz 2023) bis zu Perzeptionsstudien (z. B. Lenz/Soukup/Koppensteiner 2022) wird unter diesem Dach eine breite Palette an variations- und perzeptionslinguistischer Forschung abgedeckt.

Die Studien der objektiven Messung von Sprachgebrauch besitzen den Vorteil, dass Sprachgebrauch anhand fester linguistischer Kriterien klassifiziert werden kann, allerdings geht dieses Vorgehen mit immensem Ressourcenaufwand einher, wodurch Stichproben häufig klein sind und es schwierig ist, die Sprachverwendung auf eine breitere Bevölkerungsschicht zu inferieren. Dies ist in Form der subjektiven Selbsteinschätzung einfacher.

Frühe Daten zum subjektiven Sprachgebrauch liefert Wiesinger (1992: 293). Er berichtet von einer Umfrage der Jahre 1984/1985 in Österreich, in der sich 78 % der Befragten als Dialektsprechende bezeichneten und 22 % dies verneinten. Für den Alltagsgebrauch gaben 49 % an, Dialekt, ebenso viele Umgangssprache und 2 % Standardsprache zu verwenden. Eine umfangreiche Erhebung des subjektiven Sprachgebrauchs in Österreich liefert Steinegger (1998). So gaben 49,5 % Dialekt, 45,5 % Umgangssprache und 5 % Hochdeutsch auf die Frage nach der bevorzugten Sprechweise an (Steinegger 1998: 90). Im Vergleich der Bundesländer findet er im Land Salzburg niedrigere Werte der dialektalen Sprechweise, allerdings sind die Stichproben nicht homogen genug, um eine klare Einschätzung treffen zu können. Er findet jedoch ein West-Ost-Gefälle mit höherer Dialektverwendung im Westen Österreichs (Steinegger 1998: 201 ff) und höhere Dialektverwendung im ländlichen Bereich (Steinegger 1998: 219 ff).<sup>1</sup>

---

1 S. für eine detailliertere Beschreibung Abschnitt 4.4.

Ender/Kaiser (2009: 283) befragten Personen zur Sprachverwendung im bairischsprachigen und alemannischsprachigen Österreich sowie der Schweiz. In ersterer Gruppe gaben 6 % fast ausschließliche, 11 % häufige, 30 % seltene und 52 % so gut wie keine Hochdeutschverwendung im Alltag an.

Eine aktuelle Studie zur subjektiven Einschätzung der Sprachverwendung (n = 350) in Salzburg und dem angrenzenden Bayern in der Primarstufe und der Sekundarstufe I liefern Unterberger/Pribyl-Resch (im Druck). Wenn man die Personen abzieht, die als Familiensprache eine andere Sprache als Deutsch verwenden, geben in Salzburg mehr als 95 % der Primarschüler:innen an, als Familiensprache entweder *Dialekt* oder *eine Form zwischen Dialekt und Hochdeutsch* zu verwenden. In der Sekundarstufe I sinkt dieser Wert auf etwas über 90 %.

Auch wenn man über die Jahre einen Zuwachs der standardsprachlichen Sprachverwendung findet, sind dialektale und umgangssprachliche Sprachverwendung in Österreich allgegenwärtig und stabil, was im Licht der historisch erfolgten Stigmatisierung dieser Varietäten, der Sprachbarrierendebatte in der 2. Hälfte des 20. Jh. (s. Abschnitt 5.1) und der latent vorhandenen Spracheinstellungen (s. Abschnitt 4.1) durchaus erstaunlich ist. Dabei ist davon auszugehen, dass der Großteil der Österreicher:innen mit deutscher Erstsprache nicht nur eine Varietät beherrscht, sondern sich zu einem gewissen Grad im Shifting-Kontinuum, das die bairischen Dialekte auszeichnet, bewegen kann (vgl. Ammon 2003). So zeigen Ender/Kaiser (2014) und Kaiser/Ender (2015) mitunter ausgeprägte situative, areale sowie individuelle Unterschiede im Shifting- bzw. Akkommodationsverhalten österreichischer Sprecher:innen. Daneben existiert zusätzlich die in der ersten Welle der soziolinguistischen Forschung gut erforschte soziale Variation (s. Kapitel 4).

Sprachvariation findet man überdies auf Ebene der Standardsprache. Im Forschungsdiskurs wird schon länger von einer plurizentrischen bzw. pluriarealen Ausrichtung der deutschen Standardsprache ausgegangen (vgl. z. B. Wiesinger 2010: 361). Dies bedeutet aber nicht, dass die unterschiedlichen Ausformungen der Standardsprachen in den deutschsprachigen Ländern über das gleiche Prestige verfügen. In Österreich ist zu beobachten, dass der österreichische Standard von der Bevölkerung als weniger prestigeträchtig im Vergleich zum bundesdeutschen Standard gesehen wird (vgl. Kaiser 2006; Pfrehm 2007; Peter 2015; Fink/Ransmayr/de Cillia 2017; s. auch Abschnitt 4.1). Somit besitzen die Österreicher:innen nicht nur negative Spracheinstellungen zu ihren Dialekten, sondern auch zu ihrer Form der Standardsprache.

Zur innerdeutschen Sprachvariation kommen noch viele andere Sprachen hinzu, die im Untersuchungsgebiet im Alltag gesprochen werden. Dies betrifft vor allem die Sprachen der Migrant:innen, die sich in den letzten Jahrzehnten hier ansiedelten. Interessanterweise sind gewisse Ähnlichkeiten zwischen der Einschätzung von lernervarietären Akzenten bzw. jugendsprachlichen Varietäten aus dem Migrantenumfeld und Dialekten zu beobachten (Wiese 2015). Zum Teil sind diese beiden Bevölkerungsgruppen also ähnlichen Stigmatisierungsmustern unterworfen – dazu aber mehr in Abschnitt 5.1. Der nächste Abschnitt widmet sich zuvor den theoretischen Hintergründen von Spracheinstellungen.



## 3 Dimensionen von Spracheinstellungen

### 3.1 Grundlegende Eigenschaften von Einstellungen

Bei Spracheinstellungen handelt es sich um eine spezielle Form von Einstellungen, die sich auf Sprachen und ihre unterschiedlichen Varietäten beziehen. Sie teilen grundlegende Eigenschaften mit Einstellungen aus anderen Inhaltsbereichen, daher wird zur Einführung in dieses Thema der breitere Blick zur Sozialpsychologie gewagt und begonnen, allgemeine Eigenschaften von Einstellungen zu beschreiben. Bei der Beschreibung der Inhalte von Einstellungen hat sich in der Forschung das Multikomponentenmodell (Eagly/Chaiken 1993) durchgesetzt: Dieses unterteilt Einstellungen in drei Komponenten: eine affektive, eine kognitive und eine Verhaltenskomponente. Die affektive Komponente umfasst dabei die Gefühls- bzw. emotionale Ebene. Auf Spracheinstellungen umgelegt wäre dies z. B. die Zuneigung oder Abneigung zu einer bestimmten Sprache oder Varietät. Die kognitive Komponente umfasst „Gedanken, Überzeugungen und Eigenschaften, die wir mit einem bestimmten Gegenstand assoziieren“ (Stroebe/Stroebe/Hewstone 2007: 192). Einen zentralen Aspekt der kognitiven Einstellungskomponente stellen dabei Stereotype dar. Im Bereich der Spracheinstellungen wären dies z. B. die bekannten Vorurteile, dass Dialektsprechende weniger gebildet, derber, aber auch sympathisch, Standardsprechende höher gebildet, aber auch unnahbar seien. Die dritte Komponente von Einstellungen bildet die Verhaltenskomponente. So beeinflussen affektive und kognitive Komponenten das Verhalten, der Einfluss des Verhaltens auf die beiden anderen Komponenten, etwa im Sinne einer Verhaltensänderung, ist jedoch ebenfalls gegeben, was speziell in Hinsicht auf Einstellungsveränderungen einen möglichen Ansatzpunkt darstellt (vgl. Stroebe/Stroebe/Hewstone 2007: 193 f).

Was die Ausprägung von Einstellungen betrifft, gibt es in der Forschung zwei Sichtweisen. In der ersten geht man davon aus, dass negative und positive Einstellungen auf einer einzelnen Dimension angesiedelt sind. Dies würde dazu führen, dass positive Einstellungen negative hemmen würden und umgekehrt. In der zweiten wird davon ausgegangen, dass negative und positive Einstellungen separat voneinander gespeichert sind, was graduelle Abstufungen und koexistierende positive und negative Einstellungen auf beiden Einstellungsdimensionen möglich machen würde. Da für beide Sichtweisen Belege existieren, gibt es hierzu keine abschließende wissenschaftliche Meinung (vgl. Stroebe/Stroebe/Hewstone 2007: 155 f).

Eine weitere Frage, der man in Bezug zu Einstellungen nachgehen kann, ist die nach deren Funktionen. Der zentrale Grund liegt in der Eigenschaft begründet, dass die Belegung einer Entität mit positiven und/oder negativen Urteilen eine einfach durchzuführende Klassifizierungsstrategie darstellt und dadurch als Energiesparmechanismus fungiert (vgl. Stroebe/Stroebe/Hewstone 2007: 199 ff). Somit wäre eine negative Spracheinstellung Dialekt gegenüber beispielsweise praktisch, wenn man eine Karriere in

einem Bereich anstrebt, in dem dialektaler Sprachgebrauch stigmatisiert ist. Einer Person unvoreingenommen zu begegnen, diese rein nach interaktionalen Verhaltensinhalten einzuschätzen und eben nicht nach vorgefertigten (stereotypen) Kategorien in Bezug auf Aussehen und Sprachwahl, ist dagegen mit höherer kognitiver Anstrengung verbunden.

Grundlegend kann zwischen zwei funktionalen Typen von Einstellungen unterschieden werden: instrumentellen und Wertausdruckseinstellungen: „Bei instrumentellen Einstellungen werden Einstellungsgegenstände danach klassifiziert, ob sie dem eigenen Interesse förderlich sind, während in Wertausdruckseinstellungen Fragen zum Selbstbild und zu persönlichen Werten zum Vorschein kommen“ (Stroebe/Stroebe/Hewstone 2007: 202). Obwohl Spracheinstellungen prinzipiell beiden Typen zugeordnet werden können (als soziolinguistischer Forscher ist es z. B. durchaus möglich, instrumentelle Einstellungen zum Gegenstand „Spracheinstellungen“ zu besitzen), sind sie doch zumeist als Wertausdruckseinstellungen einzuschätzen, wie etwa der enge Zusammenhang von Sprache und Identität zeigt (s. für einen umfassenden Überblick über dieses Thema Edwards 2012). Zusätzlich sind Einstellungen kontextsensitiv: „Menschen, die sehr ambivalent gegenüber einem Gegenstand sind, lassen sich stärker von Merkmalen ihrer Umwelt beeinflussen, die ihre positiven und negativen Eigenschaften salient machen“ (Stroebe/Stroebe/Hewstone 2007: 196). Dabei muss diese Kontextsensitivität nicht nur vom Einstellungsgegenstand ausgehen. Sie kann auch von der Stärke der Selbstüberwachung (self monitoring) der Person herrühren. Personen mit starker Selbstüberwachung stimmen ihr Verhalten – und auch ihre Einstellungen – auf die Situation ab, in der sie sich gerade befinden. Auf Personen, bei denen die Selbstüberwachung schwächer ausgeprägt ist, zeigt der Kontext weniger Einfluss (vgl. Stroebe/Stroebe/Hewstone 2007: 196).

Ein weiterer Punkt betrifft die Stabilität von Einstellungen. Dabei gilt als zentrale Eigenschaft, dass die Stärke einer Einstellung mit ihrer Dauerhaftigkeit und Resistenz gegen Veränderung zu korrelieren scheint. Gleichzeitig beeinflusst eine starke Einstellung das Verhalten von Personen eher (vgl. Stroebe/Stroebe/Hewstone 2007: 204). Grundsätzlich ist also ein Zusammenhang zwischen dem Verhalten von Personen und ihren Einstellungen beobachtbar. Allerdings ist die Möglichkeit, Verhalten durch die Kenntnis von Einstellungen vorherzusagen, stark durch weitere Faktoren wie Einstellungsstärke, Persönlichkeitsmerkmale oder der Art des Verhaltens beeinflusst. In der Sozialpsychologie wurden daher verschiedene Modelle entwickelt, die den Zusammenhang zwischen Verhalten und Einstellungen erklären sollten (s. für einen Überblick Kapitel 6.8.2 in Stroebe/Stroebe/Hewstone 2007). Da in der Spracheinstellungsforschung eigenständige und zum Teil recht komplexe Modelle in diesem Zusammenhang entwickelt wurden (s. dafür Abschnitt 3.3), ist eine detaillierte Behandlung der Modelle aus der Sozialpsychologie nicht erforderlich.

## 3.2 Spezifika von Spracheinstellungen

Spracheinstellungen teilen sich grundlegende Eigenschaften mit Einstellungen aus anderen Inhaltsbereichen. So wird weitgehend in der Literatur die Ansicht vertreten, dass sie ebenfalls affektive, kognitive und Verhaltenskomponenten beinhalten (s. z. B. Soukup 2019: 90). Aber schon wenn es um die Frage nach der Menge der Dimensionen geht, wird in der soziolinguistischen Forschung davon ausgegangen, dass es sich bei Spracheinstellungen zu einer bestimmten Varietät oder Sprache um Einstellungen mit mehreren Dimensionen handeln muss. Speziell im englischsprachigen Raum, in dem die Einstellungsforschung lange Tradition besitzt und viele Ergebnisse vorgebracht hat, geht man von mehrdimensionalen Gebilden aus. Eine vielzitierte Einteilung liefert Preston (2006: 41). Er geht von den drei Dimensionen „competence“, „personal integrity“ und „social attractiveness“ aus, die er zu einem späteren Zeitpunkt in zwei Dimensionen umgruppiert: „The concepts of ‚correctness‘ (i. e. prestige, standardness, competence) and ‚pleasantness‘ (i. e. solidarity, social attractiveness) are the two most frequent characteristics of speech that have been noticed in a great deal of social psychological research on attitudes to language“ (Preston 2017: 26 f). Dragojevic (2017: 1) benennt die Dimensionen etwas anders: „Standard and nonstandard varieties elicit different evaluative reactions along the status and solidarity dimensions“. Solche oder zumindest ähnlich lautende Unterteilungen der Dimensionen werden auch von den Forscher:innen im deutschsprachigen Sprachraum übernommen (s. für eine theoretische Auseinandersetzung mit dem Thema Soukup 2019: 93).

Diese mehrdimensionale Sichtweise von Spracheinstellungen hängt stark mit den Konzepten des „overt and covert prestige“ zusammen, die aus Erkenntnissen der frühen Spracheinstellungsforschung stammen. Diese Formen von Prestige wurden schon von Labov in einer seiner frühen Studien „The Social Stratification of English in New York City“ (eine Zusammenfassung der originalen Studie findet man in Labov 2006) erkannt, da die Proband:innen die Standardvarietät zwar als prestigereich einschätzten (overt prestige), es aber auch verborgene Gründe dafür geben musste, warum sich die Varietäten mit niedrigem Prestige hartnäckiger hielten, als durch die Aussagen angenommen werden konnte. Es musste also (zumindest) eine weitere Dimension von Spracheinstellungen existieren, die schwieriger zu elizitieren ist, für Benutzer:innen einer Varietät mit geringem Prestige jedoch in ihrem Sprachumfeld ebenfalls Relevanz besitzt. In Verschränkung mit Prestons Dimensionen kann die soziale Attraktivitätsdimension mit „covert prestige“ in Verbindung gebracht werden. So werden z. B. Afroamerikaner, die ihre Varietät zugunsten des als „weiß“ empfundenen standardamerikanischen Englisch aufgeben, von ihrer Herkunftsgruppe als „cultural traitors“ bezeichnet (Siegel 2006: 161). Ähnliches findet man auch über lateinamerikanische Einwanderer in den USA und französischsprachige Kanadier:innen in Kanada, die sich ihren romanischen Erstsprachen ab- und dem Amerikanischen Englisch zuwenden (s. Edwards 2012).

Spracheinstellungen scheinen also zumindest zwei Dimensionen aufzuweisen: eine Kompetenz-, Korrektheits- bzw. Statusdimension, die über „overt prestige“ und eine Solidaritäts- bzw. soziale Attraktivitätsdimension, die über „covert prestige“ verfügt. Diese beiden Dimensionen stehen ebenfalls im Zentrum der Spracheinstellungsforschung im österreichischen Kontext (s. Abschnitt 4.1).

### 3.3 Spracheinstellungen, soziale Gruppen und Identität

Die theoretische Auseinandersetzung mit Spracheinstellungen kann ebenfalls auf eine reiche Tradition zurückblicken. Speziell in der zweiten Hälfte des 19. Jh. kamen in der Sozialpsychologie und der soziolinguistischen Forschung eine Reihe von Theorien auf, die bestimmte Verhaltensmuster im sprachlichen Verhalten mit Spracheinstellungen und dem sozialen Gefüge in Einklang bringen wollten. Dieser ganze Bereich ist jedoch etwas unübersichtlich. Manche der Theorien bauen aufeinander auf, andere nicht; manche beziehen sich konkret auf sprachliches Handeln und Spracheinstellungen, andere streifen diese Themen nur implizit. Auch in Überblickswerken zur Soziolinguistik scheint es hier keine klare Linie zu geben, wie man mit den existierenden Modellen umgehen soll. So sucht man im Kapitel von Llamas/Mullany/Stockwell (2006) zu Spracheinstellungen vergeblich nach theoretischen Erklärungsversuchen, Meyerhoff (2006) beruft sich neben der Thematisierung des linguistischen Relativismus sowohl auf die Social Identity Theory als auch die Akkommodationstheorie. Mesthrie et al. (2009) nehmen in ihrer Einführung in die Soziolinguistik Spracheinstellungen als Thema gleich gar nicht auf, Holmes (2013) nimmt als theoretische Fundierung die Konzepte des overt und covert prestige und in der Einführung von Wardhaugh/Fuller (2015) sucht man wiederum vergeblich nach einem eigenen Kapitel für Spracheinstellungen. Auch in deutschsprachigen Einführungen in die Soziolinguistik ist das Thema rar gesät. So werden Spracheinstellungen bei Dittmar (1997) nur kurz im Kontext soziologischer Aspekte, die in der Feldarbeit eine Rolle spielen können, erwähnt. Ähnlich sieht dies bei Spitzmüller (2022) aus, der Spracheinstellung ebenfalls nur kurz im historischen Kapitel zur Entwicklung der Soziolinguistik anführt. Daher werde ich in den folgenden Unterabschnitten – ohne Anspruch auf Vollständigkeit – Theorien vorstellen, die häufig als theoretische Fundierung von Spracheinstellungen herangezogen werden.

#### 3.3.1 Social Identity Theory

Spracheinstellungen stehen in engem Kontakt zu sozialen Gruppen, Gruppenbildung und gruppenspezifischem evaluativem Verhalten. Um die Zusammenhänge zwischen diesen Bereichen zu beleuchten, lohnt es sich, die Social Identity Theory (SIT) näher zu betrachten. Auch wenn Spracheinstellungen nicht explizit zum Thema in dieser Theorie gemacht werden, lassen sie sich nahtlos einfügen. Im Folgenden beziehe ich mich auf die Arbeit von Tajfel/Turner (1979), die als grundlegend für die Theorie gilt.

Die SIT geht von einem gut belegbaren Paradigma der Sozialpsychologie aus: dem „Ingroup Bias“, also der Differenzminimierung innerhalb der Ingroup und der Differenzmaximierung zwischen *Ingroup* und *Outgroup*.<sup>2</sup> Diese gruppenspezifischen Evaluationsmechanismen können sehr einfach und zuverlässig eliziert werden, sogar, wenn die Gruppenzuteilung völlig zufällig erfolgt. „The basic and highly reliable finding is that the trivial ad hoc intergroup categorization leads to ingroup favoritism and discrimination against the outgroup“ (Tajfel/Turner 1979: 39). Gleichzeitig ist – für die Prognose der Handlung einzelner Gruppenmitglieder – das Interpersonelle-Intergruppen-Handlungskontinuum bedeutend. Ist die Handlungssituation an der interpersonellen Seite des Spektrums ausgerichtet, handelt man aufgrund der persönlichen Beziehung, die man mit dem Interaktanten pflegt – die Interaktion zwischen Ehepartnern wäre hier ein Beispiel. Bei Intergruppen-Verhalten handelt man beeinflusst durch gruppenspezifische Einstellungs- und Handlungscharakteristika, wie dies z. B. bei einer Polizeikontrolle vorkommt (vgl. Tajfel/Turner 1979: 36).

In Zusammenhang mit dem Interpersonellen-Intergruppen-Kontinuum steht auch ein weiteres Kontinuum im Sozialverhalten: dasjenige zwischen sozialer Mobilität (social mobility) und sozialer Veränderung (social change). Soziale Mobilität stellt dabei die Möglichkeit dar, bei Unzufriedenheit mit der eigenen Ingroup diese zu verlassen. Dies ist jedoch nicht in jeder kulturspezifischen Ausgangssituation möglich, man denke z. B. an die Kasten Indiens, in denen sozialer Aufstieg bzw. ein Übertritt von einer zu einer anderen Kaste nicht vorgesehen ist. Befindet man sich in einer solchen Situation (in der keine Möglichkeit des Ausstiegs besteht), bleibt die Möglichkeit, soziale Veränderung anzustoßen. Veränderungsbemühungen sind aber konfliktrichtig und Personen setzen sich potenziell dem Vorwurf des kulturellen Verrates aus (vgl. Tajfel/Turner 1979: 35 f). Tajfel/Turner fassen zusammen:

[T]he nearer are members of a group to the ‚social change‘ extreme of the belief-systems continuum and the intergroup extreme of the behavioral continuum, the more uniformity they will show in their behavior toward members of the relevant out-group [...] [and] the more they will tend to treat members of the out-group as undifferentiated items in a unified social category, rather than in terms of their individual characteristics (Tajfel/Turner 1979: 36).

---

2 Die Termini Ingroup, Outgroup und Ingroup Bias stammen aus den Forschungen von Tajfel und Turner zum Minimal Group Paradigm, das zentral für die Entwicklung der Social Identity Theory war. In diesem Paradigma werden Personen anhand arbiträrer Kriterien (z. B. Münzwurf) Gruppen zugewiesen, deren Gruppenmitglieder sie jedoch nicht kennen. Im Anschluss werden die Personen aufgefordert, Punkte an Mitglieder ihrer eigenen Gruppe (Ingroup) oder anderen Gruppen (Outgroups) zu verteilen. Das erstaunliche Ergebnis ist, dass die Personen ihre Gruppenmitglieder bevorzugten und andere Gruppen benachteiligten (ingroup bias) (Hornsey 2008: 205 ff).

Warum kommt es aber zu potenzieller Unzufriedenheit der eigenen Ingroup gegenüber? Auch hierauf geben Tajfel/Turner (1979: 40) Antworten, die sie auf folgende Annahmen stützen:

1. Individuen bemühen sich um ein positives Selbstbild.
2. Soziale Gruppen/Kategorien und die Teilnahme in diesen sind verbunden mit positiven und negativen Einstellungen.
3. Die Einstellungen der eigenen Gruppe gegenüber werden im Vergleich mit anderen Gruppen festgelegt.

Diese Annahmen nehmen sie zur Grundlage, um folgende Prinzipien abzuleiten:

1. Individuen bemühen sich um positive soziale Identität.
2. Positive soziale Identität stützt sich hauptsächlich auf den Vergleich mit untergeordneten sozialen Gruppen.
3. Wenn die soziale Identität nicht befriedigend ist, versuchen Individuen, diese Gruppe zu verlassen oder ihr positivere Werte zuzuschreiben.

Daraufhin formulieren sie ihre zentrale Hypothese der Theorie: „[P]ressures to evaluate one's own group positively through in-group/out-group comparisons lead social groups to attempt to differentiate themselves from each other“ (Tajfel/Turner 1979: 40 f). Die Mechanismen, wie diese Gruppenvergleiche in konkreten Situationen angestellt werden, bleiben Tajfel/Turner ebenfalls nicht schuldig:

1. Gruppenmitglieder müssen die Mitgliedschaft als Teil des Selbstkonzeptes verinnerlicht haben.
2. Die soziale Situation muss Gruppenvergleiche zulassen können.
3. Ingroups vergleichen sich nur mit Outgroups, die Relevanz, ein bestimmtes Näheverhältnis und gewisse Salienz besitzen.

„The aim of differentiation is to maintain or achieve superiority over an out-group on some dimensions“ (Tajfel/Turner 1979: 41). Dieser Punkt ist für Spracheinstellungen überaus bedeutend, könnte er doch eine Erklärung der Kompensationshypothese (Yzerbyt/Provost/Corneille 2005) darstellen. Die Kompensationshypothese geht davon aus, dass die Besserbewertung von Dialektsprecher:innen auf Sympathieebene daher stammt, dass sie auf Kompetenzebene schlechter abschneiden. Schoel/Stahlberg (2012: 216) zeigen in einem experimentellen Setting, dass die „Wärmedimension“, wie sie die Sympathieebene nennen, ähnliche Ergebnisse für Dialekt- und Standardsprechende erbringt, wenn sie elitisiert wird, ohne die Kompetenzebene zuvor kognitiv zu primen. Tut man dies, schneiden Dialektsprechende besser ab.

Die SIT geht jedoch über rein mechanistische Erklärungen hinaus und zeigt auch mögliche Formen auf, die soziale Veränderung hervorrufen können:

1. Individuelle Mobilität: Die einfachste Form der Veränderung für das Individuum besteht darin, die soziale Gruppe zu verlassen und sich in eine einzufügen, die zufriedenstellender für das Selbstkonzept ist.
2. Soziale Kreativität: Die soziale Gruppe kann von Mitgliedern auf drei verschiedene Arten gestärkt werden:
  - a) Ein positiver Gruppenvergleich wird auf einer neuen Dimension angestrebt.
  - b) Die Änderung von negativen Einstellungen, sodass der neuerliche Vergleich eine positivere Wahrnehmung der Ingroup herbeiführt.
  - c) Die Änderung der Vergleichsgruppe hin zu einer weniger prestigeträchtigen Gruppe.
3. Sozialer Wettbewerb: Die konfliktträchtigste Möglichkeit der Veränderung besteht darin, in einen offenen Wettbewerb um bestimmte Ressourcen mit der Vergleichsgruppe einzutreten.

Die SIT stellt somit einen Erklärungsansatz für die Entstehung und die Veränderung von Spracheinstellungen dar. Positiv wahrgenommene soziale Identität basiert auf dem Vergleich von Merkmalen der eigenen Gruppe mit einer anderen, mit dem Ziel der überlegenen Wahrnehmung der Ingroup. Gruppen konstituieren sich gleichzeitig aus solchen salienten und relevanten Merkmalen ihrer Mitglieder. Diese Merkmale können aus unterschiedlichen Dimensionen stammen. Sprach- und Varietätenwahl gehört dabei zu einem wichtigen Merkmal in der sozialen Indexikaliserbarkeit von Personen. So nimmt man in der evolutionspsychologischen Forschung an, dass Sprache in der fernen Vergangenheit deswegen ein wichtiges indexikalisches Zeichen war, da sich benachbarte menschliche Gruppen eher durch ihre Sprache als durch ihr Aussehen unterschieden (Dragojevic 2017: 5). Gleichzeitig ist es sehr schwer, eine Fremdsprache oder eine fremde Varietät akzentfrei zu erwerben. Dadurch unterscheidet sich Sprache als Index von solchen wie der Kleidungswahl, in denen ohne größeren Aufwand „geschummelt“ werden kann oder auch von physiologisch salienten Zeichen wie der Hautfarbe in Mitteleuropa, die zwar nicht verborgen werden kann, aber auch nicht unbedingt den Sozialstatus indexikalisiert. So zeigen etwa fünfjährige Kinder eine Präferenz für Personen, die die eigene Varietät sprechen – Merkmale wie Hautfarbe (race) spielen dabei eine untergeordnete Rolle (Dragojevic 2017: 5).

Die Sprache, Akzente bzw. Varietäten verraten als stereotypes kognitives Konzept vermeintlich vieles über die soziale Zugehörigkeit und damit, ob es sich bei der Person um eine Ingroup- oder Outgroup-Angehörige handelt.

### 3.3.2 Kontextualisierungstheorie

Handlungskontexte und deren Einfluss auf die Semantik sind sogleich das Stichwort und zentrale Thema für die Interaktionstheorie von Gumperz (1982). Dem Kontext wird in semantisch-pragmatischen Theorien unterschiedlich starker Einfluss auf die semantische Interpretierbarkeit von sprachlichen Zeichen zugesprochen. Mit den Deiktika existiert sogar eine Gruppe von Ausdrücken, die nur im Kontext interpretierbar sind. Dazu zählen Demonstrativa, Lokal- und Temporaladverbien wie da, gestern, heute, hier

oder dort und, häufig als Personaldeixis bezeichnet, Personenbezeichnungen wie Personalpronomen oder Personennamen. „In vielen traditionellen und modernen Deixistheorien wird die Beziehung zwischen dem Kontext und dem deiktischen Ausdruck bzw. seiner Bedeutung (Interpretation) unidirektional und deterministisch gesehen: der Kontext bestimmt die Interpretation des sprachlichen Ausdrucks“ (Auer 2013: 172). So definiert z. B. Löbner (2015: 32 f) Sozialdeixis folgendermaßen: „Der Begriff soziale Bedeutung bezieht sich nicht auf diesen allgemeinen Aspekt sprachlicher Interaktion [er referiert hier auf allgemeine Kommunikationsakte] und ist daher nicht mit der Ebene des kommunikativen Sinns sprachlicher Handlungen zu verwechseln. Vielmehr liegt die soziale Bedeutung auf derselben Ebene wie deskriptive Bedeutung: Sie ist eine Dimension der Ausdrucksbedeutung“. Für Gumperz wäre diese Aussage so nicht haltbar. Er gesteht dem Kontext größeres Gewicht für das Zustandekommen von Semantik zu. Gelingende Kommunikation hängt seiner Meinung nach stark davon ab, dass man innerhalb kürzester Zeit Hinweise sammelt, die eine Einschätzung der Situation möglich machen. Gumperz nennt diese Hinweise contextualisation cues:

[C]onstellations of surface features of message form are the means by which speakers signal and listeners interpret what the activity is, how semantic content is to be understood and how each sentence relates to what precedes or follows (Gumperz 1982: 131).

Damit liefert er eine ausgesprochen linguistische Vorstellung dieser, wie Auer sie nennt, Kontextualisierungshinweise. Auer erweitert diese und damit auch die Reichweite der gumperzischen Theorie:

Bei den Kontextualisierungshinweisen sind zu unterscheiden: Kinetik und Proxemik, Prosodie (Tonhöhenverlauf, Lautstärke, Geschwindigkeit, Rhythmus und Gliederung in Tongruppen, Akzent), Blickverhalten, zeitliche Platzierung (Pausen, Simultansprechen), Varietäten-/Sprachwahl, lexikalische Variation sowie sprachliche Formulierungen (Auer 1986: 26).

Gumperz entwickelt seine Theorie speziell aus Untersuchungen von Fällen, in denen die Kommunikation aufgrund von fehlender oder falscher Deutung der Kontextualisierungshinweise fehlschlägt. Häufig beschäftigt er sich dabei mit der Prosodie. So berichtet er von Beobachtungen in Schulklassen, in denen ein afroamerikanisches Kind bei Aufforderung etwas zu lesen mit „I don’t wanna read“ mit steigender Intonation am Ende antwortet. Dies wurde von der Lehrkraft als unkooperatives Verhalten interpretiert. Diese prosodische Form bei derartigen Verneinungen scheint in afroamerikanischen Familien jedoch anders als unter Weißen indexikalisiert zu sein. Kinder zeigen wohl dadurch, dass sie sehr wohl vorlesen möchten, jedoch noch etwas Ermutigung benötigen (vgl. Gumperz 1982: 147). Ein ähnliches Ergebnis kultureller Unterschiede in der Verwendung prosodischer Muster liefert er in der Diskursanalyse amerikanischer und indischer Studierender beim Gebrauch ihrer jeweiligen englischen Varietäten. Indische Studierende wurden regelmäßig von amerikanischen unterbrochen, da Letz-

tere aufgrund der Intonation davon ausgingen, dass die Aussage beendet war, was von den indischen Studierenden als sehr unhöflich eingeschätzt wurde (vgl. Gumperz 1982: 148 f).

Aber nicht nur die Prosodie, auch andere „Verfehlungen“ wie die Varietätenwahl können negative Konsequenzen für die Kommunikation mit sich bringen. Gumperz berichtet von einem schwarzen Studenten, der zum Interview bei einer ebenfalls schwarzen Informantin erscheint und seinem Gegenüber nicht mit der ihm entgegengebrachten Varietät der Afroamerikaner antwortet, sondern in der Standardsprache, was die Interviewsituation steif und unbefriedigend macht (vgl. Gumperz 1982: 133).

Die angemessene Interpretation von und Reaktion auf Kontextualisierungshinweise scheint stark am Erfolg kommunikativen Handelns mitbeteiligt zu sein, aber nicht nur das. Durch Kleidungswahl, Auftreten, Varietäten- bzw. Sprachwahl, dem Griff zur Höflichkeitsform und Wahl der stilistischen Sprechlage sind Personen aktiv am Aufbau des Kontexts beteiligt. Auer (2013: 179) fasst dies folgendermaßen zusammen: „‚Kontextualisierung‘ [kann man] also als eine dreistellige Relation definieren zwischen Ausdrucksmitteln (Idiomatik, Gestik, Prosodie etc.), der Bedeutung (Interpretation) bestimmter Handlungen und Wissensbeständen (frames), die diese Interpretation ermöglichen, indem sie als ihr Kontext relevant gemacht werden“.

### 3.3.3 Akkommodationstheorie

Die Akkommodationstheorie, die in Kapitel 9 der Monografie von Giles/Powesland (1975) detailliert ausgeführt wird, beschreibt den Prozess, der eintritt, wenn sich in einer kommunikativen Situation zwischen den Interaktionspartnern die Sprach- bzw. Varietätenwahl aneinander anpasst. Grundsätzlich ist sprachliche Reaktion in zwei Richtungen möglich: „Convergence“, also das Angleichen der Varietäten/Sprachen der jeweiligen Interaktanten und damit die eigentliche Akkommodation, und „divergence“, ein Auseinanderdriften der Varietäten/Sprachen. Im Original klingt das folgendermaßen: „Speech convergence is a strategy of identification with the speech patterns of an individual internal to the social interaction, whereas speech divergence may be regarded as a strategy of identification with regard to the linguistic norms of some reference group external to the immediate situation“ (Giles/Powesland 1975: 156). Dem Spannungsfeld zwischen Individuum und sozialer Gruppe kommt damit in der Akkommodationstheorie eine wesentliche Rolle zu, die auf Differenzminimierung innerhalb von Ingroup-Mitgliedern und der Differenzmaximierung zwischen Ingroup- und Outgroup-Mitgliedern basiert und damit von den gleichen Prämissen wie die SIT ausgeht. Giles/Powesland (1975) unterscheiden zwischen Akkommodation im mehrsprachigen Kontext und Akkommodation innerhalb einer Einzelsprache, also zwischen Varietäten. Die beiden Modelle werden aus einer Perzeptionsperspektive gedacht; Ersteres basierend auf empirischen Ergebnissen, die in Experimentalstudien generiert wurden. Im mehrsprachigen Modell findet man eine primäre Unterteilung in die Möglichkeiten der Akkommodation und der Nichtakkommodation im Gesprächskon-

text. Wird von Seite des Kommunikators akkommodiert, gibt es von Rezipient:innen-seite her zwei Möglichkeiten, dies zu interpretieren (vgl. Giles/Powesland 1975: 164):

1. Wird es als ungezwungene Bemühung (effort) des Sprechenden interpretiert, ist die Wahrnehmung positiv und die Reaktion ist ebenso Akkommodation, falls möglich.
2. Wird die Akkommodation aufgrund des Vorhandenseins äußeren Drucks interpretiert, ist die Wahrnehmung ebenfalls positiv und es wird ebenfalls akkommodiert, jedoch weniger intensiv aufgrund der unklaren Motive des Gegenübers.

Wird nun aber von Seiten des Gegenübers nicht akkommodiert, stehen folgende Interpretationsmöglichkeiten zur Verfügung (vgl. Giles/Powesland 1975: 164):

1. Wird sie als fehlende Bemühung interpretiert, ist die Wahrnehmung negativ und es wird ebenfalls nicht akkommodiert.
2. Wird sie als fehlende Fähigkeit interpretiert, ist die Wahrnehmung negativ, es wird zwar akkommodiert, aber nur, da dem Sprechenden die Fähigkeit abgesprochen wird, nichtakkommodierte Sprache zu verstehen.
3. Wird sie aufgrund äußeren Drucks interpretiert, ist die Wahrnehmung negativ, es kann aber wegen der unklaren Motive des Gegenübers akkommodiert werden.

Das zweite Modell, „Accommodation through accent“, besitzt ein Konvergenz- und ein Divergenzparadigma. Die Spannbreite, innerhalb der akkommodiert werden kann, liegt im österreichischen Kontext potenziell zwischen dem Basisdialekt und der Standardsprache, wobei die idiolektale Kommunikation der Individuen mitberücksichtigt wird. In beiden Paradigmen kann nun nach oben, also in Richtung der Standardsprache, oder nach unten, in Richtung des Basisdialekts, konvergiert oder divergiert werden. Damit sind vier unterschiedliche Verhaltensweisen möglich (vgl. Giles/Powesland 1975: 175):

1. „Upward Convergence“: Der dem Basisdialekt nahestehende Sendende akkommodiert in Richtung des der Standardsprache nahestehenden Rezipierenden.
2. „Downward Convergence“: Der der Standardsprache nahestehende Sendende akkommodiert in Richtung des dem Basisdialekt nahestehenden Rezipierenden.
3. „Upward Divergence“: Der der Standardsprache nahestehende Sendende entfernt sich noch weiter vom dem Basisdialekt nahestehenden Rezipierenden.
4. „Downward Divergence“: Der dem Basisdialekt nahestehende Sendende entfernt sich noch weiter vom der Standardsprache nahestehenden Rezipierenden.

Auslösende Gründe für die unterschiedlichen Handlungsarten geben Giles/Powesland anhand von Beispielen. Entscheidend hierfür sind verschiedene Konstellationen der Variablen Sprachprestige der verwendeten Varietäten/Sprachen und sozialer Status bzw. Machtverhältnis der Interagierenden in unterschiedlichen Handlungskontexten.

### 3.3.4 Audience Design

Bei der Audience-Design-Theorie handelt es sich um eine Erweiterung der Akkommodationstheorie. Haben Giles/Powesland (1975) Akkommodation im dyadischen Kontext konzeptualisiert, erweitert Bell (1984) dies um die Spezifizierung des möglichen Publikums, das damit über die Einzelperson als Adressat:innen in einer dyadischen Struktur hinausgeht. Dabei richtet er sich nicht gegen diese fehlende Perspektive der Akkommodationstheorie, sondern gegen die Aussage eines anderen großen Namens in der Soziolinguistik: „Styles can be ranged along a single dimension, measured by the amount of attention paid to speech“ (Labov 1972: 208 zit. n. Bell 1984: 147).

Für Bell ist Aufmerksamkeit kein hinreichendes Kriterium, um intraindividuelle Sprachvariation erklären zu können. Seine Theorie stützt sich vielmehr auf folgende Annahme: „Variation on the style dimension within the speech of a single speaker derives from and echoes the variation which exists between speakers on the ‚social‘ dimension“ (Bell 1984: 151). Die soziale Dimension ist für Bell untrennbar mit Spracheinstellungen verbunden: „Style shift and evaluation of a variable always co-occur and presuppose that the variable is socially differentiated. This is not surprising, since historically, style differentiation of a variable is derived from social differentiation by way of social evaluation“ (Bell 1984: 151). Diese Zeichen mit sozial differenzierender und evaluativer Semantik nennt er Marker (markers) und trennt sie von Indikatoren (indicators), die zwar ebenfalls unterschiedlich in Sozialgruppen verwendet werden, jedoch nicht diese Form der Semantik besitzen. Diese seien auch aufgrund ihres quantitativ geringen Vorkommens wenig relevant. Im Mittelpunkt von Bells Theorie steht die Taxonomie der Rezipientenseite. Das Zentrum bildet die sehende Person, die ihre Sprach- bzw. Varietätenwahl aufgrund der wahrgenommenen oder hypothetischen Einschätzung ihres Publikums vornimmt. Das Publikum ist dabei jedoch nicht homogen, sondern in unterschiedlicher Weise bedeutsam für die Akkommodation der sehenden Person. In absteigender Reihenfolge sind das (vgl. Bell 1984: 159):

1. Adressat:in (addressee): bekannt, anerkannt und adressiert
2. Zuhörende (auditor): bekannt und anerkannt
3. Lauschende (overhearer): bekannt
4. Horchende (eavesdropper): nicht bekannt.

Die Variationsbandbreite selbst wird vom Sendenden bereitgestellt und kann individuell sehr unterschiedlich sein. Daher kommt ihm auch das größte Gewicht in der potenziell im Umfeld vorhandenen Variationsmöglichkeit zu: „The speaker accounts for most variation, the addressee for less than the speaker, the auditor less again, and the overhearer least of all“ (Bell 1984: 160). Bell stützt sich dabei auf seine Forschung zu Nachrichtensprecher:innen aus Neuseeland. Hier scheint es zur Regel zu gehören, dass die gleichen Nachrichtensprecher:innen für unterschiedliche Sender Nachrichten einsprechen und je nach intendiertem Adressaten/intendierter Adressatin des jeweiligen Radiosenders ihre Sprachwahl anpassen (vgl. Bell 1984: 171 f). Dieses Beispiel zeigt sogleich, dass für Sprachakkommodation der/die Adressat:in nicht zwingend physisch anwesend sein muss, eine imaginierte Form ist hierfür ausreichend.

Diese reaktive Form der adressatenbezogenen Akkommodation ergänzt Bell durch eine zweite, initiative Möglichkeit, die er „Referee Design“ nennt. Dabei wird zu einer Personengruppe hin akkommodiert, die in der Kommunikationssituation nicht anwesend ist, aber für den Sendenden solches Gewicht hat, dass sie seine Sprachwahl beeinflusst. Bell unterscheidet primär in Ingroup und Outgroup Referee Design. Bei Ersterem spricht der Sendende mit einer Outgroup, bedient sich jedoch seiner Ingroupsprache bzw. -varietät. Letzteres unterteilt er in kurzzeitiges oder langfristiges Outgroup Referee Design. Hier benutzt der Sendende eine prestigeträchtige Sprache bzw. Varietät einer Outgroup, um einen bestimmten Effekt zu erzielen. Für gewöhnlich ist dieses Sprachverwendungsmuster nur von kurzer Dauer, doch können, speziell wenn es sich um institutionalisierte Sprachen bzw. Varietäten handelt, auch langfristige Akkommodationsmuster auftreten (vgl. Bell 1984: 186–189). Die beiden Formen, die reaktive und die initiative, erinnern dabei stark an die von Giles/Powesland getroffene Unterscheidung zwischen konvergentem und divergentem Sprachhandeln.

Die Audience-Design-Theorie besitzt auch Ähnlichkeiten mit dem Speaker-Design-Phänomen, das in den letzten Jahren vermehrt untersucht wurde. Hier verlagert sich der Fokus auf gezieltes Style-Shifting von Personen, um in kommunikativen Kontexten bestimmte Effekte erzielen zu können (s. für eine Diskussion im österreichischen Kontext Soukup 2015).

### 3.3.5 Ethnolinguistic Identity Theory

Die Ethnolinguistic Identity Theory (EIT) versucht primär, den Zusammenhang zwischen individueller ethnischer Identität und Spracherhalt bzw. -abbau zu erklären. Sie stützt sich dabei – wie man unschwer am Namen erkennen kann – auf die Annahmen der SIT und führt diese mit denjenigen der Akkommodationstheorie zusammen. Die EIT soll dabei ein Erklärungsmodell bereitstellen, warum manche Mitglieder einer Gruppe in gewissen Situationen ihre ethnische Herkunft mit der Varietäten- bzw. Sprachwahl betonen, während sich andere hier an die Outgroup anpassen (vgl. Giles/Johnson 1987: 69). Sie stützen sich dabei auf Befragungsergebnisse (einer kleinen Stichprobe) von in Wales wohnhaften zweisprachigen Personen und leiten daraufhin folgende Kategorisierung ab. Sie unterscheiden drei psychokognitive Voraussetzungen, anhand derer sich Konvergenz- bzw. Divergenzphänomene erklären lassen. Alle beziehen sich auf Mitglieder untergeordneter ethnischer Gruppen, die folgende Eigenschaften in sich vereinen müssen (vgl. Giles/Johnson 1987: 84 f):

1. Personen werden am ehesten aufgrund ethnischer Solidarität handeln und damit die distinkten Sprachmerkmale (bzw. Sprachen) erhalten, wenn ...
  - a) starke Identifizierung mit der Ingroup besteht.
  - b) unsichere soziale Vergleiche angestellt werden (sie sich auch Alternativen bewusst sind).
  - c) sich Identifizierung mit anderen sozialen Kategorien nicht bzw. nur schwach auf die primäre Identifizierung auswirkt.
 und zusätzlich:

- d) Ingroup-Grenzen als hart und geschlossen wahrgenommen werden und der ethnischen Gruppe niedrige Vitalität zugeschrieben wird.  
oder:
- e) Ingroup-Grenzen als weich und offen wahrgenommen werden und der ethnischen Gruppe hohe Vitalität zugeschrieben wird.
- 2. Personen, die sich weniger stark mit ihrer ethnischen Gruppe identifizieren, die Gruppengrenzen jedoch als hart und geschlossen wahrnehmen, werden sich ebenfalls konform zu ihrer Ingroup verhalten. Distinkte Sprachmerkmale werden erhalten, wenn ...
  - a) moderate Identifizierung mit der Ingroup besteht.
  - b) unsichere soziale Vergleiche angestellt werden (sie sich auch Alternativen bewusst sind).
  - c) sich Identifizierung mit anderen sozialen Kategorien nicht bzw. nur schwach auf die primäre Identifizierung auswirkt.
  - d) Ingroup-Grenzen als hart und geschlossen wahrgenommen werden, der ethnischen Gruppe hohe Vitalität zugeschrieben wird.
- 3. Personen, die sich weniger stark mit ihrer ethnischen Gruppe identifizieren, die Gruppengrenzen als weich und offen wahrnehmen, werden sich an gesellschaftliche Regeln halten. Die distinktiven Sprachmerkmale werden abgebaut, wenn ...
  - a) moderate oder schwache Identifizierung mit der Ingroup besteht.
  - b) sichere soziale Vergleiche angestellt werden (Alternativen werden nicht wahrgenommen).
  - c) sich Identifizierung mit anderen sozialen Kategorien stark auf die primäre Identifizierung auswirkt.
  - d) Ingroup-Grenzen als weich und offen wahrgenommen werden.
  - e) Der ethnischen Gruppe niedrige Vitalität zugeschrieben wird.

Diese dispositionalen Eigenschaften der drei Gruppen reichen jedoch nicht aus, um jegliches Akkommodationsverhalten erklären zu können. Dafür ist auch die Einbettung in einen Kontext ausschlaggebend. Dies wird anhand situationeller Variablen vorgenommen. Die beiden dargestellten Situationen 4. und 5. können das Verhalten der Personengruppen 1. – 3. modifizieren (vgl. Giles/Johnson 1987: 87):

- 4. Wenige Mitglieder einer untergeordneten ethnischen Gruppe zeigen in interethnischen Kommunikationssituationen zur Norm divergierendes (Sprach)handeln, wenn sie ...
  - a) nicht stark eine Einstellungsposition vertreten, die gerade herausgefordert wird.
  - b) Zeichen als statusbetonend in der Situation deuten.
  - c) kommunikative Effizienz beibehalten wollen.
  - d) große Überschneidung zu anderen Gruppenmitgliedschaften wahrnehmen.
- 5. Viele Mitglieder einer untergeordneten ethnischen Gruppe zeigen in interethnischen Kommunikationssituationen zur Norm divergierendes (Sprach)handeln, wenn sie ...

- a) stark eine Einstellungsposition vertreten, die gerade herausgefordert wird.
- b) Zeichen nicht als statusbetonend in der Situation deuten oder der Interagierende soziale Normen verletzt hat.
- c) kommunikative Effizienz nicht beibehalten wollen.
- d) geringe Überschneidung zu anderen Gruppenmitgliedschaften wahrnehmen.

Giles/Johnson (1987) binden dieses Verhaltensmodell nun auch an die Entstehung von Einstellungen, was speziell im Kontext dieser Arbeit von Interesse ist. Hier wird wiederum auf die drei Personengruppen (1. – 3.) zurückgegriffen und deren Evaluations-schemata beschrieben:

1. Personen, die aufgrund ethnischer Solidarität handeln, bewerten Personen, die die ethnische Gruppe unterstützen, und gleichzeitig die Ingroup-Sprachverwendung mit hoher Wahrscheinlichkeit positiv, speziell auf der salienten Ingroup-Dimension Solidarität.
2. Personen, die sich konform zu ihrer Ingroup verhalten, bewerten die Ingroup-Sprachverwendung mit hoher Wahrscheinlichkeit positiv die Solidarität betreffend und negativ den Status betreffend.
3. Personen, die sich an gesellschaftliche Regeln halten, bewerten die prestigehöhere Sprachverwendung auf der Statusdimension mit hoher Wahrscheinlichkeit positiver als die prestigieniedere Ingroup-Sprachverwendung.

Damit liefern Giles/Johnson (1987) ein differenziertes Erklärungsmodell, welches Sprachhandeln aufgrund von Spracheinstellungen und weiterer sozialer Faktoren erklärt.

### 3.3.6 Selbstkategorisierungstheorie

Bei der Selbstkategorisierungstheorie (SCT) handelt es sich ebenfalls um eine Theorie, die in engem Zusammenhang mit der SIT steht. Sie soll allerdings keine Weiterentwicklung der SIT darstellen, sondern diese ergänzen, indem sie die psychologisch-kognitiven Vorgänge der sozialen Identitätsbildung in den Fokus nimmt. Entwickelt wurde sie vom Tajfel-Schüler John Turner, der sie mit den Mitgliedern seiner Arbeitsgruppe in der Monografie Turner/Hogg et al. (1987) der Öffentlichkeit präsentierte. Zusammenfassungen der Kernelemente dieser im Detail recht komplexen Theorie findet man bei Hornsey (2008) und Turner/Reynolds (2012). Gleichzeitig werden in letzterem Aufsatz auch die aktuelleren Forschungsinteressen der SCT-Forschenden aufgelistet.

Die Theorie geht, ähnlich der SIT, von einem Kontinuum der Identität aus. Sie beschreibt dieses im Unterschied dazu jedoch in drei Stufen und fokussiert dabei die kognitiven Prozesse, andere Personen oder auch sich selbst in Kategorien einzuteilen. Bei der Kategorisierung bezieht sie sich auf die Prototypentheorie von Rosch (1978). Die oberste, superordinierende Ebene kategorisiert das Selbst als menschliches Wesen in Abgrenzung zur Tierwelt, die mittlere oder Intergruppenebene als Gruppenmitglied im Kontrast zu einer Outgroup und die unterste, interpersonelle Ebene als Indi-

viduum im Kontrast mit anderen Individuen. Dies sind jedoch nur die zentralsten Beschreibungsebenen. Es ist z. B. ohne Weiteres möglich, eine intraindividuelle Ebene zu beschreiben, in der das Individuum Einzelkategorien innerhalb des Selbstbildes gegeneinander abwägt (vgl. Turner/Reynolds 2012: 404). Die einzelnen Ebenen stehen auch in einer Beziehung des „funktionellen Antagonismus“. Gewinnt eine Ebene an Salienz, verlieren die anderen diese (vgl. Hornsey 2008: 208). Aber welche der Ebenen dieses Selbstkategorisierungs-Kontinuums ist in der jeweiligen Situation nun relevant? Die SCT erklärt dies mit Zugänglichkeit (accessibility) und zwei Formen der Passgenauigkeit (fit): vergleichende (comparative) und normative (normative) Passgenauigkeit (vgl. Hornsey 2008: 208):

1. Dabei ist mit Zugänglichkeit gemeint, ob und wie häufig in der vorliegenden Situation relevante Kategorien evoziert werden.
2. Die vergleichende Passgenauigkeit folgt dem Metakontrastprinzip. Dies bedeutet, dass die Passgenauigkeit dann hoch ist, wenn die Unterschiede innerhalb einer Kategorie so gering und die Unterschiede zwischen den Kategorien so groß wie möglich sind.
3. Die normative Passgenauigkeit ist dann hoch, wenn das Gruppenverhalten und die Gruppenmitgliedschaft sich an stereotypen Erwartungshaltungen orientieren.

Es existiert nun ein prägnanter Unterschied zwischen der individuellen und der Gruppenkategorisierung, der als Depersonalisation bezeichnet wird und dem zentrale Bedeutung in der SCT zukommt. Soziale Gruppen werden kognitiv als Prototypen mit bestimmten Eigenschaften repräsentiert und eben nicht als Konglomerat verschiedener Einzelindividuen, die ähnliche Eigenschaftsbündel miteinander vereint. Dies führt dazu, dass Gruppenmitglieder weniger als Individuen, sondern als „austauschbare Exemplare“ des Gruppenprototypen betrachtet werden (vgl. Hornsey 2008: 208; Turner/Reynolds 2012: 403).

Turner/Reynolds (2012: 406) fassen die Kernelemente der SCT folgendermaßen zusammen:

1. Human beings are both individuals and group members and therefore have both personal identity and social identity.
2. People can define or categorize themselves at different levels of abstraction.
3. Salience explains the way a particular situation (that includes the self) is categorized and given ‚meaning‘.

Einen interessanten Anknüpfungspunkt an Spracheinstellungen, spezifisch an die Normierung von Einstellungen und Verhalten, findet man ebenfalls in der SCT. Die Stufen werden folgendermaßen beschrieben (Turner 1987 zit. n. Turner/Reynolds 2012: 408):

1. Individuals define themselves as members of a distinct social category.
2. They learn or develop the appropriate, expected, desirable behaviours that are correlated with category membership, and differentiate it from other categories (e. g. the stereotypical norm).

3. They assign the norms and attributes of the category to themselves (internalization) through the process of depersonalization and self-stereotyping.
4. Their behaviour therefore becomes normative as their category membership becomes more salient.

Damit stellt die SCT eine breite und allgemeine Interpretations- und Kontextualisierungsmöglichkeit für soziale Prozesse bereit, unter die auch Spracheinstellungen und die damit verbundenen Auf- und Abwertungsmuster zwischen Sprachgruppen fallen. Auch die Analyse von Machtverhältnissen in Bezug auf Sprachgruppen (z. B. Godley/Reaser 2018; Murray 2001) und die Sprachideologieforschung (Irvine/Gal 2000) kann von der SCT profitieren, da sich einer der aktuelleren Forschungsbereiche der SCT mit Führungsqualitäten auseinandersetzt. „The breakthrough idea is that leadership rests on an individual's ability to be seen as prototypical of a shared social identity and hence will have greater influence as a result of such categorization processes“ (Turner/Reynolds 2012: 410).

### 3.3.7 Sprachideologie

Den Abschluss in diesem Kapitel bildet der Bereich der Sprachideologie. Für die Definition und die Funktionsweise von Sprachideologie ist der Aufsatz von Irvine/Gal (2000) als wegweisend zu nennen. Daher bietet es sich an, anhand der Systematisierung und der Funktionsweise von Sprachideologien, die in diesem Aufsatz unternommen werden, in dieses Forschungsthema einzuführen.

Um zu verstehen, wie Sprachideologien funktionieren, ist es in einem ersten Schritt wichtig, die grundsätzliche Funktionsweise von Ideologien zu betrachten. Sie beschreiben Ideologien als „conceptual schemes“ oder auch „folk theories“ (Irvine/Gal 2000: 35 bzw. 78), die sich aus einer Sammlung von Überzeugungen zusammensetzen und die aus Differenz- bzw. Oppositionspaaren bestehen, wie es etwa beim konzeptuellen Schema von Männlichkeit und Weiblichkeit zu beobachten ist. So existieren bestimmte Konventionen in der Gesellschaft, wie man sich als Mann/als Frau zu kleiden hat. Die Aktivierung dieses Schemas, etwa beim Einkaufen von neuem Gewand, oder der Verletzung des Schemas durch eine Person in der Öffentlichkeit, beinhaltet immer auch die Aktivierung des gegenteiligen Schemas, also wie sich das andere Geschlecht zu kleiden hat. Irvine/Gal (2000) definieren drei Prozesse, mit denen die Funktionsweise von Ideologien (und eben auch Sprachideologien) beschrieben werden kann: „Fractal Recursivity“, „Erasure“ und „Iconization“.

Den Prozess der Fractal Recursivity beschreiben sie als „projection of an opposition, salient at some level of relationship, onto some other level“ (Irvine/Gal 2000: 38). Bei der konventionellen Vorstellung von angemessener Kleidung von Männern und Frauen wären etwa Unterschiede festzustellen, ob die Person sich im privaten oder im öffentlichen Raum befindet. Im privaten Bereich liegen andere, eventuell lockere Konventionen für die Kleidung vor als im öffentlichen Raum. Es können aber in den einzelnen Domänen noch feinere Unterschiede ausgemacht werden: Im öffentlichen Raum etwa existieren große Unterschiede in der Kleidungskonvention zwischen einem

Strand- und einem Opernbesuch. Trotz der unterschiedlichen Domänen existieren grundsätzliche Regeln: Nur in wenigen Fällen wäre es etwa angemessen, nackt zu erscheinen, und es scheint ausgeschlossen, mit einem großen Karton bekleidet in jeglicher Situation angemessen erscheinen zu können (außer vielleicht im Fasching oder im Theater).

Erasure beschreiben Irvine/Gal (2000: 38) als „rendering invisible some persons or activities that do not fit a given ideological understanding“. Dieser Prozess ist typisch für politische Argumentation und kann vielfach beobachtet werden. Klassisch wäre etwa die Migrationsdebatte in Österreich, in der für gewöhnlich die Zuwanderung aus islamischen Ländern kritisiert wird. Dass Zuwanderung nach Österreich aber auch aus Ländern wie den USA, Frankreich oder Norwegen stattfindet und die mit Abstand größte Zuwanderergruppe in Österreich Personen aus Deutschland sind, wird in der Debatte ausgeblendet.

Den dritten Prozess beschreiben Irvine/Gal (2000: 37) folgendermaßen: „*Iconization* involves a transformation of the sign relationship between linguistic features (or varieties) and the social images with which they are linked [Kursivsetzung i. O.].“

Diese drei Prozesse können nun auch in der österreichischen Varietätenlandschaft beobachtet werden. Die bairischen Dialekte unterscheiden sich zum Standarddeutschen etwa durch unterschiedliche Genuszuweisung zu Substantiven (*der Butter, der Teller*) oder auch dem Fehlen von grammatischen Merkmalen wie dem nicht vorhandenen Genitiv (std.: *der Hut des Vaters* vs. bair.: *der Hut vom Vater, dem Vater sein Hut*). Bekannte Vorurteile schätzen diese Merkmale als „schlampige Sprache“ oder „defekte Sprachverwendung“ ein (vgl. Abschnitt 2.2).

Fractal Recursivity liegt einerseits vor, da ein Gegensatz „schlampig“ vs. „rein“ konstruiert wird, andererseits werden diese Bewertungen auch auf andere Ebenen projiziert. So werden sie in einem ersten Schritt auf die individuellen Sprecher:innen und in einem zweiten auf die ganze Sprachgruppe übertragen: Urteile wie „Dialektsprecher:innen sind schlampig“ oder sogar „weniger intelligent“ lassen sich empirisch nachweisen (s. Kapitel 4). Um eine solche Ideologie langfristig zu bewahren, benötigt es zusätzlich den Prozess der Iconization. Dadurch wird eine Zeichenverbindung zwischen linguistischen Eigenschaften wie „fehlerhafte Genus- und Kasusverwendung“ und sozialen Eigenschaften wie „fehlerhafte (also schlampige, unintelligente) Dialektsprecher:innen“ aufgebaut und aufrechterhalten. Der Prozess der Erasure kommt ins Spiel und „bereinigt“ die ideologische Vorstellung um die Variabilität in der bairischen Dialektlandschaft und auch um die Tatsache, dass viele Personen in dieser Sprachregion nicht nur Dialekt, sondern auch die Standardsprache beherrschen.

Die Theorie der Sprachideologie stellt damit ein nützliches Werkzeug bereit, die unterschiedlichen Einstellungsmuster, die in der Gesellschaft vorliegen, analysieren zu können.

### 3.3.8 Zwischenfazit

Die präsentierten Theorien liefern Belege für das Mehrkomponentenmodell. So bietet die SCT einen Erklärungsansatz für die kognitive Komponente von Spracheinstellungen: Spracheinstellungen werden als Prototypen gespeichert. Des Weiteren muss der Gruppenkategorisierung eine stärkere Rolle als der individuellen Kategorisierung zukommen, da sich Spracheinstellungen primär aus dieser Gruppenkomponente speisen. Damit kommt der Depersonalisation eine wesentliche Rolle in der Genese von Spracheinstellungen zu – der Schritt zur Stereotypisierung bzw. zur Bildung einer Sprachideologie ist somit nur mehr ein kleiner. Bei Dragojevic (2017: 1), der sich nicht auf die SCT bezieht, aber zu einer ähnlichen Schlussfolgerung kommt, klingt das folgendermaßen:

„Language attitudes are evaluative reactions to different language varieties. They reflect, at least in part, two sequential cognitive processes: social categorization and stereotyping. First, listeners use linguistic cues (e. g., accent) to infer speakers' social group membership(s). Second, based on that categorization, they attribute to speakers stereotypic traits associated with those inferred group membership(s)“.

Auch Soukup (2019) liefert ein primär kognitiv orientiertes Modell von Spracheinstellungen. Sie geht von „human epistemological constructs (HEC)“ als zentrale Einheit von Einstellungen aus: „HECs sind im Prinzip in der sozialen Interaktion entwickelte, zu kognitiven Gebilden zusammengefügte und als solche auch speicherbare sprachreflexive Diskurse, die als kommunikative ‚Orientierungshilfen, Bezugspunkte, [!] und Ordnungsstrukturen‘ dienen und welche ‚wir Menschen [...] erlernen, erstellen, verwenden, [!] und anpassen“ (Soukup 2019: 96 f).

Viele der anderen Theorien sind an der Schnittstelle der affektiven und der Verhaltenskomponente angesiedelt. So liefert die SIT eine allgemeine Erklärung des Zusammenhangs von sozialer Identität und Handlung. Die Ethnolinguistic Identity Theory spezifiziert die SIT und schafft eine Verknüpfung von Spracheinstellungen und Sprachhandeln, indem sie sich der Erkenntnisse aus der Akkommodationstheorie bedient. Die Akkommodationstheorie selbst sieht Sprachprestige und den sozialen Status der Interagierenden als zentrale Faktoren.

Auch der Kontext scheint eine bedeutende Rolle bei Spracheinstellungen zu spielen. So kann man die Kontextualisierung zwar ebenfalls als kognitive Komponente von Spracheinstellungen beschreiben, die jedoch auch eine zusätzliche Ebene öffnet, da der Kontext zwar kognitiv verarbeitet wird, jedoch nicht in der vorher besprochenen Weise, z. B. als Stereotyp, gespeichert ist. Ähnlich verhält es sich auch mit der Theorie des Audience Design, in deren Zentrum die reale, intendierte oder imaginierte Adressat:innen schaft steht, die ebenfalls Einfluss auf Register- bzw. Varietätenwahl hat. Schließlich scheinen auch intraindividuelle psychokognitive Faktoren wie die Stärke des Self-Monitoring eine Rolle dabei zu spielen, wie stark die Kontextsensitivität beim Individuum ist.

Bei Spracheinstellungen handelt es sich also um einen konzeptuell vielschichtigen und theoretisch nicht leicht zu fassenden Untersuchungsgegenstand. Das Mehrkompo-

nentenmodell eignet sich, um die zentralen Bestandteile zu beschreiben. Die kognitive Komponente verarbeitet eine Einstellung zu einer Varietät//Sprache und speichert diese energiesparend als Prototyp (bzw. Stereotyp). Solche Einstellungsbündel können als Sprachideologien beschrieben werden. Die affektive Komponente stellt die Verbindung zwischen Sprache//Varietät und persönlicher Identität her. Beide Komponenten wirken sich auf die Verhaltenskomponente aus. Wie stark die Verhaltensanpassung ausfällt und wie gut sie vorhergesagt werden kann, hängt von individuellen und kontextuellen Faktoren ab.

Spracheinstellungen verfügen darüber hinaus zumindest über zwei Dimensionen, eine Dimension der Kompetenz//Korrektheit//Status und eine der sozialen Attraktivität//Solidarität/Freundlichkeit. Ich werde im praktischen Teil der Arbeit die beiden Dimensionen der Einfachheit halber als Kompetenz- und Sympathieebene bezeichnen.

### 3.4 Methoden der (Sprach)-Einstellungsforschung

Bei der Messung von Einstellungen wird zwischen expliziten und impliziten Methoden unterschieden. Typischerweise werden explizite Einstellungen über direkte Fragen und deren Beantwortung auf einer mehrteiligen Skala erhoben. Dabei kommen zumeist Likert-Skalen oder semantische Differenziale zur Anwendung. Auf einer Likert-Skala kann die Zustimmung oder Ablehnung in gradueller Abstufung gegeben werden. Die Skala beim semantischen Differenzial funktioniert insofern anders, da an ihren Extrempolen durch zwei konträre Adjektive (z. B. gut vs. schlecht) angegeben werden und damit unterschiedliche Dimensionen von Zustimmung und Ablehnung zur gleichen Frage elitiziert werden können (s. auch Stroebe/Stroebe/Hewstone 2007: 207 ff). Auch Spracheinstellungen werden häufig auf diese Weise erhoben.

Bei impliziten Einstellungsmaßen gehen die Methoden zwischen Sozialpsychologie und Soziolinguistik etwas auseinander. In der allgemeinen Einstellungsforschung werden hier Methoden wie das evaluative Priming oder der Implicit Association Test herangezogen (Stroebe/Stroebe/Hewstone 2007: 210 f). In der Spracheinstellungsforschung ist hier als zentrale Methode die Matched-Guise-Technik zu nennen, die auf Lambert et al. (1960) zurückgeführt werden kann. Sie wurde seither in der soziolinguistischen Einstellungsforschung häufig verwendet.<sup>3</sup> Die Technik besteht darin, dass Audioaufnahmen mit semantisch gleichem Inhalt in unterschiedlichen Varietäten oder Sprachen aufgezeichnet werden. Bei den sprechenden Personen handelt es sich um bivarietäre Sprecher:innen, die die Inhalte in beiden Varietäten einsprechen. Dadurch erhält man die sogenannten Matched Guises. Den ratenden Personen werden die Aufnahmen der jeweiligen Personen aber nicht nacheinander, sondern durch Füllaufnahmen oder Durchmischung der Aufnahmen getrennt voneinander vorgespielt. Ob erstere oder letztere Variante, das Ziel ist, dass es den ratenden Personen nicht auffällt, dass es sich um die gleiche Person handelt, die das Guisepaar eingesprochen

---

3 S. für einen Überblick Dragojevic 2017: 3, im deutschsprachigen Raum z. B. Häcki Buhofer/Studer 1993, Bellamy 2012, Kaiser/Kasberger 2021.

hat. So können Ratingunterschiede rein auf die Varietäten zurückgeführt und persohnenspezifische Faktoren wie Stimme, Stimmlage, Betonung usw. ausgeschlossen werden. Die Guises werden im Regelfall auf einer Likert-Skala auf einer oder mehreren Dimensionen eingeschätzt. Eng verwandt mit der Matched-Guise-Technik ist die Verbal-Guise-Technik, bei der unterschiedliche Sprecher:innen die Guises einsprechen.

Daneben existieren auch andere Methoden wie die direkte Elizitierung durch ethnolinguistische Gruppenlabel, Inhaltsanalyse von Interviews und ethnografische Studien (vgl. Dragojevic 2017: 3) sowie perzeptionslinguistische Ansätze (s. Preston 2010; Preston 2019).

## 4 Erkenntnisse aus der Spracheinstellungsforschung

Die Entwicklung der Spracheinstellungen zu den Varietäten Dialekt und Standardsprache im deutschen Sprachraum lässt sich darauf zurückführen, dass gegen Ende des 18. Jhs. die Schriftsprache allmählich den Status einer Standardsprache erreichte und diesen im 19. Jh. durchsetzte (vgl. Besch 2003: 2252 f, s. auch Abschnitt 2.2). Ab dieser Zeit sind vermehrt pejorative Aussagen über dialektales Sprechen auszumachen (vgl. Ziegler 1999: 83). So ist es einigermaßen verwunderlich, dass soziolinguistische Forschung im Allgemeinen und Spracheinstellungsforschung im Speziellen erst in den 60er-Jahren des 20. Jhs. aufkamen. In dieses Jahrzehnt fallen, wie schon erwähnt, die Entwicklung der Matched-Guise-Technik (Lambert et al. 1960) sowie die ersten Studien von Labov (nachzulesen in Labov 1969), der prägend für die Soziolinguistik war. Die Erforschung der Spracheinstellungen lief dabei parallel zur Variablen- und Variantenanalyse, die den vorherrschenden Schwerpunkt der soziolinguistischen Forschung bildete, und lieferte einen weiteren Blickwinkel auf die Zusammenhänge zwischen sozialen Strukturen und sprachlichen Verwendungsmustern. Da die beiden Forschungsgebiete eng miteinander verflochten sind, lohnt es sich, einen kurzen historischen Überblick über die Entwicklung der soziolinguistischen Forschung zu geben, der die Spracheinstellungsforschung häufig zugeordnet wird.

Eckert (2012) macht in dieser Hinsicht drei Wellen der variationslinguistischen Forschung aus. In der ersten Welle wurde versucht, Korrelationen zwischen der Verwendung bestimmter linguistischer Variablen und sozialen Kategorien wie sozialer Schicht, Alter, Geschlecht, ethnischer Zugehörigkeit usw. zu untersuchen (vgl. Eckert 2012: 88–91; Wardhaugh/Fuller 2015: 170 f). Als Vorbild für die Ausrichtung der Forschung in dieser Zeit gilt die Kaufhausstudie von Labov in New York (zu finden in Labov 1969), in der er die Aussprache von /r/ in unterschiedlichen Kaufhäusern untersuchte, die einer unteren, einer mittleren und einer oberen Schicht zugeordnet werden konnten.

Ein weiterer wichtiger Vertreter dieser Zeit ist Trudgill (1972), der in seiner Norwich-Studie ebenfalls die unterschiedliche Verwendung von Variablen verschiedener sozialer Gruppen im Visier hatte. Er erweiterte das Forschungssetting jedoch um potenzielle Geschlechterunterschiede in der Varietätenverwendung. Es schlossen sich viele weitere Studien an, auch außerhalb des englischsprachigen Kontexts, in denen sich die folgenden Muster herauszeichneten: In niederen sozialen Schichten werden Variablen verwendet, die kleinräumig differenziert sind und denen eine gewisse Stigmatisierung anhaftet. Je höher sich Personen auf der sozialen Hierarchieleiter befinden, desto mehr werden diese stigmatisierten Variablen gemieden und stattdessen standardsprachliche Variablen verwendet, die diese kleinräumige Differenzierung nicht besitzen (vgl. Eckert 2012: 88). In Bezug auf Geschlechterunterschiede wurde in

dieser Zeit ebenfalls ein Muster entdeckt: Frauen bzw. Mädchen verwenden mehr standardnahe Formen als Männer bzw. Buben der gleichen sozialen Schicht (vgl. Wardhaugh/Fuller 2015: 171 f).

In der zweiten Welle änderte sich der Fokus der Forschung insofern, als mehr auf die soziale Kraft der Verwendung von unterschiedlichen Varietäten geachtet wurde. Diese Veränderung wurde begleitet durch einen Methodenwechsel hin zu ethnografischen Beobachtungen. Ausgelöst wurde die zweite Welle durch das Interesse von James und Lesley Milroy (z. B. Milroy 1989) am Zusammenhang von Varietätenverwendung von Individuen und der sozialen Netzwerke, in die diese Personen eingebunden sind (vgl. Eckert 2012: 91). Sie fanden einen Zusammenhang zwischen der Dichte der sozialen Netzwerke und der normativen Kraft, die diese auf die Sprachverwendung der einzelnen Mitglieder ausübte. Ebenso zeigte sich, dass die Verwendung von lokalen Varietäten eng mit dem Engagement in lokalen Netzwerken zusammenhing. Es wurde damit der Fokus von allgemeinen sozialen Kategorien hin zu lokalen bzw. regionalen Faktoren innerhalb des sozialen Netzwerkes gelenkt. So wurde als eines der zentralen Ergebnisse die enge Verflechtung von Identität, Varietätenverwendung und sozialer Gruppe gefunden (vgl. Eckert 2012: 91 ff).

Allerdings konnte mit diesem Fokus die oft gezielte stilistische Variation der kommunizierenden Personen nicht erklärt werden. In der dritten Welle nahm man sich also der Indexikalitätsrelation zwischen Variablen und sozialen Kategorien an, indem man stilistische Anpassungen an den sozialen Kontext von Personen in den Blick nahm (vgl. Eckert 2012: 93 f). Eckert spricht in diesem Zusammenhang von einem indexikalischen Feld:

I argue that the meanings of variables are not precise or fixed but rather constitute a field of potential meanings – an *indexical field*, or constellation of ideologically related meanings, any one of which can be activated in the situated use of the variable [Kursivierung im Original] (Eckert 2008: 453).

Sie stützt sich dabei auf die Forschung von Agha (2003)/Agha (2004) und sein zentrales Konzept des Enregisterment, das wiederum auf dem Konzept der sprachlichen Register aufbaut: „A register is a linguistic repertoire that is associated, culture-internally, with particular social practices and with persons who engage in such practices“ (Agha 2004: 24). Enregisterment beschreibt den Prozess, der zur Herausbildung eines solchen Registers führt:

A register exists as a bounded object only to a degree set by sociohistorical processes of enregisterment, processes by which the forms and values of a register become differentiable from the rest of the language (i. e., recognizable as distinct, linked to typefiable social personae or practices) for a given population of speakers [Heraushebung im Original] (Agha 2004: 37).

Diese Konzepte, Enregisterment wie auch das indexikalische Feld, zeigen die wandelbare und kontextuelle Struktur von sozialen Bedeutungszuschreibungen zu sprachlichen Variablen. Wie schon in Abschnitt 3.3 ausgeführt wurde, sind auch Sprachein-

stellungen kontextsensitiv. Gleichzeitig fungieren sie als zentraler evaluativer Mechanismus, durch den soziale Indexikalität erst möglich wird.

Die Ergebnisse der Spracheinstellungsforschung selbst können meiner Ansicht nach nicht so trennscharf in die drei Wellen der Variationsforschung eingeordnet werden, wie dies mit den Ergebnissen aus den Variantenanalysen möglich ist. Dies lässt sich meiner Ansicht nach mit der unterschiedlichen Methodologie der beiden Forschungsbereiche erklären. In der Spracheinstellungsforschung wurde ein großer Teil der Ergebnisse mit der Matched-Guise-Technik erhoben (vgl. Dragojevic 2017: 3).<sup>4</sup> Das Interesse an der Anwendung dieser Methode lag nun darin, Eigenschaften von den Proband:innen zu elizitieren, die sie z. B. dialektsprechenden oder standardsprechenden Personen zuschrieben (s. Abschnitt 3.2). Diese Eigenschaften (Dialektsprechende schneiden besser auf Sympathieebene ab; Standardsprechende auf Kompetenzebene) komplettierten das Bild, das in der Soziolinguistik z. B. zwischen Schichtzugehörigkeit und Sprachverwendung gezeichnet wurde. Eine spezifische Korrelation von Schichtzugehörigkeit mit den Einstellungen lag daher nicht im Zentrum des Forschungsinteresses.

Unterschiedliche soziale Schichten oder Gruppen, die sich unterschiedlicher Varietäten bzw. Sprachen bedienen und in einem bestimmten Nähe- bzw. Konkurrenzverhältnis stehen, spielen jedoch in der Konzeption der Guises eine wesentliche Rolle. So wurden im Zuge der erstmaligen Anwendung dieser Methode unterschiedliche Einstellungen von französischsprachigen und englischsprachigen Kanadier:innen in Quebec eliziert. Dabei eignen sich prinzipiell alle regionalen Kontexte, in denen entweder innerhalb einer Einzelsprache unterschiedliche Varietäten (sowie Lerner Sprachen und Akzente) oder auch unterschiedliche Sprachen (und auch eine Mischung von beiden) vorkommen. (Einen detaillierten Überblick über die Forschung im englischsprachigen Raum mit einem kurzen Blick über den Tellerrand liefert Garrett 2010: 53–87; einen Kurzüberblick liefert Dragojevic 2017: 3).

Die Forschung im deutschsprachigen Raum (in dieser Arbeit mit einem Fokus auf Österreich) lief teils ebenfalls in diese Richtung, teils wurden auch spezifische Fragestellungen wie die Einstellungen zum bundesdeutschen Standard in Österreich (Kaiser 2006) oder unterschiedliche Normvorstellungen in den deutschsprachigen Ländern (Peter 2015) unter die Lupe genommen. Gleichzeitig findet man vereinzelt Ergebnisse zu Spracheinstellungen bei unterschiedlichen Geschlechtern bzw. im Migrationskontext (z. B. Fuchs/Elspaß 2019; Vergeiner et al. 2019; Pribyl-Resch 2021).

---

4 Diese Technik verwendet Sprachproben in zumindest zwei Varietäten (z. B. Dialekt und Standardsprache), die von der gleichen Person gesprochen wurden. Dass es sich um die gleiche Person handelt, wird mit Füllsätzen bzw. vielen unterschiedlichen Guises verschleiert (s. auch Abschnitt 3.4).

## 4.1 Dialekt und Standardsprache

Das zentrale Interesse der Spracheinstellungsforschung liegt darin, unterschiedliche Einstellungen zu prestigereicheren und -ärmeren Varietäten bzw. Sprachen besser zu verstehen. Im Kontext der inneren Mehrsprachigkeit geht dies häufig mit dem Varietätenkontinuum Dialekt – Standardsprache einher – so auch in Österreich. Eine frühe Arbeit dazu liefert Moosmüller (1990). Sie lässt verschiedene Sprachproben von Akademiker:innen aus Graz, Wien, Innsbruck und Salzburg von 84 Personen, die ebenfalls aus diesen Städten stammten, sowohl regional (die neun Bundeshauptstädte) als auch sozial (verschiedene Berufsgruppen) zuordnen. Dabei findet sie heraus, dass den Sprachproben aus Wien und Salzburg der höchste Status zugesprochen wurde und schlussfolgert, dass diese Varietäten am ehesten prototypisch als österreichischer Standard bezeichnet werden können (vgl. Moosmüller 1990: 116).

Im selben Jahr präsentiert Mattheier (1990) Sprachverwendung in Deutschland, in dem er neben der abnehmenden Dialektkompetenz von Süden nach Norden, abnehmendem Dialektgebrauch zwischen Land und Stadt, Zusammenhängen zwischen Sprachverwendung und Beruf und Zusammenhängen zwischen Sprachverwendung und Geschlecht auch Daten zu Spracheinstellungserhebungen in den Städten Heidelberg und Bonn zeigt. Dabei ist ein Zusammenhang von positiven Spracheinstellungen gegenüber dialektalen Varietäten mit Ortsloyalität auf der einen und jüngeren Sprecher:innen auf der anderen Seite zu erkennen.

Aktuellere Untersuchungen von Spracheinstellungen in Deutschland liefern Adler/Plewina (2019). Sie erheben anhand semantischer Differenzskalen die Einschätzungen von Deutsch (was für die Standardsprache steht), Bayrisch, Sächsisch und Plattdeutsch und stellen fest, dass Deutsch in den Kategorien logisch und systematisch weit besser abschneidet als die dialektalen Varietäten (Adler/Plewina 2019: 144).

Soukup (2009) fokussiert ebenfalls Spracheinstellungen zu Dialekt- und Standardsprecher:innen im österreichischen Kontext und findet die klassischen Zuschreibungen. Sie verwendet dafür eine direkte Methode via semantischem Differenzial:

The results showed that the dialectal speakers were perceived as more natural, honest, emotional, relaxed, and likeable than their standard speaking peers, as well as having a better sense of humor. Yet they were also judged to sound more aggressive. By contrast, the standard speakers were perceived as more polite, intelligent, educated, gentle, serious, and refined, but also as sounding more arrogant. The results furthermore suggest that standard speakers could be regarded as tendentially more competent, more industrious, cleverer, but also stricter than dialect speakers, although the outcomes were not as conclusive in these respects (Soukup 2009: 127).

Bellamy (2012) untersucht Spracheinstellungen in Wien, die er mit der Matched-Guise-Technik erhebt. Interessanterweise findet er Abweichungen zu den typischen Ergebnissen. So wird die Status-Eigenschaft „confidence“ mehr den dialektalen Varietäten zugeschrieben und Standardsprecher:innen werden ehrlicher und großzügiger

eingeschätzt (vgl. Bellamy 2012: 210). In anderen Status- und Solidaritätsdimensionen findet Bellamy die typischen Zuschreibungen zu Dialekt- und Standardsprechenden.

Fuchs/Elspaß (2019) erheben Spracheinstellungen in direktem Messverfahren bei Schüler:innen einer 10. Schulstufe in Handelsschulen und Handelsakademien in Österreich. Es zeigen sich die altbekannten Einstellungsmuster: So wird Sprechen im Dialekt als glaubwürdig, echt und sicher eingeschätzt, jedoch als weniger kompetent. In standardsprachlicher Sprechweise steigen die Werte auf Staturebene an und die Werte auf Solidaritätsebene gehen zurück, wobei man erkennt, dass letztere Werte auch in standardsprachlicher Sprechweise vergleichsweise hoch sind (vgl. Fuchs/Elspaß 2019: 22).

Kasberger/Kaiser (2019) und Kaiser/Kasberger (2021) gehen der Frage nach, in welchem Alter Spracheinstellungen erworben werden. Sie sehen sich hierzu Kindergarten- und Schulkinder im Alter zwischen drei und zehn Jahren an und kommen zum Schluss, dass „bekannte Einstellungen gegenüber den Varietäten vereinzelt bereits im Kindergartenalter ausgedrückt werden und dass diese bis zum Ende der Volksschulzeit eine deutliche Steigerung erfahren“ (vgl. Kasberger/Kaiser 2019: 17).

Im Rahmen einer qualitativen Studie geht Kraftl (2020) Spracheinstellungen von drei Müttern in Österreich nach. Auch sie findet die altbekannten Muster, indem die Standardsprache (speziell der bundesdeutsche Standard) auf Staturebene, der Dialekt auf Solidaritätsebene positiver abschneiden. Sie verschränkt diese Erkenntnisse mit der Sprachverwendung von Eltern mit ihren Kindern und findet hier insofern Zusammenhänge, dass zwei der drei Mütter angeben, mit ihren Kindern Standard zu sprechen, obwohl sie selbst dialektal aufgewachsen sind (vgl. Kraftl 2020: 75).

Ender (2020) untersucht den Zusammenhang zwischen Dialektkompetenz und -bewertung bei Personen mit Deutsch als L1 und L2. Sie wendet für die Messung von Spracheinstellungen die Matched-Guise-Technik an, in der sie Sprachproben von Verkäufer:innen und Ärzt:innen verwendet und somit unterschiedliche Situationen evokiert. Die Ergebnisse zeigen nun, dass dialektale Sprechweise zwar generell etwas niedrigere Werte erzielt, diese im Verkaufskontext jedoch sehr nahe an der standardsprachlichen liegen, im medizinischen Kontext die Verwendung der Standardsprache jedoch deutlich positiver eingeschätzt wird (vgl. Ender 2020: 97). Diese Ergebnisse bestätigen das bisher gezeigte Bild von Spracheinstellungen in Österreich.

Pribyl-Resch (2021) untersucht qualitativ erhobene Spracheinstellungen von Lehrkräften in Handelsschulen und Handelsakademien in Österreich. Sie findet einerseits die Einschätzungen, dass Dialekt allgemein förderlich, ein Reichtum der deutschen Sprache bzw. des Kulturguts und bedeutend für die Identität sei, wenn es allerdings um den Dialekt im schulischen Kontext geht, dann überwiegen die negativen Einstellungen (vgl. Pribyl-Resch 2021: 42 ff).

Einen etwas anderen Weg hat Kaiser (2006) eingeschlagen, wenn sie eine spezielle Form von Spracheinstellung erhebt, die in Österreich relevant ist. Sie befragt Österreicher:innen nach ihrer Meinung zur Verwendung bundesdeutschen Standards in unterschiedlichen sozialen Situationen und Medien in Österreich. Dabei wird der bundesdeutsche Standard auf der Status- und Dynamik-Dimension recht positiv einge-

schätzt, am stärksten von gebildeten Personen. Auf der Attraktivitätsdimension wird das Bundesdeutsche eher negativ eingeschätzt (Kaiser 2006: 241 f).

Auch die Ergebnisse von Lenz (2014) gehen in eine ähnliche Richtung, wenn sie Spracheinstellungen bei Lehrkräften im deutschsprachigen Raum zum Thema Sprachwandel erhebt. Eine der Erkenntnisse ist, dass neben Englisch und zum geringeren Teil auch anderen Sprachen in Österreich und der Schweiz dem Bundesdeutschen eine wichtige Rolle bei Sprachwandelprozessen in diesen Regionen zugeschrieben wird. Dagegen kann der umgekehrte Fall nicht beobachtet werden (Lenz 2014: 350).

Ähnliche Ergebnisse lassen sich beobachten, wenn man sich Sprachnormvorstellungen im deutschsprachigen Raum ansieht. So zeigt Peter (2015) anhand des Textbewertungsverhaltens von Proband:innen, dass der bundesdeutschen Standardvarietät das höchste Prestige zugeschrieben wird (vgl. Peter 2015: 145). Ähnliche Ergebnisse finden auch Fink/Ransmayr/de Cillia (2017). Zwar wird dem „österreichischen Deutsch“ eine bedeutende identitätstiftende Funktion zugeschrieben, andererseits wird österreichisches Standarddeutsch als weniger korrekt empfunden als der bundesdeutsche Standard (vgl. Fink/Ransmayr/de Cillia 2017: 93).

Normerwartungen im schulischen sowie universitären Kontext Österreichs untersuchen Vergeiner et al. (2019). Sie finden heraus, dass sich in beiden Institutionen gesamtgesellschaftliche Erwartungen spiegeln: So spielen die Formalität der Situation als auch die Adressat:innen eine wesentliche Rolle bei der Varietätenwahl, gleichzeitig existieren Unterschiede insofern, dass die Standarderwartung an Universitäten höher ist. Die Einstellungen zu den Sprachvarietäten decken sich wiederum stark mit den schon altbekannten Mustern (vgl. Vergeiner et al. 2019: 33 ff).

Spracheinstellungen zu Dialekt und Standardsprache stellen einen gut erforschten Gegenstand in Österreich dar. Die unterschiedlichen Zuschreibungen, die international zu Dialekten und Standardsprachen gefunden werden, scheinen sich auch in dieser Sprachregion zu bestätigen. Auffällig sind die Befunde von Bellamy (2012), der z. T. unklare Einschätzungen auf der Solidaritätsdimension findet, sowie jene von Fuchs/Elspaß (2019), bei denen der Abstand zwischen den Varietäten auf gleicher Dimension ebenfalls nicht allzu groß ausfällt.

## 4.2 Spracheinstellungen und Gender

Schon in einer der ersten Arbeiten der Variationslinguistik untersuchte Trudgill (1972) den Unterschied im Sprachverhalten von männlichen und weiblichen Proband:innen. Er zeigt, dass Frauen standardsprachliche Varianten häufiger verwenden, als dies Männer tun (vgl. Trudgill 1972: 174). Dieses Ergebnis wurde im englischsprachigen Raum auch mehrfach bestätigt (vgl. Wardhaugh/Fuller 2015: 171 f)<sup>5</sup> und ist so auch im deutschsprachigen Raum zu finden (z. B. Steinegger 1998: 259). Die Datenlage wird jedoch um einiges dünner, wenn man Geschlechterunterschiede bei Spracheinstellungen fokus-

---

5 Einen detaillierten Überblick über variablenbasierte Geschlechterforschung gibt Cheshire (2008).

siert. Brouwer (1989) liefert eine Studie, in der dieser Zusammenhang direkt analysiert wird. Dafür sieht sie sich einerseits die Sprachverwendung von in Amsterdam lebenden Paaren an und andererseits auch die zugehörigen Spracheinstellungen, die sie mittels der Matched-Guise-Technik erhob. Sie stößt ebenfalls auf das Muster, dass Frauen im Vergleich zu Männern vermehrt standardniederländische Varianten verwendeten. Gleichzeitig zeigt sie einen positiven Zusammenhang zwischen Bildungsgrad und Verwendung der Standardsprache. Bei den Spracheinstellungen stellt sich heraus, dass Frauen positivere Einstellungen zu Standardniederländisch und Männer positivere Einstellungen zu Dialekt besitzen (vgl. Brouwer 1989: 103 f). Zusätzlich erkennt sie Folgendes: „The standard variety was associated with both high social status, particularly for men, and femininity, and the non-standard variety with both low social status, particularly for women, and masculinity“ (Brouwer 1989: 104).

Soukup (2009) untersucht Geschlechterunterschiede der Proband:innen in Bezug auf deren Ratingverhalten österreichischer Sprachvarietäten gegenüber: „[F]emale informants were consistently rating the speakers higher/closer to the more desirable adjective pole than the male informants, without a single exception“ (Soukup 2009: 126). Auch Bellamy (2012) analysiert das Ratingverhalten von Männern und Frauen im österreichischen Kontext und erkennt dasselbe Muster wie Soukup (vgl. Bellamy 2012: 190). Zusätzlich vergleicht er die unterschiedlichen Bewertungsmuster von männlichen und weiblichen Guises:

When comparing the responses to the female guises with the responses to the male guises, the differences in the evaluations of the high and low-prestige guises for the male speakers are generally larger in margin for the status traits than for the high and low-prestige guises for the female speakers. Besides this, the female high-prestige guises are perceived as more religious than the female low-prestige guises and for the male guises it is the reverse (vgl. Bellamy 2012: 211).

Vergeiner et al. (2019) untersuchen die Angemessenheitsurteile zur Sprachverwendung an Universitäten und kommen zu folgendem Urteil:

In Bezug auf den Faktor ‚Gender‘ erweisen sich männliche Studierende [...] als vielfach toleranter als weibliche, insofern sie die Standardsprache stärker in informellen und alltags-sprachlichen und Dialekt sowie Umgangssprache eher in formellen Situationen akzeptieren. Die Geschlechterunterschiede sind dabei jedoch insgesamt gering und stellen sich bei einer übergreifenden Betrachtung als weitgehend irrelevant heraus (Vergeiner et al. 2019: 20).

Bei Betrachtung von Geschlechterunterschieden in Bezug zu Spracheinstellungen scheinen sich zwei Muster herauszukristallisieren: So scheinen Frauen Sprachvarietäten generell positiver einzuschätzen, andererseits zeigt Brouwer (1989), dass Frauen nicht nur häufiger Standardvarietäten verwenden, sondern diese auch positiver einschätzen als Männer.

### 4.3 Spracheinstellungen im Migrationskontext

In den vorhergehenden Abschnitten wurden Spracheinstellungen von (autochthonen) Sprecher:innen mit einem Fokus auf Österreich besprochen. Dabei lebt in Österreich auch eine große Gruppe an Personen, die nicht mit Deutsch als Erstsprache aufgewachsen sind. Wie sieht es nun mit Spracheinstellungen bei dieser, zumeist allochthonen Gruppe aus?

Baßler/Spiekermann (2001) befragen hierzu 15 DaZ-Lehrkräfte und 151 DaZ-Lernende. Auch wenn sich beide Stichproben nicht gut vergleichen lassen (speziell die Lehrkräfte-Stichprobe kann wegen der geringen Anzahl und der Profession nicht als repräsentativ eingeschätzt werden), lohnt sich ein Blick auf den Vergleich der beiden Gruppen. Dabei fällt auf, dass die Lernenden Dialekt generell negativer einschätzen als die Lehrkräfte. Einzige Ausnahme betrifft das Adjektiv einsetzbar (gemeint ist die Einsetzbarkeit im Berufsalltag), welches von den Lernenden höher eingeschätzt wird. Deutliche Unterschiede sieht man auf Sympathieebene, auf der die Lehrkräfte den Dialekt deutlich positiver einschätzen (vgl. Baßler/Spiekermann 2001: 6).

Ender/Kasberger/Kaiser (2017) untersuchen Spracheinstellungen von Jugendlichen mit Deutsch als L1 und L2, die sie mit der Matched-Guise-Methode elizitieren. Sie verwenden Guises von einem Verkäufer, einer Verkäuferin sowie eines Arztes und einer Ärztin und ließen diese von den Jugendlichen auf einer siebenteiligen Skala einschätzen. Dabei zeigt sich, dass sowohl im Verkaufs- als auch im medizinischen Kontext die standardsprachlichen Guises von beiden Gruppen sehr ähnlich eingeschätzt wurden. Bei den dialektalen Guises zeigt sich ein anderes Bild: So werden die dialekt-sprechenden Personen von den Jugendlichen mit Deutsch als L2 negativer eingeschätzt als von denjenigen mit Deutsch als L1 (vgl. Ender/Kasberger/Kaiser 2017: 105 f). Schon diese beiden Studien verdeutlichen, dass Personen, die nicht Deutsch als Erstsprache besitzen, dialektale Sprechweise tendenziell negativer einschätzen als Personen mit Deutsch als L1. Dass die unterschiedliche Einschätzung dieser Personengruppen jedoch nicht primär mit der Erstsprache zusammenhängen muss, zeigt Ender (2020). Sie nimmt bei 108 Personen im mittelbairischen Raum sowohl die Spracheinstellungen zu Dialekt und Standard sowie die Verwendung dieser Varietäten im Alltag in den Fokus. Die Ergebnisse zeigen, dass die Prognose der Einstellungen gegenüber Dialekt mit der Variable Dialektkompetenz günstig ist: Je höher die Dialektkompetenz, desto positiver sind die Einschätzungen der dialekt-sprechenden Personen (vgl. Ender 2020: 98).

### 4.4 Spracheinstellungen in städtischen und ländlichen Regionen

Unterscheiden sich Spracheinstellungen zu den Varietäten Dialekt und Standarddeutsch bei Personen, die am Land oder in der Stadt leben? Diese Frage drängt sich aufgrund der unterschiedlichen arealen Verwendungsmuster der Varietäten auf. Da-

bei sind mir jedoch keine Untersuchungen bekannt, die (zumindest im deutschsprachigen Raum) Spracheinstellungen in ihren arealen Unterschieden systematisch untersucht haben.

Was die areale Verwendung von Dialekt und Standardsprache betrifft, berichtet schon Wiesinger (1992: 294): „Groß ist der Stadt-Land-Gegensatz von Dörfern, Märkten und Kleinstädten (bis zu 20.000 Einwohnern) mit 83 % Dialekt und 17 % Umgangssprache und Mittel- und Großstädten mit nur 32 % Dialekt und dafür 63 % Umgangssprache und noch 5 % Standardsprache im durchschnittlichen alltäglichen Sprachgebrauch“.

Steinegger (1998) untersucht den Sprachgebrauch in unterschiedlichen österreichischen Regionen und gelangt zu folgendem Resümee:

Dörfer und Kleinstädte ähneln einander in Sprachgebrauch und Sprachbeurteilung sehr. Sie tendieren generell zu höherem Dialektgebrauch und positiverer Einschätzung der dialektalen Varietät. Die Bewohner aller größeren Orte zeigen eine deutlich höhere Dialekt-distanz. Ihre Kurven liegen v. a. beim Dialektgebrauch und bei der positiven Dialekteinschätzung eng beieinander. Sowohl die Großstädte [...] als auch Wien [...] zeigen allerdings eine Eigendynamik, die es unmöglich macht, alle größeren Städte in einen Topf zu werfen. (vgl. Steinegger 1998: 180)

Wiesinger (2014a) berichtet von Umfragen, die er 1984/85 und 1991 durchführte, dass die Sprachverwendung in Dörfern bei 62 % Dialekt, 35 % Umgangssprache und 3 % Standard; in Kleinstädten bei 55 % Dialekt, 42 % Umgangssprache und 3 % Standard und in Großstädten bei 27 % Dialekt, 65 % Umgangssprache und 8 % Standard liegt (Wiesinger 2014a: 94). Dass im ländlichen Raum signifikant dialektaler gesprochen wird als in urbanen Zentren, findet auch Huesmann (1998: 125 f) in unterschiedlichen Regionen Deutschlands.

Die Verwendung von Dialekt und Standardsprache scheint also in enger Verbindung mit der Urbanität einer Region zu stehen – je ländlicher die Region, desto häufiger wird der Dialekt verwendet. In Österreich fällt zusätzlich auf, dass es neben einem Stadt-Land-Gefälle in der Dialektverwendung ebenfalls ein West-Ost-Gefälle zu geben scheint (vgl. Steinegger 1998: 202), das für diese Untersuchung jedoch keine Relevanz hat.



## 5 Spracheinstellungen und Sprachvariation im schulischen Kontext

In Österreich herrschen neun Jahre Schulpflicht. Auch wenn man nach Ablauf dieser Zeit keines der gut ausgebauten Angebote an allgemeinbildenden oder berufsbildenden weiterführenden Schulen nutzt, verbringt jede Person dieses Landes viel Zeit der Kindheit und der Jugend in diesen Institutionen. Dabei ist eines der Ziele des Unterrichts, Lese-, Schreib- und Sprechkompetenzen in der Erstsprache zu erwerben und zu erweitern. So liest man schon im aktuellen Lehrplan der österreichischen Volksschulen, dass es ein Ziel sein sollte „die Bereitschaft und Fähigkeit der Schüler zur mündlichen Kommunikation allmählich zu erhöhen und sie zu einem möglichst sicheren Gebrauch der Standardsprache zu führen“.<sup>6</sup> Wie dies in der Praxis umgesetzt wird und welche Schwierigkeiten sowie Potenziale sich in einer Sprachregion mit lebendigem Dialekt-Standard-Kontinuum ergeben, wird in den folgenden Abschnitten beleuchtet.

### 5.1 Die Sprachbarrierendiskussion

In Abschnitt 2.2 wurde gezeigt, dass die Entstehung von Spracheinstellungen im deutschsprachigen Raum in enger Verbindung mit der Sprachverwendung des Bildungsbürgertums steht. Da auch Lehrkräfte dieser Gesellschaftsschicht angehören, ist es nicht weiter verwunderlich, dass Spracheinstellungen in die Schulen getragen und dort – direkt oder indirekt – an die Schüler:innen weitergegeben werden. Unter diesen Voraussetzungen veröffentlichte Bernstein (1964) seine Theorie des restringierten und elaborierten Codes. Er geht davon aus, dass sich in sozialen Gruppen, die sich aufgrund eines spezifischen Indexes eng miteinander identifizieren, ein restringierter Code entwickelt (vgl. Bernstein 1964: 61), den er in weiterer Folge jedoch nur auf bildungsferne Schichten projiziert. Dieser eignet sich zwar dafür, sich in dieser sozialen Gruppe zurechtzufinden, beinhaltet jedoch nicht die Möglichkeit des elaborierten Codes „to elaborate verbally and to make explicit his discrete intent“ (Bernstein 1964: 63). Er spezifiziert auch die sozialen Schichten, denen er die Codes zurechnet. So verwenden den elaborierten Code die „middle-class“ und ihr zugeordnete Schichten, der restringierte Code werde stattdessen hauptsächlich von der Arbeiterklasse verwendet (Bernstein 1964: 66). So würden Kinder der unterschiedlichen Schichten auch mit den zugehörigen Codes sozialisiert, was Bernstein zufolge Einfluss auf deren Intelligenzentwicklung zeigt:

---

6 S. [https://www.bmbwf.gv.at/dam/jcr:0a2bae62-d73c-4cd1-b3ca-55b0bc8bdccc/VS7T\\_Deutsch\\_3994.pdf](https://www.bmbwf.gv.at/dam/jcr:0a2bae62-d73c-4cd1-b3ca-55b0bc8bdccc/VS7T_Deutsch_3994.pdf), abgerufen am 29. Mai 2024.

This deterioration in verbal I. Q., discrepancy between verbal and nonverbal I. Q. tests and failure to profit from formal education on the part of working-class children, particularly those of lower working-class origins, is thought to be closely related to the control on types of learning induced by a restricted code (Bernstein 1964: 67).

Diese problematische Analyse führte im deutschsprachigen Raum zur sogenannten Sprachbarrierendiskussion, die die soziolinguistische Forschung, aber auch Empfehlungen für die Deutschdidaktik zum Umgang mit Sprachvariation im Unterricht maßgeblich beeinflusste. Bevor wir uns diese Diskussion näher ansehen, sei noch ergänzt, dass die „Defizithypothese“, wie sie auch genannt wird, im englischsprachigen Raum schnell auf Widerstand stieß. Die bekannte Replik auf Bernstein kam dabei von Labov (1969), indem er dialektale Sprachvarietäten nicht als defizitär darstellte, sondern deren Differenz auf den unterschiedlichen linguistischen Ebenen herausarbeitete und sie als ähnlich funktionale Systeme beschrieb. Labov war zehn Jahre später auch in einem Rechtsprozess als Gutachter involviert, der in den USA Aufsehen erregte: In der als Ann-Arbor-Decision bekannt gewordenen Entscheidung urteilte der Richter, dass die Martin-Luther King Jr. Elementary School in Ann Arbor, Michigan, afroamerikanischen Kindern nur unzureichende Ausbildungsmöglichkeiten zur Verfügung stellte, da die Sprachunterschiede, die diese Bevölkerungsgruppe im Vergleich zur weißen Bevölkerung aufweist, nicht miteinbezogen wurden. Dies führte in den USA zu einer bis heute fortdauernden Forschungstradition, die sich mit Sprachvariation im Englischunterricht auseinandersetzt (s. z. B. Fasold 2001; Mufwene 2001; Murray 2001; Charity/Scarborough/Griffin 2004; Craig/Kolenic/Hensel 2014; Smith et al. 2019), dazu aber mehr in Abschnitt 6.2.

Auch im deutschsprachigen Kontext schlug die Sprachbarrierendiskussion hohe Wellen. So griffen z. B. Ammon (1972) und Löffler (1972) die Defizithypothese auf und übertrugen das Konzept des restringierten Codes auf die Dialekte im deutschsprachigen Raum. Ammon (1972) zeigt dies an Normabweichungen in Schulaufsätzen und findet bei Kindern der Unterschicht eine höhere Anzahl von Normabweichungen und dialektinduzierten Fehlern (Ammon 1972: 236). Neuere Forschungen scheinen dies jedoch nicht zu bestätigen. So zeigt z. B. Noack (2003: 19), dass sich Rechtschreibfehler in der 5. und 7. Schulstufe vor allem auf die Unterschiede zwischen gesprochener und geschriebener Sprache zurückführen lassen und dass dialektal bedingte Fehler generell recht selten vorkommen.

Im deutschsprachigen Raum regte sich schon bald Widerstand dagegen, Dialekte pauschal als defizitär darzustellen. So diagnostiziert Besch (1975), dass Dialekte nur dann eine Barriere bei der Erlernung der Standardsprache darstellen, wenn sie „durch soziale Diskriminierung in der Schultube und außerhalb und durch Ignorierung der primärsprachlichen Interferenzen zum Nachteil der dialektsprechenden Kinder [werden]“ (Besch 1975: 163).

Im Zusammenhang mit der Sprachbarrierendiskussion ist auch der Vortrag von Wandruszka (1975) zu sehen, der dem Konzept der Sprachbarriere dasjenige der inneren Mehrsprachigkeit gegenüberstellt und den Auftrag der schulischen Sprachbildung folgendermaßen sieht:

Aufgabe der Schule ist es daher, die Ohren und Augen der Kinder zu öffnen für die Verschiedenheit der Sprachen, die uns in unserer Sprache zur Verfügung stehen, ihre Mehrsprachigkeit in der Erstsprache zu entfalten und zu erweitern, Schranken abzubauen und dabei nicht neue zu errichten durch die Überbetonung einer in allen Lebenslagen allein gültigen Hochsprache, einer einzig richtigen monosystematischen Norm (Wandruszka 1975: 322).

Auf diesem Denkmuster aufbauend zeigt Bausinger (1983), wie man das Wissen um Dialekte in den Deutschunterricht integrieren könnte. Auch Bücherl (1993) sieht „Dialekt als Chance“. Er stellt dabei den Vorteil von bivarietären Sprecher:innen heraus, die, sowohl auf gesprochen- als auch geschriebensprachlicher Ebene, einen Vorteil darin besitzen, angemessener auf die Kommunikationssituationen reagieren zu können (vgl. Bücherl 1993: 74 f). Glantschnig (2011) zeigt unterschiedliche grammatische Bereiche im bairischen Dialektraum, die man für kontrastiven Grammatikunterricht heranziehen könnte. Damit würde man ihrer Meinung nach auch den Grammatikunterricht generell aufwerten (vgl. Glantschnig 2011: 240).

Die Sprachbarrierendiskussion ließ dann in den 1980er-Jahren nach (s. Ammon/Kellermeier 1997; Kellermeier-Rehbein 2017), was jedoch nicht daran lag, dass Dialekte und Standardsprachen fortan in der Wahrnehmung der Bevölkerung und der schulischen Institutionen als gleichberechtigte Varietäten galten. Die verbesserte Situation und das dadurch abnehmende Interesse für Dialekt im Unterricht sieht Berthele (2016: 9) darin begründet, dass der maximal stigmatisierte Platz, den früher dialektsprechende Kinder innehatten, heute von Kindern aus Einwandererfamilien ausgefüllt wird, die damals noch keine wesentliche Rolle in der Wahrnehmung spielten. Infolge verschob sich auch das Forschungsinteresse im Bereich Sprachvariation im Unterricht hin zu dieser Gruppe (s. z. B. Wiese 2006; Wiese 2015; Androutsopoulos 2019).

Ob sich die Diskussion im Bereich der innersprachlichen Variation damit erübrigt hat, lässt sich schwer feststellen. Es ist zumindest im Forschungskontext nur mehr selten von Sprachbarrieren zu lesen, wie man es bei Kellermeier-Rehbein (2017) findet. Sie zieht Parallelen zum derzeitigen Diskurs, der sich um das Thema Bildungssprache dreht (vgl. Kellermeier-Rehbein 2017: 22) und diagnostiziert neue „interlinguale“ Sprachbarrieren im Migrationskontext oder indem sie den Wert der englischen Sprache für den international ausgerichteten Arbeitsmarkt hervorhebt und der deutschen Sprache hier Bedeutungslosigkeit zuschreibt (vgl. Kellermeier-Rehbein 2017: 24 f). Ob in letzterer Sprachbarrierensituation-Neu dem Deutschen nun die Funktion des restringierten Codes zukommt, bleibt jedoch unbeantwortet.

## 5.2 Spracheinstellungen bei Lehrkräften in Österreich

Auch wenn man in (sozio)linguistischen Kreisen heutzutage nicht mehr von einem Defizit sprechen würde, das mit dem Erwerb und der Verwendung einer dialektalen Varietät einhergeht, heißt das nicht, dass sich diese Erkenntnis auch in der Allgemeinbevölkerung durchgesetzt hat – oder bei allen Germanist:innen. Sehen wir uns dies

nun anhand von Spracheinstellungen bei Lehrkräften an. Rastner (1997) führt eine qualitative Umfrage unter Lehrkräften zu den Varietäten Dialekt – Umgangssprache – Standardsprache in Österreich durch. Sie zeigt, dass in Ostösterreich einer „gepflegten Umgangssprache“ bzw. der Standardsprache höchstes Prestige zugesprochen wird, im Westen und Süden Österreichs die Lage jedoch anders eingeschätzt wird: Hier lautet die Empfehlung in offiziellen Situationen, die Umgangssprache zu verwenden, in inoffiziellen jedoch den Dialekt zwecks Ausdruck regionaler Identität einzusetzen (vgl. Rastner 1997: 92).

Auf Ebene der Standardvariation untersucht Fink (2014: 92) das Korrekturverhalten von Lehrkräften, indem sie ihnen einen fingierten Schüleraufsatz, der sowohl Deutschlandismen als auch Austriazismen enthält, vorlegt und evaluieren lässt. Dabei werden Deutschlandismen zwar häufiger unterstellt, Austriazismen jedoch fast dreimal so häufig korrigiert.

Eine mögliche Erklärung dafür wäre, dass Deutschlandismen für österreichische Lehrende weniger gewohnt sind und daher eher unterstellt werden; zugleich werden sie aber als korrekt empfunden und somit seltener korrigiert; auch seltener als Austriazismen (Fink 2014: 6).

Lenz (2014) erhebt Spracheinstellungen von Lehrkräften in Deutschland, Österreich und der Schweiz zum Thema Sprachwandel des Deutschen. Auch wenn sie keine Fragen zum Umgang mit Sprachvariation im Unterricht stellt, sind zwei Ergebnisse der Studie erwähnenswert: Der Vergleich der Stichprobe mit einer Studie der Universität Mannheim zeigt, dass Lehrkräfte sehr ähnliche Einschätzungen wie die Gesamtbevölkerung zeigen (vgl. Lenz 2014: 331). Im Licht der oben zitierten Studien ist es daher wahrscheinlich, dass es sich bei Einstellungen zu Sprachvarietäten ähnlich verhält. Zusätzlich ist auffällig, dass sich die Lehrkräfte selbst nur eine geringe sprachbeeinflussende Rolle zuschreiben (vgl. Lenz 2014: 349). Dabei haben Studien von Kasberger/Kaiser (2019) und Kaiser/Kasberger (2021) gezeigt, dass Spracheinstellungen vor allem in der Primarstufe erworben werden. Dass Lehrkräfte hier einen Einfluss besitzen, liegt also auf der Hand. Pribyl-Resch (2021) untersucht Spracheinstellungen von österreichischen Lehrkräften anhand qualitativer Interviews. Sie legt dar, dass der Großteil der Proband:innen Dialekt als allgemein förderlich in Bezug auf Identität, Verwurzelung, Zusammengehörigkeit sowie die Verständigung und Integration eingeschätzt wird. Nur selten wird er als negativ für den beruflichen Erfolg bzw. unadäquates Mittel der Kommunikation dargestellt (vgl. Pribyl-Resch 2021: 42–47). Anders verhält es sich bei der Einschätzung von Dialekt im schulischen Kontext. Sehr selten wird Dialekt als förderlich für den schulischen Erfolg und auch als Mittel der Integration für DaZ-Schüler:innen beschrieben. Die Berichte über negative Auswirkungen überwiegen deutlich. Dabei werden vor allem die vermeintlichen Probleme von Dialektsprecher:innen beim Erwerb der Schriftsprache angeführt sowie Dialekt mitverantwortlich gemacht für die schlechten Rechtschreibleistungen von Schüler:innen mit einer anderen Erstsprache als Deutsch (vgl. Pribyl-Resch 2021: 53 ff).

Es erweist sich, dass Dialekte im österreichischen Schulunterricht eine ambivalente Stellung genießen. Zum einen tritt wiederum das österreichtypische West-Ost-Gefälle hervor (s. Abschnitt 4.4), zum anderen wird auch bei den Lehrkräften ersichtlich, dass sie Dialekt förderlich auf Solidaritäts- bzw. Sympathieebene sehen, für den Erwerb von Kompetenzen (z. B. schriftsprachlicher Natur) aber eher negativ.

### 5.3 Die Rolle der Schule bei der Weitertradierung stereotyper Spracheinstellungen

Im vorhergehenden Kapitel hat sich gezeigt, dass Lehrkräfte ähnliche Spracheinstellungen besitzen, wie sie auch in der Allgemeinbevölkerung vorkommen, sowie spezifische, was den Schriftspracherwerb betrifft. Alleine schon dadurch liegt es nahe, dass Lehrkräfte als Normautoritäten (Davies 2007) ihre Spracheinstellungen an die Schüler:innen weitergeben. Man findet aber auch ganz konkrete Beispiele in Unterrichtsmaterialien, in denen sie – zum Teil in recht stereotyper Form – vermittelt werden. Dies ist häufig dann der Fall, wenn Sprachvariation zum Thema gemacht wird.

So kommen Bekes/Neuland (2006: 510) zum Schluss, dass „Dialekt oft als Quelle kommunikativer Missverständnisse oder komischer Effekte“ präsentiert wird. In dem Beispiel, das sie präsentieren, werden die Dialektsprecher zusätzlich als rüpelhaft und derb im Vergleich zu den Hochdeutschsprechenden dargestellt. Auch Maitz (2015: 224 ff) widmet sich diesem Thema und zeigt, wie sehr die Darstellung von Sprachvariation von sprachlichen Ideologien geprägt ist. Das Dialektbeispiel, das er angibt, ist einerseits aus linguistischer Sicht fragwürdig, da hier dialektale, umgangssprachliche und gesprochensprachliche Phänomene vermischt werden und andererseits diese meiner Ansicht nach sehr gut allgemein verständlichen Beispiele davon leben, dass sie offensichtlich nicht vom Gegenüber verstanden werden. Gleichzeitig werden auch Beispiele der Jugendsprache und Kiezdeutsch besprochen, die vor allem Stereotype bedienen und nicht der Alltagsverwendung dieser Varietäten entsprechen. Kaiser/Ender (2020: 13) untersuchen die gängigsten Schulbücher, die in Österreich zur Anwendung kommen, und stellen fest, dass das Thema Sprachvariation nur in sehr geringem Umfang berücksichtigt wird. Wenn es vorkommt, dann werden Dialekte häufig als dysfunktionale Sprachsysteme dargestellt. Sie finden jedoch auch Lehrwerke, in denen der soziolinguistische Kenntnisstand berücksichtigt ist.

## 5.4 Dialekte im Deutschunterricht: Potenzial oder vergebene Liebesmühe?

In den letzten Abschnitten wurde viel über Dialekt als Sprachbarriere, über potenzielle Probleme von Dialektsprechenden in der Schule und stereotype Spracheinstellungen gesprochen. Dabei gibt es genügend Erkenntnisse, die Gegenargumente gegen diese weitverbreiteten Meinungen liefern.

### 5.4.1 Dialekt als Barriere für Bildungserfolg?

Die Sprachbarrierendiskussion wurde von Beginn an von Bernsteins (1964) Theorie des restringierten und elaborierten Codes getrieben, die im deutschsprachigen Raum sogleich auf die Dialekte und Standardsprachen angewendet wurde. Wenn man diese Verflechtung zwischen Sprachvarietät, sozialer Schicht und kognitiven Fähigkeiten (Intelligenz) auflöst und neutral die Frage stellt, ob Dialekte potenzielle Sprachbarrieren im Sinne von Verständnisbarrieren darstellen, dann ist die Antwort schlicht ja. Da im deutschsprachigen Raum eine relativ große Dialektvielfalt herrscht (bzw. herrschte) und die Unterschiede zwischen den Dialekten alle linguistischen Beschreibungsebenen betreffen können (lautliche, grammatische und lexikalische Unterschiede), kann von einer allgemeinen Interkomprehensionsmöglichkeit nicht ausgegangen werden. Genauso wenig kann von einer Interkomprehension der (noch lebendigen) westgermanischen Sprachen Deutsch, Niederländisch und Englisch auf Ebene der Standardsprachen ausgegangen werden. Der Argumentation von Kellermeier-Rehbein (2017: s. Abschnitt 5.1) kann also insofern zugestimmt werden, dass auch zwischen dem Deutschen und dem Englischen eine Sprachbarriere vorliegt. Interessanterweise ist nun jedoch die Konsequenz aus didaktischen Kreisen bezüglich dieser beiden sehr ähnlichen Situationen eine völlig andere. So ist das Ziel der europäischen Bildungspolitik, Sprachkompetenz in mindestens drei Sprachen zu fördern (Europäische Kommission 1995, zit. n. Berthele 2010: 45). Auch in Österreich wird schon am Ende der Primarstufe damit begonnen, Englischkenntnisse zu vermitteln, mutmaßlich, um die von Kellermeier-Rehbein angesprochene Sprachbarriere zum international ausgerichteten Arbeitsmarkt hin abzubauen. Dabei könnte man auch die blasphemische Frage stellen, wie groß der Prozentsatz der Personen ist, denen die im Englischunterricht erworbenen Kenntnisse in der späteren Arbeitswelt wirklich nützen.

Im Gegensatz dazu nimmt im Deutschunterricht das Thema Sprachvariation und innere Mehrsprachigkeit zumeist eine Randstellung ein und, wenn es einmal zur Sprache kommt, werden häufig unreflektierte Stereotype wiedergegeben (s. Abschnitt 5.3). Dabei ist in der Realität Dialekt (bzw. Umgangssprache) in Österreich für die meisten Bewohner:innen die Erstsprache und die im Alltag zumeist verwendete Varietät (s. Abschnitt 2.1). Brisant wird diese Ausgangssituation im Kontext der Migration. Da der Fokus des Deutschunterrichts primär auf dem Erwerb der Standardsprache liegt (s. Abschnitt 5.2) und Migrantenkindern somit der gesteuerte Erwerb der im Alltag zumeist gesprochenen Varietäten unzugänglich bleibt, ist es nicht auszuschließen, dass dadurch neue „Sprachbarrieren“ entstehen. Da Dialektkommunikation in Österreich als

Sprache der Nähe und Standardsprache als Sprache der Distanz gilt (vgl. Koch/Oesterreicher 1985: s. auch die klassischen Spracheinstellungen zu Dialekt und Standardsprache, Abschnitt 4.1), kann der Nichterwerb des Dialektes eine Integrationshürde sowie eine Hürde am Arbeitsplatz in den westlichen Bundesländern und den ländlich geprägten Regionen Österreichs darstellen.

Dies stellt jedoch nur einen problematischen Aspekt des Umgangs mit Variation bzw. Variationserwerb im Deutschunterricht für eine bestimmte Bevölkerungsgruppe dar. Die unsachgemäße Behandlung des Dialekts im Unterricht kann auch schädliche Auswirkungen für den Schulerfolg von Dialektsprecher:innen mit sich bringen. So berichtet Yiakoumetti (2012: 300), dass das Verbannen von Dialekt im Unterrichtskontext in einer Region mit lebendiger Dialektkultur (in ihrem Fall handelt es sich um zypriotisches Griechisch vs. Standardgriechisch) dazu führt, dass Schüler:innen mehr Dialektinterferenzen in gesprochener und geschriebener Sprache zeigen, negative Einstellungen gegenüber ihren Dialekten annehmen und im Fremdsprachenunterricht schlechtere Leistungen zeigen.

Wichtige Gründe hierfür liegen darin, dass Kinder, die eine dialektale Varietät sprechen, Angst davor haben müssen, dass sie vor der Klasse korrigiert oder herabgesetzt werden. Das führt bei manchen Kindern dazu, dass sie Hemmungen entwickeln, sich im Unterricht frei auszudrücken, was die kognitive Entwicklung negativ beeinflusst (Siegel 2006: 160). Im schlimmsten Fall kann es dazu kommen, dass die Schüler:innen „verstummen“, eine Konsequenz, die schon lange bekannt ist. Zehetner (1980: 7) beschreibt dies z. B. folgendermaßen:

Die extremste Art der Dialektvermeidung in der Schule ist es, sich jeder verbalen Äußerung zu enthalten, um sich vor Blamage mit dem Dialekt zu bewahren. Solches Verhalten wird dann möglicherweise von Seiten des Lehrers als mangelnde Sachkenntnis und ungenügende Mitarbeit gewertet. Und der Weg vom Verstummen zum schulischen Versagen ist nicht weit.

Eine logische Konsequenz dieses Umgangs mit Sprachvariation kann sein, dass Schüler:innen eine Abwehrhaltung gegenüber der Standardsprache entwickeln (Kaiser/ENDER 2020: 14). Diese Sprachbarriere, die hier konstruiert wird, geht in diesem Fall klar von Vertreter:innen einer verfehlten Sprachdidaktik aus.

Dabei gibt es genügend Hinweise darauf, wie günstig es ist, die regionalen Dialekte im Erstsprachenunterricht zu integrieren (s. für einen Überblick Tegegne 2015). Wheeler/Swords (2004) zeigen eine Methode der kontrastiven Analyse in den USA, in der Elemente von formeller und informeller Sprache (informelle Sprache bezieht sich die Varietät der schwarzen Bevölkerung in Vergleich mit standardamerikanischem Englisch) ohne Besser- oder Schlechterbewertung einer Varietät kontrastiert wurden. Die Ergebnisse dieses Unterrichts bestehen darin, dass die afroamerikanischen Kinder nach einem Jahr gleiche Schulnoten aufwiesen wie die weißen und diese in Mathematik sogar übertrafen.

Auch Yiakoumetti (2007) untersucht in einer Studie, inwiefern kontrastiv analytischer Sprachunterricht die Schreibfertigkeiten von elf- bis zwölfjährigen Schüler:in-

nen beeinflussen kann. Sie stellt eine generelle Verbesserung der Aufsätze fest, speziell bei Kindern aus ländlichen Regionen, sowie einen Rückgang dialektaler Interferenzen (Yiakoumetti 2007: 58 ff).

Neben den Ansätzen der kontrastiven Analyse gibt es eine ganze Reihe an Ansätzen, die „dialektale Bewusstheit“, „kritische Bewusstheit“ oder generell „Sprachbewusstheit“ in den Mittelpunkt der Erstsprachdidaktik stellen. Da diese Methoden (allen voran der Critical Awareness Approach von Siegel 2006) auch in der Konzeption der in dieser Arbeit verwendeten Unterrichtsmaterialien als Vorbild galten, werden sie separat in Abschnitt 6.2 behandelt.

Vom Dialekt selbst geht in der Erstsprachendidaktik also keine Sprachbarriere aus, vielmehr stammt die Barriere von einer Sprachdidaktik, die Dialekte als defizitäre Sprachsysteme bewertet, dialektsprechende Schüler:innen daher korrigiert und verunglimpft, bis sich diese vom Schulsystem abwenden. Dabei wäre es so einfach, mit der richtigen Methodenwahl (und dem entsprechenden Mindset) die Leistungen dieser Bevölkerungsgruppe zu fördern oder zumindest Chancengleichheit zu gewährleisten.

#### **5.4.2 Vor- und Nachteile innerer Mehrsprachigkeit im Kontext der Bildungslandschaft**

Der Forschungsdiskurs hat sich jedoch nicht nur mit didaktischen Fragestellungen zum Umgang mit Sprachvariation im Erstsprachenunterricht auseinandergesetzt. Auch die kognitionslinguistische Forschung hat sich des Themas der Mehrsprachigkeit angenommen. Traditionell wurde zwar die äußere Mehrsprachigkeit in den Blick genommen, man findet jedoch auch vereinzelt Forschungsergebnisse zur inneren Mehrsprachigkeit. So berichtet Bialystok (2001), dass Multilinguale im Vergleich mit monolingualen Kindern eine höhere Fähigkeit der Aufmerksamkeitskontrolle aufweisen. Diese Fähigkeit, die auch als exekutive Kontrolle bezeichnet wird, konnte in bidialektalen Personen nachgewiesen werden. So belegen dies Garbin et al. (2010) in einer Non-verbal-task-switching-Studie, in der die Aktivierung der Hirnregionen bei monolingual Spanischsprechenden mit bidialektal Spanisch-Katalan-Sprechenden verglichen wurden. Dabei stellte sich heraus, dass einerseits das Switching den bidialektalen Personen geringere kognitive Anstrengung bereitet, andererseits sie dabei Hirnregionen verwenden, die bei Monolingualen nicht aktiviert wurden.

Papapavlou/Phili (2009) untersuchen metasprachliche Fähigkeiten von Kindern, die entweder monolingual mit Standardgriechisch, bidialektal mit Standardgriechisch und zypriotischem Griechisch oder bilingual mit Griechisch und einer weiteren Sprache aufwachsen. Dabei stellen sie eine höhere phonologische Bewusstheit von bidialektalen Kindern im Vergleich mit den monolingualen fest – die Ergebnisse der bidialektalen und der bilingualen Kinder unterscheiden sich jedoch kaum voneinander.

Antoniou et al. (2016) fokussieren ebenfalls Kinder im zypriotisch-griechischen Setting und finden erhöhte Fähigkeit zur exekutiven Kontrolle bei den bidialektalen und den bilingualen Kindern im Vergleich zu den monolingualen, wobei letztere über ein etwas höheres Sprachniveau verfügen. Zwischen Ersteren bestanden wiederum nur geringe Unterschiede in den Ergebnissen.

Zu etwas abweichenden Ergebnissen kommen Poarch/Vanhove/Berthele (2019). Sie untersuchen exekutive Kontrolle von Erwachsenen anhand des Simon- und Flanker-Tasks im schwäbischen Sprachraum. Dabei schneiden die Dialektsprechenden besser ab als die bidialektal Schwäbisch-Standarddeutsche Gruppe. Das abweichende Ergebnis – wie die Autoren selbst anmerken – könnte jedoch mit dem Umstand zu tun haben, da in dieser Region außerhalb der eigenen vier Wände und in Medien die standarddeutsche Sprache ubiquitär vorhanden ist. Daher könnte es in diesem Sprachraum der Fall sein, dass die „monolektalen“ Personen häufiger zwischen den Varietäten wechseln müssen als die bidialektalen, sie also de facto gar nicht monolektal sind (vgl. Poarch/Vanhove/Berthele 2019: 625).

Daneben scheint Bidialektalität auch andere Vorteile aufzuweisen. Vangsnes/Söderlund/Bleckesaune (2017) untersuchen im norwegischen Setting die Schulleistungen von Schüler:innen, die entweder in Nynorsk oder Bokmål literalisiert sind. Bei diesen zwei Varietäten handelt es sich um die beiden schriftsprachlichen Normen, die in Norwegen verwendet werden. Nynorsk stellt dabei die Minderheitenvarietät dar (ca. 12,5 % Verwendung im schulischen Bereich) – Kinder, die mit dieser Varietät aufwachsen, lernen simultan Bokmål, da sie dieser Varietät im Alltag ständig ausgesetzt sind. Die Ergebnisse zeigen nun, dass die generellen Schulnoten in den Nynorsk-Zentren über denjenigen in den Bokmål-Zentren liegen, wenn sozioökonomische Faktoren kontrolliert werden. Wenn man diese Tendenz in Norwegen auf den deutschsprachigen Raum umlegt, dann zeigt sich, dass die Länder in Deutschland, in denen es noch lebendige Dialekte gibt (Baden-Württemberg, Bayern, Sachsen), auch bei den PISA-Tests besser abschneiden (vgl. Bayerisches Staatsministerium für Bildung und Kultus 2015: 5).

Bialystok/Craik/Freedman (2007) zeigen, dass bilinguale Personen im Durchschnitt vier Jahre später an Demenz erkranken als monolinguale. Die Bidialektalität scheint ebenfalls einen positiven Effekt auf die Erschließungskompetenz zu haben. So untersuchte Berthele (2008) die Fähigkeit, niederländische Texte zu erschließen bei Germanistik-Studierenden mit bzw. ohne Dialektkenntnis und kommt zu folgendem Ergebnis:

Dialektnahe Personen erschließen tendenziell etwas mehr unbekannte Wörter, sie tun dies auch etwas spontaner, signifikant korrekter und sie verstehen einen Lesetext in der unbekannten Fremdsprache Niederländisch signifikant besser als dialektferne Sprecher:innen (vgl. Berthele 2008: 100).

Daneben existiert auch der offensichtliche Vorteil, dass bidialektale Sprecher:innen besser in der Lage sind, sich an sprachliche Situationen anzupassen. Sie können stärker auf einen angemessenen Einsatz der Varietät in einem Umfeld reagieren, indem die Varietätenpole Dialekt und Standarddeutsch für nächsprachliche und distanzsprachliche Kommunikation eingesetzt werden. Dies eröffnet maximale Flexibilität im soziolinguistischen Umgang mit anderen Menschen und erhöht damit die Chancen auf kommunikativen Erfolg.

Man findet auch manche Eigenschaften, die mit (innerer) Mehrsprachigkeit einhergehen, die man als Nachteile auslegen könnte. Zum einen zeigen Studien, dass sich bei bidialektalen Sprecher:innen im Vergleich zu denjenigen, die nur mit der Standardvarietät aufwachsen, dialektale Interferenzen nachweisen lassen. Bei kontrastivem Unterricht werden diese, wie wir oben gesehen haben, zwar stark reduziert, sie können jedoch auch bei hohem Sprachniveau beider Varietäten vorkommen (vgl. Yiakoumetti 2012: 298).

Gleichzeitig weisen bilinguale Kinder einen kleineren Wortschatz pro Einzelsprache auf als monolinguale (Klassert/Kauschke/Gagarina 2009), es ist also durchaus denkbar, dass dies bei bidialektalen Kindern auch der Fall ist – überprüft wurde dies jedoch meines Wissens noch nicht. Dabei ist es wichtig anzumerken, dass der Mythos der „doppelten Halbsprachigkeit“ nicht haltbar ist. Bei bilingualen Kindern legt sich das Gehirn zwar erst später auf die Sprachmuster der Erstsprachen fest, allerdings entsteht dadurch keine Verwirrung, eher im Gegenteil: „Das Organisieren bilingualer Lernkontexte befördert diese effiziente kognitive Verarbeitungskompetenz“ (Böttger 2016: 179). Zum anderen ist es teilweise der Fall, dass dialekt sprechende Personen schwerer eine Anstellung bekommen. Maitz/Elspaß (2011) zeigen für Deutschland, dass dialekt sprechende Personen, wenn es um Schule, Studium und Bewerbungen geht, diskriminiert werden. Grogger/Steinmayr/J. Winter (2020) untersuchen ebenfalls in Deutschland die Einkommensverhältnisse von Personen mit dialektaler Färbung und standardsprechenden Personen und finden einen ähnlichen Gehaltsunterschied dieser beiden, wie man dies beim Gender-Pay-Gap zwischen Frauen und Männern vorfindet.

Bidialektalität scheint ähnlich wie Bilingualität mehrere Vorteile und manche Nachteile zu besitzen. Auch wenn die Forschungslage ausführlicher sein könnte, scheinen bidialektale Sprecher:innen Vorteile in der Aufmerksamkeitssteuerung/exekutiven Funktionen zu besitzen, sie schneiden eventuell besser in der Schule ab (wenn sie nicht aufgrund ihrer Sprechweise diskriminiert werden), haben eventuell den medizinischen Vorteil, dass eine etwaige Demenzerkrankung sich erst in späterem Lebensalter manifestiert (wobei es in diesem Fall nur Untersuchungen aus der Bilingualitätsforschung z. B. Bialystok/Craik/Freedman 2007 gibt) und besitzen eventuell den Vorteil, sich eine ähnliche Fremdsprache besser erschließen zu können. Daneben wird man bei bidialektalen Personen Interferenzen vorfinden, wenn durch fehlenden kontrastiven Grammatikunterricht die potenziellen Interferenzbereiche nicht bewusst gemacht wurden. Ob man gelegentliche Interferenzen allerdings als Nachteil einschätzt, ist aus meiner Perspektive nur aus einer stark standardsprachlichen Ideologieposition heraus vertretbar. Ähnlich verhält es sich mit dem Nachteil am Arbeitsmarkt, der sich ebenfalls nur aus ideologischen und nicht aus fachlichen Argumenten speist. Zuletzt existiert auch ein Problem identitärer Natur. Wie wir in Abschnitt 3.2 gesehen haben, werden Personen, die sich der prestigeträchtigen Varietät zuwenden, von Personen ihrer eigenen Ingroup potenziell als kulturelle Verräter gebrandmarkt. Dieses Problem existiert jedoch auch nur dann, wenn die Herkunftsvarietät zugunsten der Prestigevarietät abgelegt wird. Im

Falle ausgebauter innerer Mehrsprachigkeit kann man diesem Problem durch die dem Kontext angemessene Varietätenverwendung entgehen.

Zwar wären noch mehr Forschungsergebnisse in diesem Bereich – vor allem auch aus dem deutschsprachigen Raum – wünschenswert, alles in allem überwiegen meiner Meinung nach aber die Vorteile deutlich und sprechen dafür, zumindest in Regionen mit lebendiger Dialektlandschaft die Alltagsvarietäten im Erstsprachenunterricht als festen Bestandteil zu verankern. Eine mögliche Umsetzung findet man in Kapitel 7.



## 6 Veränderung von Spracheinstellungen

Ein zentrales Ziel dieser Arbeit besteht darin, zu untersuchen, ob sich Spracheinstellungen durch Bewusstmachung und Reflexion im Deutschunterricht verändern lassen. Die Einstellungsveränderung stellt in Kombination mit allgemeinen Einstellungen ein zentrales Forschungsfeld in der Sozialpsychologie dar. Unter dem Begriff Persuasion weist die Sozialpsychologie (und speziell die Werbepsychologie) eine lange Forschungstradition auf. Die Untersuchungen drehen sich dabei jedoch nicht ausschließlich um die Frage, wie man den Verkauf eines Produkts durch gezielte Werbemaßnahmen fördern kann. Auch Fragestellungen, wie man beispielsweise einen Teil der Bevölkerung durch Werbemaßnahmen von einer gesünderen Lebensweise überzeugen kann, wurden untersucht. Daher werde ich im ersten Abschnitt dieses Kapitels die Funktionsweise der Persuasion aus sozialpsychologischer Sicht beschreiben.

Der nächste Abschnitt widmet sich einer Richtung der Sprachdidaktik, den Critical-Awareness-Ansätzen. Nach dem Ann-Arbor-Urteil stand die Didaktik in den USA unter Zugzwang, somit gab es einen Bedarf, Unterrichtsmodelle speziell in Hinblick auf den Umgang mit Varietäten im Erstsprachenunterricht zu entwickeln und diese auch in der Praxis zu implementieren. Diese Ansätze stellten die Grundlage für die Entwicklung der Unterrichtsmaterialien dieser Studie (s. Kapitel 7) dar.

### 6.1 Persuasion

Zur Erklärung von Einstellungsveränderung/Persuasion haben sich in der Sozialpsychologie zwei Modelle durchgesetzt: das Modell der Elaborationswahrscheinlichkeit (ELM) und das heuristisch-systematische Modell. Bei beiden Modellen handelt es sich um Zwei-Prozess-Modelle. Da sie sich in wesentlichen Bereichen entsprechen, werde ich nur auf das erste Modell eingehen. Die folgenden Ausführungen stammen allesamt aus Stroebe (2014: 237 ff).

Die Zwei-Prozess-Theorien gehen von zwei unterschiedlichen Möglichkeiten der Informationsverarbeitung aus: einer systematischen und einer nicht systematischen. Wie die Information verarbeitet wird, hängt einerseits von der Verarbeitungsmotivation ab, da Informationsverarbeitung mit Aufwand und Zeit verbunden ist. Zugleich spielt auch die Verarbeitungsfähigkeit eine Rolle, da Menschen ja nicht über gleiches Wissen verfügen und im Moment eventuell keine Zeit für die Verarbeitung zur Verfügung steht. Sind nun Verarbeitungsmotivation und -fähigkeit hoch, so verläuft die Informationsverarbeitung entlang der zentralen Route der Persuasion. Sind sie es nicht, so verläuft sie über die periphere Route der Persuasion. Diese beiden Routen unterscheiden sich nun stark darin, welche persuasiven Effekte greifen.

Bei der zentralen Route bzw. der systematischen Verarbeitung zählen Argumente. Starke Argumente lösen überwiegend positive/zustimmende Gedanken aus und können so zu einer Einstellungsveränderung führen. Dagegen lösen schwache Argumente vor allem negative/ablehnende Gedanken aus, was keine Einstellungsveränderung auslöst.

Bei der peripheren Route zählen der Inhalt der Botschaft und die Qualität der Argumente weniger, da die Motivation, Fähigkeit oder Zeit fehlt, sich den Argumenten zu widmen. Daher spielen Hinweisreize hier eine wichtige Rolle. Bei der heuristischen Verarbeitung z. B. kann die Glaubwürdigkeit einer Person einen positiven Hinweisreiz darstellen, der die heuristische Assoziation „auf Experten kann man sich verlassen“ nach sich zieht und so eine Einstellungsveränderung herbeiführt. Das Gegenbeispiel, dass geringe Glaubwürdigkeit eine negative Assoziation auslöst und dies keine Einstellungsveränderung bewirkt, ist ebenso möglich. Daneben können bei der Einstellungsveränderung über die periphere Route Informationen über die Einstellungen von Bezugspersonen eine Rolle spielen sowie die Häufigkeit der Darbietung eines Reizes und die (daraus resultierende) evaluative Konditionierung.

Wie aber entscheidet eine Person, welcher Route der Persuasion sie folgt? Dazu findet man ebenfalls mehrere Faktoren. Neben situativen Komponenten wie Zeitdruck sind dies die persönliche Relevanz der Kommunikation, das Kognitionsbedürfnis sowie das Bedürfnis nach einem abschließenden Urteil. Ist der Inhalt einer Kommunikation relevant für die rezipierende Person, wird sie eher die zentrale Route der Verarbeitung einschlagen. Diese Route wird auch eher von Personen eingeschlagen, die ein hohes Kognitionsbedürfnis<sup>7</sup> aufweisen. Bei Personen, die ein hohes Bedürfnis nach einem abschließenden Urteil aufweisen, ist die Sachlage etwas komplexer. Diese Personen sind in der Phase der Entscheidungsfindung besonders zugänglich für Argumente, da sie das Bedürfnis verspüren, schnell zu einer Entscheidung zu kommen. Ist die Entscheidung dann einmal gefällt, wenden sie sich von Argumenten und Persuasionsversuchen ab, da sie ebenfalls das Bedürfnis nach Dauerhaftigkeit empfinden.

Abschließend sei noch angemerkt, dass Persuasion, die durch systematische Verarbeitung und Abwägung von Argumenten aufgebaut wurde, stabiler ist als diejenige durch periphere Verarbeitung.

## 6.2 Critical-Language-Awareness/Sprachbewusstheit

Wie lässt sich nun aber der Zusammenhang von Machtverhältnissen, sozialen Schichten und Spracheinstellungen bewusst machen, um Letztere beeinflussen zu können? Diese Frage haben sich Vertreter der Critical-Awareness-Ansätze (CLA) gestellt. Wie schon angesprochen, entstand in den USA durch die Ann-Arbor-Entscheidung 1979 der Bedarf, Programme für Kinder zu entwickeln, die eine andere Varietät als das standardamerikanische Englisch als Erstsprache besitzen. 1981 wurde z. B. in Kalifornien

7 Personen mit hohem Kognitionsbedürfnis gehen gerne aufwändigen kognitiven Aktivitäten nach, vgl. Stroebe 2014: 242.

ein Gesetz erlassen, wonach der Erwerb des Standardenglischen von „Speakers of Black Language“ im Schulsystem möglich sein müsse und dafür deren Varietät als Grundlage dienen solle (vgl. Deák 2007: 107). Ich werde im Folgenden einige der Ansätze vorstellen.

Die Vorzüge von Critical-Language-Awareness-Ansätzen werden gerne im Kontrast zu „gewöhnlichen“ Language-Awareness-Ansätzen (LA) vorgestellt. Dieses Vorgehen verfolgen auch Clark et al. (1990). Die Tabelle in Abbildung 6.1 liefert hierzu einen Überblick (entnommen aus Clark et al. 1990: 250). Ein zentrales Unterscheidungskriterium der beiden Ansätze besteht darin, dass in einem CLA die soziale Ordnung als veränderbar dargestellt wird. Clark et al. (1990: 250) beschreiben den Nutzen von CLA folgendermaßen:

|     | <i>Objectives</i>   | <i>Motivations</i>                          | <i>Schooling</i>                               | <i>Language</i>   | <i>Learning</i>                    |
|-----|---------------------|---|--|-------------------|------------------------------------|
| LA  | Social integration  | Legitimation of social & socling order      | Fitting children into social order             | Natural order     | Knowledge isolated from practice   |
| CLA | Social emancipation | Critique & change of social & socling order | Fitting children to work in & change soc order | Naturalised order | Knowledge integrated with practice |

Abbildung 6.1: Vergleich Language Awareness mit Critical Awareness

In arguing for a critical LA, we are applying to language a general view of what the main objective of schooling ought to be: developing a critical awareness of the world, and of the possibilities for changing it. In the case of language, it is a matter of coming to see the existing practices of a given ‚sociolinguistic order‘ as socially created within particular social relationships, and therefore socially changeable.

Dabei stellen sie speziell drei Bereiche heraus, die in einem CLA-Lehrplan behandelt werden sollten (vgl. Clark et al. 1991: 48 f):

1. Bewusstmachung des sozialen Diskurses
  - a) Wie beeinflusst der soziale Diskurs den sozialen Kontext?
  - b) Wie trägt der soziale Diskurs zur Reproduktion oder Veränderung von Macht- und Dominanzverhältnissen bei?
  - c) Wie steht die Varietätenwahl von Sprecher:innen und Schreiber:innen damit in Zusammenhang?
2. Kritische Bewusstmachung der Diversität
  - a) Warum werden manche Varietäten als hochwertiger und funktionaler in prestigeträchtigen Situationen eingeschätzt?
  - b) Wie hat sich die aktuelle Bewertung von Varietäten, einschließlich des Standards, historisch entwickelt?

3. Bewusstmachung und praktische Anleitung für Veränderung
  - a) Wie hängen die Veränderung von Machtverhältnissen, sozialen Kämpfen und Spracheinstellungen zusammen?
  - b) Welche Möglichkeiten und Beschränkungen gibt es für Veränderungen?
  - c) Wie können Sprachpraktiken darauf ausgerichtet werden, Veränderung herbeizuführen?

Zur Umsetzung im Erstsprachenunterricht schlagen sie vor, einen Aufsatz über die Familiengeschichte zu schreiben, da diese Aufgabe folgendermaßen in der Lage ist, die drei genannten Bereiche zu reflektieren:

1. Der Austausch mit der Familie stärkt die Identifikation mit den sozialen Wurzeln und macht die eigene Situation im sozialen Raum bewusst (Punkt 1 und 3).
2. Die Aufgabe beinhaltet Wertschätzung der eigenen Erfahrungen des Lernenden gegenüber (Punkt 3).
3. Die Schreibaufgabe habe ein echtes Ziel, echte Leser und eine echte Verbreitung. [Im Original: „the writing would have a real purpose, real readers and real distribution.“]
4. Es erlaubt den Lernenden, viele der Beschränkungen und Konventionen der Gesellschaft zu reflektieren, während sie selbst entscheiden, was für sie relevante und wichtige Inhalte sind, welche Varietäten und Repertoires sie verwenden usw. (Punkt 1, 2 und 3) (vgl. Clark et al. 1991: 51).

Wolfram/Friday (1997) liefern einen Vorschlag, wie ein Critical-Awareness-Programm für die Primarstufe aussehen kann. Drei primäre Ziele werden in diesem Programm verfolgt: humanistische, wissenschaftliche und soziohistorische. In der humanistischen Einheit soll es darum gehen, dass die Schüler:innen mit Dialekt-Stereotypen konfrontiert werden. Als Methode wird vorgeschlagen, unterschiedliche Sprachproben vorzuspielen, und im Anschluss sollen die Schüler:innen in Kleingruppen über die Bedeutung von Sprachvariation und über Vorurteile gegenüber dialektalen Varietäten diskutieren.

In der wissenschaftlichen Einheit sollen Unterschiede zwischen dem Standard und dialektalen Varietäten analysiert werden, um Generalisierungen über die Funktionsweisen anstellen zu können. Dieser Sprachkontrast könne auch den klassischen Grammatikunterricht ersetzen. Die sozialhistorische Einheit handelt von der historischen Herkunft von Varietäten und den Umständen, die zu den aktuellen Verwendungsweisen und Spracheinstellungen geführt haben (vgl. Wolfram/Friday 1997: 146–152).

Eine differenzierte Beschreibung dieser Ansätze liefert Siegel (2006). Er beschreibt in einem ersten Schritt den Awareness-Approach: „[Varieties] are seen as a resource to be used for learning the standard and for education in general“ (Siegel 2006: 163). Ein solcher Ansatz besteht laut Siegel aus drei Komponenten: einer soziolinguistischen, einer Akkommodations- und einer kontrastiven Komponente. Die soziolinguistische Komponente soll dabei den Zusammenhang zwischen Varietäten, Spracheinstellungen und den sozialen Implikationen demonstrieren, im Akkommodationsteil soll Literatur

und Musik in der „eigenen“ Varietät gezeigt werden, und im kontrastiven Teil werden die linguistischen Charakteristika unterschiedlicher Varietäten in kontrastivem Unterricht vermittelt (vgl. Siegel 2006: 163). Er listet verschiedene Programme auf, die mit diesem Ansatz Erfolg hatten, und erklärt dies damit, dass der soziolinguistische Teil die Akzeptanz linguistischer Diversität steigert, der Akkommodationsteil die Motivation steigert und der kontrastive Teil den Schüler:innen die Unterschiede bewusst macht (Siegel 2006: 165). Dieses Vorgehen selbst – wenn auch erfolgreich – ist für ihn jedoch nicht ausreichend, ihm fehlt darin „Empowerment“ zu demonstrieren, also die Erlangung von Macht in bestimmten sozialen Bereichen, in denen Sprachminoritäten ausgeschlossen sind (ähnlich, wie es auch Clark et al. 1991 in Punkt 3 vorschlagen). Der Critical-Awareness-Approach besteht für ihn schließlich aus folgenden Bestandteilen (Siegel 2006: 168 f):

1. Reflexion und Interpretation von Erfahrungen, die mit unterschiedlicher Varietätenverwendung gemacht wurden
2. Verwendung von Varietäten im Unterricht, wo sie normalerweise nicht verwendet werden
3. Als Grundlage für den Unterricht sollen die Varietäten der Schüler:innen dienen, was mit kontrastiver Analyse einhergehen kann.
4. Auch historisches Wissen über Varietäten ist wichtig.

Dieses Vorgehen führt aus seiner Sicht zu folgenden Veränderungen:

1. Die Lehrkräfte werden sich der Sprachideologien bewusst. Dadurch verändern sich deren Spracheinstellungen zum „Positiven“.
2. „Ethnic pride“ wird verstärkt.
3. Die Wahrnehmung, dass das Schulsystem nur die dominante Sprache und Gruppe fördert, wird abgebaut.
4. Wissen über Spracheinstellungen und Sprachideologien ist ebenfalls förderlich für Personen der dominanten Gruppe.

Mallinson/Charity Hudley (2010) erweitern den Fokus auf das Umfeld, in dem die Critical-Awareness-Programme stattfinden. Sie zählen fünf Punkte auf, die sie für unerlässlich für eine erfolgreiche Umsetzung halten (vgl. Mallinson/Charity Hudley 2010: 254):

1. Partnerschaften zwischen Wissenschaftler:innen, Lehrkräften sowie kulturell-diversen Schulen sollen aufgebaut werden.
2. Linguistisches Wissen über Varietäten soll sowohl Lehrkräften als auch Schüler:innen vermittelt werden.
3. Funktionierende Praktiken sollen implementiert werden, mit denen linguistische Informationen an Lehrkräfte weitergegeben werden können.
4. Die Resultate der Lehrkräfteausbildung sollen ausgewertet und ...
5. an Einrichtungen der Bildungspolitik weitergegeben werden.

Einen ähnlichen holistisch ausgerichteten Ansatz verfolgen Bucholtz et al. (2014). Sie beschreiben ihn folgendermaßen:

The SKILLS [School Kids Investigating Language in Life and Society] program is designed to be a genuine partnership that benefits all participants by combining sociocultural linguistic research, graduate training, undergraduate service learning, community engagement, and academic outreach to underserved students in Santa Barbara County. The program is inspired by the educational insight that meaningful and lasting learning is most readily fostered when academic experiences build on students' existing 'funds of knowledge' – especially the linguistic and cultural knowledge they bring from home (Bucholtz et al. 2014: 145).

Das Programm selbst soll fünf Ziele erfüllen (vgl. Bucholtz et al. 2014: 146 f):

1. Bewusst machen und anerkennen von linguistischer Variation mit gleichzeitiger Vermittlung der Systematizität und des kulturellen Werts der Varietäten.
2. Den Wert des linguistischen Repertoires in unterschiedlichen formellen und informellen Situationen herausstellen.
3. Die Schüler:innen anregen, das gelernte Wissen über die Varietäten an andere Personen weiterzugeben.
4. Wissen über die soziopolitische Macht der Varietätenverwendung vermitteln, anregen, dieses Wissen weiterzugeben sowie andere bei Interesse/Bedarf beim Varietätenerwerb zu unterstützen.
5. Alle Sprachbenutzer:innen als linguistische Experten anerkennen, die dazu fähig sind, etwas zum kulturellen Wissen beisteuern zu können sowie zu fördern, dass dieses Wissen auch eingebracht wird.

Einen etwas anderen Blickwinkel auf den Einsatz eines Critical-Awareness-Ansatzes im Unterricht bietet Rutter Strickling (2012). In ihrer Dissertation begleitet sie die Entwicklung mehrerer Lehrkräfte beim Einsatz einer solchen Methode. Sie beschreibt fünf Phasen, die die Lehrkräfte durchlaufen, während sich ihre Einstellungen verändern:

[C]onfronting new linguistic information, questioning linguistic information and previously held beliefs, integrating new linguistic information within contexts of social engagement, evaluating the relational, emotional, and cognitive aspects of the experience, and modifying beliefs about language within the specific context of the experience“ (aus dem Abstract der Dissertation von Rutter Strickling 2012, nicht nummeriert).

Zur Veranschaulichung möchte ich das Beispiel einer schwarzamerikanischen Lehrerin geben, die selbst mit African American Englisch aufwuchs, aber sehr negative Einstellungen dazu hatte (vgl. Rutter Strickling 2012: 236 f). Sie fing mit ihrer Mittelschulklasse an, kontrastive Analysen durchzuführen, und als sie bemerkte, dass dies gute Ergebnisse bei den Schüler:innen erbrachte, wurde sie leidenschaftlich dem Thema gegenüber und berichtete den Schüler:innen von ihrer eigenen Einstellungsveränderung zu African American Englisch. Dies schien für die Schüler:innen der wichtigste Veränderungsschritt der Lehrerin gewesen zu sein. Eine Schülerin sagte dazu:

Seeing [the teacher] change throughout the weeks was extremely powerful [...] [By the end] she went in there knowing that she was going to make a difference in kids' lives. So when you compare [the students'] writing samples from the beginning of the summer to the end, you see a very convincing change.

Allein die Bewusstmachung der Regelmäßigkeit der Sprachsysteme schien diese Wirkung erzielt zu haben. Die Sichtung der Literatur zeigt, dass Critical-Awareness-Ansätze im englischsprachigen Raum schon häufiger zum Einsatz gekommen sind. Empirische Untersuchungen zu den Wirkungen dieser Interventionen fehlen allerdings zumeist. Eine Ausnahme bildet die Arbeit von Johnson et al. (2017). Sie untersuchen die Wirkungen eines Programms, welches vor allem Code-Shifting und die angemessene situationale Verwendung der Varietäten bei Kindern der Primarstufe im Blickfeld hat (Studie mit Interventions- und Kontrollgruppen, n = 115). Dafür wurde der Dialektalitätsgrad anhand von geschriebenen narrativen Texten an zwei Zeitpunkten gemessen. Um die angemessene Varietätenverwendung messen zu können, wurde den Schüler:innen zusätzlich eine „Übersetzungsaufgabe“ gestellt. Dabei sollten sie Sätze für eine jeweilig vorgegebene Situation umschreiben. Die Resultate zeigen, dass dialekt-sprechende Schüler:innen, die im Unterricht spezifisch auf die unterschiedlichen sprachlichen Kontexte hingewiesen wurden, auch mehr Standardformen in Standardkontexten verwenden. Gleichzeitig ist bei den Schüler:innen der Interventionsgruppen höhere Lesekompetenz im Vergleich zu den Kontrollgruppen messbar (vgl. Johnson et al. 2017: 2032).

Auch im deutschsprachigen Raum gibt es schon seit langer Zeit die Forderung, Dialekte im Deutschunterricht zu integrieren. So wurden schon als frühe Replik auf die Bernsteinsche Defizithypothese von Besch und Löffler die Sprachhefte „Mundart/Hochsprache – kontrastiv“ (vgl. Besch 1975: 160) für weite Teile des deutschen Sprachraums entwickelt, die eine Grundlage für den kontrastiven Grammatikunterricht bilden sollten (für das Bairische s. Zehetner 1977).

Auch Bausinger (1983) forderte schon früh die Auseinandersetzung mit Dialekten im Erstsprachenunterricht. Dabei sollten eine wissenschaftliche Perspektive behandelt werden, indem die Schüler:innen den regionalen Dialekt erforschen, ebenso eine historische Perspektive, in der die Entwicklung der Dialekte und Standardsprache vermittelt wird, und eine soziolinguistische Perspektive, indem unterschiedliche Funktionen der Varietäten als Nähe- und Distanzsprache im sozialen Kontext besprochen werden (vgl. Bausinger 1983: 100 ff).

Bücherl (1993) reiht sich ebenfalls ein unter diejenigen, die die thematische Auseinandersetzung mit Dialekten im Unterricht forderten. Er streicht vor allem die Vorteile von bidialektalen Personen heraus, indem sie durch die Akkommodationsmöglichkeit monodialektalen Personen überlegen sind (vgl. Bücherl 1993: 72 ff).

Neuland (1993) schlägt vor, im Deutschunterricht weniger Sprachnorm- und Sprachmängelbewusstsein, sondern neutrales Sprachdifferenzbewusstsein zu fokussieren. Sie schlägt drei Themenbereiche vor, mit denen man dies im Unterricht fördern kann: altersspezifische Sprachvariation wie Babysprache oder Jugendsprache; regionale Sprachvariation zwischen Dialekten, Umgangssprachen und der Standard-

sprache sowie geschlechterspezifische Sprachvariation, wie dies an rollenspielenden Kindern beobachtet werden kann.

Glantschnig (2011) untersucht Normverstöße in Texten von steirischen Schüler:innen und findet diese vor allem in Morphologie und Syntax, bei der Verbalflexion und beim Gebrauch von Konjunktionen in Nebensätzen (vgl. Glantschnig 2011: 230). Ihre Kategorisierung der potenziellen Fehlerquellen liefert eine Vorlage für kontrastiven Grammatikunterricht. Ihre Empfehlung für den Deutschunterricht geht dann auch in diese Richtung:

Die Reflexion der Systemhaftigkeit dialektalen Sprachgebrauchs macht Abweichungen von der Norm transparent – sowohl für die Lehrer als auch für die SchülerInnen. Das Besprechen dialektaler Grammatik bietet die Möglichkeit, Einsichten in grammatische Variation und ihre Ursachen sowie in Sprachwandelprozesse zu gewinnen. Die SchülerInnen sollen sich dabei ihrer inneren Mehrsprachigkeit bewusst werden und diese situationsadäquat einzusetzen lernen. Grundlegende Unterschiede zwischen gesprochener und geschriebener Sprache können in einem dialektorientierten Grammatikunterricht behandelt, das Sprachdifferenzbewusstsein gesteigert und grammatische Zweifelsfälle angesprochen werden (Glantschnig 2011: 230).

Kaiser/Ender (2020) zeigen eine neue Möglichkeit auf, Sprachvariation im Unterricht zu reflektieren, indem sie als Diskussionsgrundlage das Modell der Faktoren und Funktionen von Jakobson/Halle (1960) verwenden. Dieses Modell beschreibt sechs Funktionen von Sprachverwendung. Kaiser/Ender (2020: 18 ff) geben für jede der Funktionen Beispiele, wie sie im Deutschunterricht zur Reflexion von Sprachvariation eingesetzt werden könnten. In folgender Tabelle werden diese Beispiele kurz zusammengefasst:

1. Referentielle Funktion: Ist es möglich, mit Dialekten jeden beliebigen Sachverhalt auszudrücken?
2. Expressive Funktion: Wie hängt die Varietätenverwendung mit dem Ausdruck von Identität bzw. Gruppenzugehörigkeit zusammen?
3. Konative Funktion: Was möchte ich mit meiner sprachlichen Äußerung beim Gegenüber erreichen und wie funktioniert das am besten?
4. Phatische Funktion: Wie würden Freunde reagieren, wenn man plötzlich anders (z. B. im Standard) mit ihnen sprechen würde?
5. Metasprachliche/metakommunikative Funktion: Warum werden gewisse Werbungen im Dialekt gehalten und andere nicht?
6. Poetische Funktion: Welche Wirkung haben Varietäten im literarischen Bereich (z. B. Varietätenverwendung bei österreichischen Autor:innen wie Wolf Haas)?

Daneben unterstreichen Kaiser/Ender zusätzlich Methoden, die wir schon aus anderen Ansätzen kennen. Durch forschendes Entdecken soll Schüler:innen die areale sowie die soziolinguistische Variation nähergebracht werden. Daneben sollen sie produktive und rezeptive Kompetenzen entwickeln. Als Beispiel wird das Rollenspiel genannt, das schon in der Primarstufe eingesetzt werden kann und den weiteren Vorteil besitzt, nicht die persönliche Varietätenverwendung zur Exposition stellen zu müssen. Gleichzeitig

wird auch auf das kontrastive Unterrichtskonzept gesetzt. Schlussendlich soll Sprachvariation z. B. durch Literaturunterricht integrativ erfahren werden.

Die Vermittlung von Critical Awareness bzw. Sprachbewusstheit im Erstsprachenunterricht wird schon seit einiger Zeit gefordert. Es sprechen gute Gründe und viele Belege dafür, einen solchen Ansatz im Erstsprachenunterricht zu verfolgen. Die Sichtung der bisherigen Projekte zeigt, dass vor allem folgende Maßnahmen vielversprechende Erfolge zeigen:

1. Vermittlung linguistischen Wissens durch kontrastive Analyse
2. Vermittlung historischen Wissens über die Herkunft von Varietäten
3. Vermittlung soziolinguistischen Wissens über den Zusammenhang von Varietäten, Spracheinstellungen und Sprachideologien
4. Akkommodationstraining in gesprochen sprachlichen und geschriebensprachlichen Kontexten

Dabei scheint es überdies wichtig zu sein, das Umfeld miteinzubeziehen, indem man Materialien und Schulungen für Lehrkräfte zur Verfügung stellt sowie Kooperationen mit Schulen und Bildungsinstitutionen aufbaut, damit sich das Wissen auf verschiedenen institutionellen Ebenen verbreiten kann. Zugleich ist es sinnvoll, über regionale Medien von den Erfolgen zu berichten und die Forschungserkenntnisse populärwissenschaftlich zu veröffentlichen, um auch der Allgemeinbevölkerung das Wissen über Sprachvariation zugänglich zu machen.



## 7 Materialien zur Bewusstmachung und Reflexion von Spracheinstellungen

So wie das Wissen über den Nutzen von Sprachvariation im Unterricht nichts grundlegend Neues ist, wurden ebenfalls schon nach Aufkeimen der Sprachbarrierendebatte im deutschsprachigen Raum erste Unterrichtsmaterialien entwickelt. Schon angesprochen wurde die Reihe „Mundart/Hochsprache – kontrastiv“ (vgl. Besch 1975), die auch ein Heft für den bairischen Sprachraum herausgegeben hat (vgl. Zehetner 1977). Dieses geht detailliert auf lautliche, lexikalische und grammatische Unterschiede zum Standarddeutschen ein.

Saxalber/Lanthaler (1993) liefern didaktische Empfehlungen für die sprachliche Situation in Südtirol, die sich vom restlichen bairischen Sprachraum dadurch unterscheidet, dass in dieser Region mit Italienisch eine andere Dachsprache existiert und den dort gesprochenen bairischen Dialekten der Rang einer Minderheitensprache zukommt. Auch in neuerer Zeit wurden immer wieder Didaktisierungsvorschläge gemacht, eine kurze Auswahl:

In Erlach/Schurf (2011) findet man eine Sammlung unterschiedlicher Aufgaben aus dem Bereich Sprachwandel, Sprachverfall, Medialität von Sprache und Varietätenvielfalt/innerer Mehrsprachigkeit. Allen gemeinsam ist die Tendenz des sprachreflexiven Aufbaus der Aufgaben.

Bangel (2012) regt zur Reflexion von sprachpuristischen Positionen an, indem sie Fremdwörter in der Musik- und Jugendsprache kontrastiert. In Bayern erschienen auf ministerialer Ebene zwei Handreichungen mit einer breiten Palette von Vorschlägen, wie man Dialekte im Unterricht thematisieren kann. Die Vorschläge reichen von Aufgaben zur regionalen Variation über die Vermittlung sprachhistorischen Wissens, lexikalische Kontrastübungen, der Analyse von Dialekten in der Werbung bis zur Dialekt-dichtung (Bayerisches Staatsministerium für Bildung und Kultus 2015; für Unterricht und Kultus 2019). Diese Bemühungen mündeten in einer digitalen Sammlung dieser Aufgaben auf der Homepage des Staatsinstitut für Schulqualität und Bildungsforschung (ISB).

Ganz aktuell wird unter Federführung von Hanna Fischer (2023) eine Sammlung von Unterrichtsmaterialien für Sprachvariation an der Universität Marburg unter dem Stichwort Sprachvariation@Schule eingerichtet. Wölke (2018) liefert einen umfangreichen Didaktisierungsvorschlag, der aus fünf Bausteinen besteht: einem Einstiegsmodul, in dem Sprache als Zeichensystem dargestellt wird, einem Dialektmodul, einem Fachsprachenmodul, einem soziolinguistischen Modul und einem Abschlussmodul über das linguistische Relativitätsprinzip.

Das Heft von Tophinke (2019) beinhaltet fünf Didaktisierungsvorschläge für unterschiedliche Altersgruppen, ebenfalls mit sprachreflexivem Fokus. Besonders hervorzuheben ist die Arbeit des Projektteams um Heike Wiese mit dem Titel Deutsch ist

vielseitig, das sowohl Unterrichtsmaterialien für alle Altersgruppen erarbeitet hat als auch Fortbildungen für Lehrkräfte und Erzieher:innen anbietet. Die Unterrichtsmaterialien umfassen einen breiten Themenbereich und sind didaktisch so aufbereitet, dass sie sofort im Unterricht verwendet werden können.

Österreichspezifische Materialien sind dabei eher selten aufzufinden. Eine Ausnahme bilden die Didaktisierungen aus dem Projekt Österreichisches Deutsch als Unterrichts- und Bildungssprache rund um Rudolf de Cillia (vgl. de Cillia/Fink/Ransmayr 2013; de Cillia/Fink/Ransmayr 2017; de Cillia/Ransmayr 2019). Dabei liegt ein zentraler Fokus dieses Projekts darin, Lehrkräfte auf standardsprachliche Varianten des österreichischen Deutsch zu sensibilisieren, um z. B. hyperkorrektiven Tendenzen (vgl. Fink 2014) zuvorzukommen. Einen weiteren Didaktisierungsvorschlag für das österreichische Deutsch liefert Schweiger (2014), der vorschlägt, sprachliche Auffälligkeiten von österreichischen Autoren wie Wolf Haas zum Ausgangspunkt von Sprachreflexion und dem Aufbau von Sprachbewusstheit zu verwenden.

Abschließend sei noch einmal Glantschnig (2011) erwähnt, die mit ihrer Analyse und Kategorisierung von potenziellen Normabweichungen in Texten von steirischen Schüler:innen eine Ausgangslage für kontrastiven Sprachunterricht in dieser Region liefert.

Obwohl die Forschungscommunity im deutschsprachigen Raum schon einiges ausgearbeitet hat, existieren keine Materialien, die alle Inhaltsbereiche eines Critical-Awareness-Ansatzes enthalten würden. Auch die Zusammenstückelung von existierenden Aufgaben schien keine praktikable Lösung für den kontrollierten Unterrichtseinsatz darzustellen. Daher wurden im vom Land Salzburg geförderten Projekt Sprachliche Vielfalt verstehen, reflektieren und ausbauen eigene Unterrichtsmaterialien entworfen, die inhaltlich konsistent Sprachbewusstheit (critical language awareness) im mittelbairischen Raum fördern sollen. Dabei handelt es sich um fünf Module à zwei Unterrichtseinheiten für die 10. Schulstufe, in denen in folgender Art und Weise die Maßnahmen der Critical-Awareness-Ansätze eingesetzt werden:

- Modul 1: Herstellung eines geschützten Kommunikationsrahmens, Reflexion über Sprachvariation, Sprachakkommodation sowie Sprache und Identität
- Modul 2: Vermittlung system- und areallinguistischen Wissens, Ausbau von Sprachverwendungs- und Sprachakkommodationskompetenzen
- Modul 3: Vermittlung sprachhistorischen Wissens, Ausbau von Sprachakkommodationskompetenzen
- Modul 4: Vermittlung soziolinguistischen Wissens und soziolinguistischer Kompetenzen, Ausbau zielgerichteter Verwendung von Sprachregistern
- Modul 5: Festigung des Wissens und der erworbenen Kompetenzen, kreatives Abschlussprojekt

Die folgenden Abschnitte liefern einen Kurzüberblick der Modulinhalte. Die Materialien sowie eine Lehrkräftehandreichung, die methodisch-didaktische sowie fachwissenschaftliche Hintergrundinformationen für die einzelnen Aufgaben enthält, sind unter [spravive.com](http://spravive.com) verfügbar.

## 7.1 Meine, deine, unsere Sprachen

Das erste Modul beginnt mit einer Diskussion über die individuellen Formen der Mehrsprachigkeit, die die Schüler:innen mitbringen. Dabei werden die Konzepte der inneren und äußeren Mehrsprachigkeit induktiv erarbeitet. In der nächsten Aufgabe werden potenzielle Vorteile von äußerer und danach innerer Mehrsprachigkeit skizziert. Die selbst gefundenen Vorteile sollen mithilfe der Lehrkraft zu folgender Liste erweitert werden: Verständigung, interkulturelle Kompetenz/Toleranz, Erschließungs- und Lernkompetenz, bessere Aufmerksamkeitssteuerung, medizinische Vorteile und Vorteile am Arbeitsmarkt. Auch potenzielle Nachteile wie die doppelte Halbsprachigkeit und der geringere Wortschatz von bilingualen Kindern sollen diskutiert werden. Im Anschluss wird die Erschließungskompetenz an einem niederländischen Wikipedia-Eintrag geübt.

Die nächste Aufgabe widmet sich dem Varietätenkontinuum Standarddeutsch – Umgangssprache – Dialekt, indem die Begriffe mit Beispielen erklärt werden. Danach wird das von den Schüler:innen im Alltag praktizierte situative Varietätenshifting anhand von unterschiedlichen Situationen bzw. Personen reflektiert. Der letzte Abschnitt dieses Moduls widmet sich dem Thema Sprache und Identität. Stereotype Spracheinstellungen werden hier reflektiert und die Bedeutung von Varietäten und Sprachen für die individuelle Identität bewusst gemacht. Dieses Kapitel soll einen geschützten, wertschätzenden Rahmen für die Analyse von Sprachvariation und -einstellungen bereitstellen, indem über die tägliche Sprachvariation/Akkommodation reflektiert und diese auch als Potenzial konzipiert wird. Zusätzlich werden zentrale Begriffe eingeführt und ein Zusammenhang zwischen Sprache und Identität hergestellt. Das Modul lässt sich somit nicht unter einen der zentralen Bereiche eines Critical-Awareness-Ansatzes subsumieren, bereitet aber den Weg dahin und ist für die motivationale Ebene wichtig, da es den Bezug zu den Schüler:innen selbst herstellt.

## 7.2 Gemeinsam Dialekt erforschen

Dieses Modul widmet sich im ersten Teil der Vermittlung von Wissen über areale Variation im deutschsprachigen Raum. Die zentrale Methode bildet hierfür die Erschließung einer Karte, die die wichtigsten Dialektgebiete des deutschsprachigen Raums enthält und mit Inhaltswissen, Reflexionsfragen und Sprachbeispielen kombiniert ist. Im Anschluss werden mithilfe eines Concept Cartoons<sup>8</sup> zwei Konzepte, die der wissenschaftlichen Forschung entsprechen, und drei Fehlkonzepte zum Thema der sprachlichen Korrektheit von Dialekten präsentiert. Letztere werden im Laufe des Moduls durch das vermittelte Inhaltswissen entkräftet. Zentral hierfür ist das Fragebuch. In diesem findet man 19 Wörter im Standard, die im mittelbairischen Dialektraum regelhafte Unter-

---

8 Das Konzept des Concept Cartoons stammt aus dem historischen und politwissenschaftlichen Kontext (vgl. z. B. Buchberger/Eigler/Kühberger 2020). Wir haben es im Projekt für den Deutschunterricht adaptiert.

schiede zeigen. In einem ersten Schritt sollen die Schüler:innen die dialektalen Formen bei einer dialektsprechenden Person erheben. Die Wörter sind so ausgesucht, dass sie sechs in der Region typischen und vielgebrauchten Dialektphänomenen zugeordnet werden können (*a*-Verdumpfung, *l*-Vokalisierung, nicht durchgeführte Monophthongierung, Ausbleiben der Senkung vor Nasal, Synkope im Präfix bzw. Präfixtilgung beim Partizip II). Die induktive Erarbeitung dieser Phänomene ist dann auch der Arbeitsauftrag im nächsten Schritt. Dass bei Erstsprachler:innen die Dialektphänomene automatisiert produziert werden können, zeigt eine Folgeübung, in der Fantasiewörter übersetzt werden sollen. Als Abschluss des Moduls werden zwei grammatische Muster des Dialektraums gezeigt, bei denen kein Pendant im Standard existiert. Dabei handelt es sich um das *da*-Präfix und den Inchoativ.

In diesem Modul wird linguistisches Wissen vermittelt. Dies geschieht mit der Bewusstmachung der arealen Variation im deutschsprachigen Raum, der Bewusstmachung der Regelmäßigkeit dialektaler Phänomene, dem Aufzeigen, dass Dialekt auch Elemente enthält, die in der Standardsprache nicht existieren, und der daraus folgenden Berichtigung von in der Bevölkerung weit verbreiteten Fehlkonzepten über die sprachliche Korrektheit (eher: Inkorrektheit) von Dialekt.

### 7.3 Altes Deutsch – neues Deutsch

Das dritte Modul widmet sich der Sprachgeschichte und dem Sprachwandel. Zu Beginn wird die deutsche Sprache in den Kontext der indogermanischen Sprachen gesetzt. Auf einer Karte sollen die Schüler:innen die Verwandtschaftsbeziehungen dieser Sprachfamilie kennenlernen. Dies wird begleitet mit einer Übung, in der Wörter aus unterschiedlichsten indogermanischen Sprachen miteinander verglichen werden und damit die Verwandtschaft induktiv erfahren werden soll. Danach verschiebt sich der Fokus auf die Entwicklung der deutschen Sprache. Diese wird anhand einer Abbildung in ihren wichtigsten Phasen beschrieben und durch eine Übersetzungsübung vom Niederdeutschen ins Hochdeutsche und umgekehrt können die Regeln der 2. Lautverschiebung induktiv erfahren werden. Zusätzlich wird gezeigt, dass sich das Englische in dieser Hinsicht (alter Lautstand der Plosive) wie das Niederdeutsche verhält. Ergänzend dazu zeigt eine Reliefkarte des deutschsprachigen Raums die Herkunft der Ausdrücke „Niederdeutsch“ und „Hochdeutsch“. Im Anschluss daran findet sich ein mittelhochdeutscher Text, der ins Neuhochdeutsche und in den bairischen Dialekt übersetzt werden soll. Der Text beinhaltet dabei drei der sechs aus Modul 2 bekannten Dialektphänomene (Senkung vor Nasal, alte Diphthonge, *n*-Suffix bei Nomen mit weiblichem Genus, die auf *e* enden), die sich in den Übersetzungen des Neuhochdeutschen und Bairischen unterscheiden. Damit soll gezeigt werden, dass Dialekte alte Wurzeln haben. In der nächsten Aufgabe wird durch ein Experteninterview in Form eines Videos die Entwicklung der Standardsprache skizziert. Als Folgeübung sollen in einer Mindmap die wichtigsten Schritte rekapituliert werden. Den Abschluss des Moduls bildet eine Aufgabe, in

der gezeigt wird, dass Anglizismen, die in der Jugendsprache verwendet werden, dialektal integriert werden. Gezeigt wird dies an der regelhaften Synkopierung bzw. Tilgung des Präfixes im Partizip II.

Dieses Kapitel widmet sich der Vermittlung des historischen Wissens über Sprachvarietäten. Es zeigt, dass der deutsche Sprachraum in eine große indogermanische Sprachfamilie eingebunden ist, dass Dialekt alte Wurzeln hat, aber auch durch die Integration von Fremdwörtern lebendig ist, und dass es sich bei Standarddeutsch um eine konstruierte Varietät handelt, die viel jünger ist als die Dialekte.

## 7.4 Das sagt man (nicht)!!?

Das vierte Modul widmet sich der situativen Angemessenheit von Varietäten. Dabei wird auf das Modell von Koch/Oesterreicher (1994) zur Erklärung von Nähe- und Distanzsprache zurückgegriffen. Zentral hierfür ist der Vergleich zweier geschriebensprachlicher Texte – ein Schulbuchtext über die Wurzeln der deutschen Sprache und eine Messenger-Textnachricht zwischen zwei Schulfreundinnen im Dialekt, in der der Inhalt des Schulbuchtexts wiedergegeben wird. Diese Texte sollen in einem ersten Schritt auf nahe- und distanzsprachliche Phänomene hin untersucht werden. Dadurch sollen diese Konzepte induktiv erkannt werden. In einem weiteren Schritt werden sie in einer vereinfachten Version des Modells von Koch/Oesterreicher explizit beschrieben und weitere Kontexte sollen im Modell ergänzt werden. Eine weitere Aufgabe widmet sich jugendsprachlichen Ausdrücken und der Frage, welchen Zweck, welche Zeichenfunktion diese in der sozialen Kommunikation besitzen. Zum Abschluss des Moduls werden gängige Mythen zur Sprachvariation im schulischen Kontext (dialekt-sprechende Kinder haben größere Probleme mit dem Standardspracherwerb; Leistungen der Schüler:innen verbessern sich, wenn man Dialekte aus der Schule verbannt; Dialektsprecher:innen finden schwerer Arbeit; Anglizismen bedrohen die deutsche Sprache; wenn Kinder mit zwei Sprachen aufwachsen, können sie keine richtig gut) direkt thematisiert und durch fachwissenschaftliches Wissen berichtigt.

Dieses Modul zielt konkret auf die Vermittlung soziolinguistischen Wissens ab. Es werden Kontexte diskutiert, in denen die Verwendung von Dialekt bzw. Standard als angemessener in der Bevölkerung wahrgenommen wird. Daneben wird gezeigt, dass diese Kontexte gut in einem Kontinuum von Nähe- und Distanzsprache konzeptualisiert werden können. Die Funktion von Jugendsprache wird reflektiert und Mythen, die sich um Dialektverwendung und Mehrsprachigkeit in Schule und Arbeitswelt ranken, werden entkräftet.

## 7.5 Vorhang auf: Das haben wir gelernt!

Das letzte Modul widmet sich einerseits der Wiederholung des erworbenen Wissens. Dafür wurde ein Kreuzworträtsel erstellt, das 21 teils knifflige Fragen enthält. Zur Lösung ist es erlaubt (wenn nicht sogar erwünscht), in den einzelnen Aufgaben nachzuschlagen. Andererseits finden sich in diesem Modul Vorschläge für die Umsetzung eines kreativen Abschlussprojekts, aufgegliedert nach den vier Inhaltsmodulen. Diese kreative Beschäftigung soll dazu beitragen, das erworbene Wissen noch weiter zu festigen.

## 8 Erhebungsinstrumente und Auswertungsmethoden

Für die praktische Durchführung einer empirischen Studie kommt sowohl der Wahl der Erhebungsinstrumente als auch der der Auswertungsmethoden zentraler Stellenwert zu. Im Idealfall sind beide aufeinander abgestimmt, werden gemeinsam geplant und sind in der Lage, die im Vorfeld aufgeworfenen Hypothesen zu belegen (genauer: die Nullhypothesen bestätigen oder verwerfen zu können). Daher werden in einem ersten Abschnitt dieses Kapitels die zentralen Fragestellungen und Hypothesen besprochen. Danach wird die Erstellung des Fragebogens thematisiert, der in dieser Studie das Erhebungsinstrument darstellt. Dabei wird der Fokus auf die Matched-Guise-Blöcke gelegt, die die zentralen Daten für die Analyse liefern. Im Anschluss werden die Auswahl der Schulen, die Stichprobe und die Datenerhebung erläutert, bis im letzten Abschnitt die statistischen Auswertungsmethoden beschrieben werden.

### 8.1 Fragestellungen und Hypothesen

Die zentrale Idee der Untersuchung ist es, einen Critical-Language-Awareness-Ansatz (s. Abschnitt 6.2) im deutschsprachigen Sprachraum zu implementieren und zu evaluieren. Damit soll der Frage nachgegangen werden, ob sich stereotype Spracheinstellungen zu den Varietäten Dialekt und Standarddeutsch (s. Abschnitt 4.1) auf Kompetenz- und auf Sympathieebene durch Bewusstmachung, Reflexion und den Aufbau von Wissensbeständen im Deutschunterricht verändern lassen.

Spracheinstellungen sind jedoch nicht homogen in der Bevölkerung verteilt. Wie wir schon in Kapitel 4 gesehen haben, findet man beispielsweise eine unterschiedliche Ausprägung der Spracheinstellungen bei den Geschlechtern (s. Abschnitt 4.2), bei unterschiedlichen Erstsprachen (s. Abschnitt 4.3) oder in unterschiedlichen sozialen Schichten (s. Einleitung zu Kapitel 4). In Abschnitt 4.4 wurde gezeigt, dass Stadt-Land-Unterschiede im Sprachgebrauch existieren – am Land wird mehr Dialekt gesprochen. Daher liegt es nahe, dass sich potenziell auch die Spracheinstellungen in diesen Regionen unterscheiden.<sup>9</sup>

Diese Ergebnisse aus der Spracheinstellungsforschung bildeten die Grundlage für die Bildung von übergeordneten Fragestellungen und daraus ableitbaren Hypothesen. Diese dienten als Basis für die Erstellung des Fragebogens. Dabei ist es sinnvoll, die Fragestellungen in zwei Bereiche zu unterteilen. Dabei wird im ersten Bereich der

---

<sup>9</sup> Der Fragebogen selbst beinhaltet noch weitere Frageblöcke, die in dieser Arbeit jedoch nicht analysiert und damit auch nicht thematisiert werden. Bei Interesse kann die Papiervorlage der verwendeten LimeSurvey-Umfrage auf meiner OSF-Seite (<https://osf.io/7mhwtf/>) eingesehen werden.

allgemeinen Frage nachgegangen, ob die Spracheinstellungen, die die Schüler:innen im Prätest mitbringen, denjenigen entsprechen, die in der Forschung berichtet werden. Der zweite Bereich widmet sich im Anschluss der Frage, ob sich die erhobenen Spracheinstellungen verändern lassen. Im Folgenden sind die übergeordneten Fragestellungen ( $F_x$ ), Hypothesen ( $H_x$ ) und Nullhypothesen ( $H_{xn}$ ) dargestellt:

### Welche Spracheinstellungen bringen die Schüler:innen mit?

- F<sub>1</sub>: Werden standardsprechende Personen kompetenter eingeschätzt als dialekt-sprechende Personen?
  - H<sub>1</sub>: Standardsprechende Personen werden kompetenter eingeschätzt als dialekt-sprechende Personen.
  - H<sub>1n</sub>: Es gibt keinen Unterschied in der Kompetenzeinschätzung von standard- und dialekt-sprechenden Personen.
- F<sub>2</sub>: Werden dialekt-sprechende Personen sympathischer eingeschätzt als standard-sprechende Personen?
  - H<sub>2</sub>: Dialekt-sprechende Personen werden sympathischer eingeschätzt als standardsprechende Personen.
  - H<sub>2n</sub>: Es gibt keinen Unterschied in der Sympathieeinschätzung von dialekt- und standardsprechenden Personen.
- F<sub>3</sub>: Weisen Schülerinnen unabhängig von der Sprachvarietät allgemein positivere Spracheinstellungen in Bezug auf Kompetenz- als auch Sympathieebene auf als Schüler?
  - H<sub>3</sub>: Schülerinnen weisen unabhängig von der Sprachvarietät allgemein positivere Spracheinstellungen in Bezug auf Kompetenz- als auch Sympathieebene auf als Schüler.
  - H<sub>3n</sub>: Es gibt keinen Unterschied in der Einschätzung von Spracheinstellungen auf Kompetenz- und Sympathieebene zwischen Schülerinnen und Schülern.
- F<sub>4</sub>: Schätzen Schülerinnen standardsprechende Personen sympathischer ein als Schüler?
  - H<sub>4</sub>: Schülerinnen schätzen standardsprechende Personen sympathischer ein als Schüler.
  - H<sub>4n</sub>: Es gibt keinen Unterschied in der Sympathieeinschätzung standard-sprechender Personen zwischen Schülerinnen und Schülern.
- F<sub>5</sub>: Schätzen Schüler:innen am Land dialekt-sprechende Personen kompetenter ein als Schüler:innen in der Stadt?
  - H<sub>5</sub>: Schüler:innen am Land schätzen dialekt-sprechende Personen kompetenter ein als Schüler:innen in der Stadt. Schüler:innen in Pendeldistanz liegen mit ihrer Einschätzung zwischen den beiden anderen Regionen.
  - H<sub>5n</sub>: Es existiert kein Unterschied zwischen den Kompetenzeinschätzungen von Schüler:innen am Land, in Pendeldistanz zur Stadt und der Stadt in Bezug auf dialekt-sprechende Personen.

- F<sub>6</sub>: Schätzen Schüler:innen in der Stadt standardsprechende Personen sympathischer ein als Schüler:innen am Land?
- H<sub>6</sub>: Schüler:innen in der Stadt schätzen standardsprechende Personen sympathischer ein als Schüler:innen am Land. Schüler:innen in Pendeldistanz liegen mit ihrer Einschätzung zwischen den beiden anderen Regionen.
- H<sub>6n</sub>: Es existiert kein Unterschied zwischen den Sympathieeinschätzungen von Schüler:innen am Land, in Pendeldistanz zur Stadt und der Stadt in Bezug auf standardsprechende Personen.
- F<sub>7</sub>: Gibt es einen Unterschied in der Sympathieeinschätzung von Schüler:innen mit Deutsch als Erstsprache, Deutsch und einer weiteren Erstsprache und einer anderen Erstsprache als Deutsch in Bezug auf dialektsprechende Personen?
- H<sub>7</sub>: Schüler:innen mit Deutsch als Erstsprache schätzen dialektale Sprechweise sympathischer ein als Schüler:innen mit einer anderen Erstsprache als Deutsch. Schüler:innen mit Deutsch und einer weiteren Erstsprache liegen mit ihrer Einschätzung zwischen den beiden anderen Gruppen.
- H<sub>7n</sub>: Es existiert kein Unterschied in der Sympathieeinschätzung von Schüler:innen mit Deutsch als Erstsprache, Deutsch und einer weiteren Erstsprache und einer anderen Erstsprache als Deutsch in Bezug auf dialektsprechende Personen.
- F<sub>8</sub>: Schätzen Schüler:innen mit unterschiedlichen Bildungshintergründen dialektsprechende Personen auf Kompetenzebene unterschiedlich ein?
- H<sub>8</sub>: Schüler:innen mit niedrigem Bildungshintergrund schätzen dialektsprechende Personen kompetenter ein als Schüler:innen mit hohem Bildungshintergrund.
- H<sub>8n</sub>: Es existiert kein Unterschied in der Kompetenzeinschätzung von Schüler:innen mit niedrigem oder hohem Bildungshintergrund in Bezug auf dialektsprechende Personen.
- F<sub>9</sub>: Schätzen Schüler:innen mit unterschiedlichen Bildungshintergründen standardsprechende Personen auf Sympathieebene unterschiedlich ein?
- H<sub>9</sub>: Schüler:innen mit hohem Bildungshintergrund schätzen standardsprechende Personen sympathischer ein als Schüler:innen mit niedrigem Bildungshintergrund.
- H<sub>9n</sub>: Es existiert kein Unterschied in der Sympathieeinschätzung von Schüler:innen mit niedrigem oder hohem Bildungshintergrund in Bezug auf standardsprechende Personen.

### **Lassen sich Spracheinstellungen verändern?**

Da Forschungsergebnisse zur Veränderbarkeit von Spracheinstellungen im Deutschunterricht noch nicht vorliegen, ist es nicht möglich, differenzierte, forschungsbasierte Hypothesen zu entwickeln. Daher sind die Fragestellungen und Hypothesen in diesem Bereich sehr allgemein gehalten.

- F<sub>10</sub>: Lassen sich Spracheinstellungen in Bezug auf die Kompetenzeinschätzung dialektsprechender Personen verändern?
- H<sub>10</sub>: Die Intervention sorgt dafür, dass dialektsprechende Personen nach der Intervention von den Schüler:innen kompetenter eingeschätzt werden.
- H<sub>10n</sub>: Schüler:innen schätzen nach erfolgter Intervention dialektsprechende Personen auf Kompetenzebene gleich ein wie zuvor.
- F<sub>11</sub>: Lassen sich Spracheinstellungen in Bezug auf die Sympathieeinschätzung standardsprechender Personen verändern?
- H<sub>11</sub>: Die Intervention sorgt dafür, dass standardsprechende Personen nach der Intervention von den Schüler:innen sympathischer eingeschätzt werden.
- H<sub>11n</sub>: Schüler:innen schätzen nach erfolgter Intervention standardsprechende Personen auf Sympathieebene gleich ein wie zuvor.

Auch weitere Überlegungen, die sich nicht auf bestehende Forschungsergebnisse stützen konnten, haben für die Erstellung des Fragebogens und die Auswertungen in dieser Arbeit eine Rolle gespielt (s. Abschnitte 8.2 und 8.6). Damit setzt sich diese Untersuchung aus einer Mischung von hypothesengeleiteter und explorativer Forschung zusammen.

## 8.2 Fragebogendesign

Für die Konzeption der Umfrage selbst wurde ein Prätest-Posttest-Design gewählt. An diesem Versuchsaufbau existiert zwar Kritik aus der Spracherwerbsforschung an der geringen Anzahl der Testzeitpunkte (z. B. Kliesch/Pfenninger 2021), mehrere Testzeitpunkte wären im schulischen Regelunterricht nur schwer realisierbar gewesen und hätten unter Umständen die generelle Durchführbarkeit und Genehmigung des Projekts verhindert. Gleichzeitig ist es fraglich, ob mehrere longitudinale Testungen die Entwicklung von Spracheinstellungen graduell abzubilden vermögen. Da die Erfragbarkeit von Einstellungen keine große Variation erlaubt, ist es wahrscheinlich, dass das häufige Abfragen ähnlicher Items zu Ermüdungseffekten bei den Proband:innen führt.

Das Prä-Posttest-Design wurde mit Interventions- sowie Kontrollgruppen konzipiert. Da es sich um eine Feldstudie mit Schulklassen handelt, kann man von einem quasiexperimentellen Versuchsaufbau mit Clusterstichprobe sprechen, die Ähnlichkeiten zu faktoriellen Designs aufweist.

Für die Erhebung von Spracheinstellungen wurde als zentrale Methode die Matched-Guise-Technik verwendet (s. für eine genaue Beschreibung der Methode Abschnitt 3.4). Im Fragebogen kamen zwei Guises-Blöcke, Lehrkräfte und Schüler:innen, zur Anwendung. Zwischen diesen beiden Blöcken existiert eine Trennung durch andere Items im Fragebogen. Jeder Block wurde von jeweils zwei ungeschulten Sprecherinnen und Sprechern eingeschrieben. Die Schüler:innen-Guises wurden von zwei

Schülern und zwei Schülerinnen aus dem Land Salzburg aus einer sechsten Schulstufe angesprochen, die nicht bei der Untersuchung mitmachen, uns aber von Lehrkräften als bivariable Sprecher:innen empfohlen wurden. Bei den Lehrkräfte-Guises stammten die beiden Sprecher aus dem Untersuchungsraum, die Sprecherinnen nicht. Eine stammt ursprünglich aus Tirol und eine aus Oberösterreich. Beide sind jedoch in der Lage, eine im Salzburger Zentralraum gebräuchliche Varietät zu verwenden. Die Sprecher:innen waren zum Zeitpunkt der Aufnahmen zwischen 25 und 35 Jahren alt.<sup>10</sup>

Die Blöcke enthalten unterschiedliche Formalitätsgrade – eine formelle Situation (Instruktion im Deutschunterricht) und eine informelle (z. B. Ganggespräch). Pro Block existieren also vier unterschiedliche Sätze, die einmal in einer umgangssprachlichen Varietät des Salzburger Zentralraums und einmal in österreichischem Standarddeutsch realisiert wurden, was insgesamt vier Matched-Guises bzw. acht Sätze pro Block ergibt. Die Sätze selbst sind so gestaltet, dass sie ohne lexikalische Ersetzungen und grammatische Umstellungen in Dialekt und Standarddeutsch gesprochen werden können. Die Unterscheidung liegt also primär auf lautlicher Ebene und enthält typische Phänomene des Mittelsächsischen wie *a*-Verdumpfung, *i*-Vokalisierung, Zusammenfall der stimmhaften und stimmlosen Bilabial- und Dentalplosive zu stimmlosen Lenislauten, bairische Personalpronomen, historische Diphthonge, Schwund von stimmlosen Velarfrikativen im Auslaut, unterschiedliche Flexionssuffixe in Verbalparadigmen usw. Um den Aufbau der Guises noch näher zu erläutern, folgen zwei Beispielsätze. Den Ersten spricht die Lehrerin in formeller Situation:

Heute nehmen wir einen neuen Inhalt für die nächste Schularbeit durch. Nehmt dafür das Arbeitsbuch heraus, schlagt die Seite 44 auf und lest den Text ganz oben auf der Seite.

Folgendes Beispiel spricht der Lehrer in der informellen Situation:

So, ihr zwei. Ihr habt ja vorher Fragen zur Partnerarbeit gehabt. Ist jetzt alles klar bei euch, oder braucht ihr noch was von mir?

Die Umfrage selbst wurde mit der Offlineversion von LimeSurvey *Offline Surveys* durchgeführt, um von den WLAN-Netzwerken der Schulen unabhängig zu sein. Die einzelnen Items wurden innerhalb der Blöcke randomisiert, um Reihungseffekte auszuschließen. Ein Nebeneffekt der Randomisierung liegt darin, dass die Guise-Paare einer Person in seltenen Fällen nacheinander abgespielt werden, was potenziell Einfluss auf die Ergebnisse haben kann. Zu jedem Satz wurde eine Frage zu den beiden zentralen Ebenen der Einstellungen Kompetenz und Sympathie sowie eine Frage zur persönlichen Präferenz gestellt (s. zu den Einstellungsdimensionen auch Abschnitt 3.2). Für die Befragung wurde jeder Guise auf diesen drei Ebenen geratet. Die Fragen, die dafür

<sup>10</sup> Die Wahl männlicher und weiblicher Sprechstimmen für die Guises führt in den Analysekapiteln zu einem recht geschlechterbinären Forschungssetting. Dabei ist es nicht die Absicht dieser Arbeit, sich in Fragen der Geschlechterkategorisierung zu positionieren und ein diesbezüglich traditionelles bzw. veraltetes Rollenbild zu vertreten. Vielmehr war dies eine forschungspragmatische Entscheidung dahingehend, dass wir davon ausgegangen sind, dass weibliche Stimmen als weiblich und männliche Stimmen als männlich interpretiert werden. Es gibt auch meines Wissens nach keinen Hinweis in der geschlechterbasierten soziolinguistischen Forschung darauf, dass Stimmen nichtbinär kategorisiert werden.

gewählt wurden, sind in Abbildung 8.1, einem originalen Ausschnitt des digitalen Fragebogens, ersichtlich. Dabei zielt der erste Satz auf die Ebene der persönlichen Präferenz, der zweite auf die Sympathie- und der dritte auf die Kompetenzebene.

|  | ☹️                    |                       |                       | 😐                     |                       |                       | 😊                     |
|--|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Wie <b>gerne</b> hättest du Unterricht bei dieser Lehrerin?      | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Wie <b>nett</b> kommt dir die Lehrerin vor?                      | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Was meinst du: Wie <b>gut</b> erledigt die Lehrerin ihre Arbeit? | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

**Abbildung 8.1:** Likert-Skala für die Abfrage der Präferenz-, Sympathie- und Kompetenzebene

### 8.3 Auswahl der Schulen

Für die Auswahl der Schulen war neben der Kooperationsbereitschaft der Schulen die Möglichkeit, Stadt-Land-Unterschiede analysieren zu können, ein entscheidendes Kriterium. Die Idee war es, jeweils zwei Schulen in der Stadt Salzburg, zwei in Pendeldistanz zur Stadt Salzburg und zwei außerhalb der Pendeldistanz zu finden und zur Teilnahme zu animieren. Dabei sollte es sich um Gymnasien handeln, die zumindest zwei parallele Klassen in der 10. Schulstufe führen.

Für die Suche spielte also zusätzlich die Lage und Größe der Schulen eine Rolle. Das Land Salzburg besteht aus fünf Bezirken, sogenannten Gauen, und der Landeshauptstadt, der Stadt Salzburg. Der Flachgau und der Tennengau, die sich beide größtenteils im Alpenvorland befinden, liegen in Pendelnähe zur Stadt Salzburg und bilden mit der Stadt Salzburg den zentralen Handels-, Wirtschafts- und Verwaltungsraum im Land Salzburg. Die drei Gebirgsгаue, der Pinzgau, der Pongau und der Lungau, befinden sich südlich davon und unterscheiden sich auch sprachlich (s. Kapitel 2, insbesondere Abschnitt 2.1). Die Pendeldistanz für Schüler:innen ist speziell vom Lungau wie auch vom Pinzgau in den Zentralraum praktisch unmöglich, was einerseits an der schieren Weglänge, andererseits an der Erreichbarkeit durch öffentliche Verkehrsmittel liegt, die nur sehr eingeschränkt gegeben ist.<sup>11</sup> Etwas anders sieht das für den Pongau aus. Die Weglänge der Bezirkshauptstadt St. Johann in die Stadt Salzburg ist zwar mit etwas über 60 Kilometern nicht gering, allerdings mit dem Salzburger Regionalexpress gut zu erreichen. Aufgrund dieser Zwischenlage wurde der Pongau in den Überlegungen zur Schulakquise erstmal hintenangestellt. Im Lungau ergibt sich eine andere Problematik. Diese ist in der geringen Anzahl der Einwohner dieses Bezirks zu sehen. So existiert im Lungau nur ein einziges Gymnasium, welches auch, speziell in

<sup>11</sup> Die Weglänge von den Bezirkshauptstädten Tamsweg bzw. Zell am See in die Stadt Salzburg beläuft sich jeweils auf ca. 100 Kilometer.

der Oberstufe, nicht immer über zwei Parallelklassen verfügt. Daher wurde auch das Gymnasium Tamsweg ans Ende der Liste der möglichen Schulen gesetzt. Trotz der schwierigen äußeren Umstände infolge der Coronapandemie wurden sechs Schulen, die über Parallelklassen verfügen und die in den drei Stadt-Land-Kategorien liegen, gefunden – zwei Schulen in der Stadt Salzburg, jeweils eine Schule im Flach- und eine im Tennengau und zwei Schulen im Pinzgau. Für alle drei Regionen wurden die gleichen Guises verwendet. Diese sprechen einen Dialekt bzw. eine Umgangssprache des Salzburger Zentralraums, der sich sprachlich vom Pinzgau abhebt. Anekdotische Beobachtungen in den Pinzgauer Gymnasien legten jedoch nahe, dass dieser Varietät des Salzburger Zentralraums im gymnasialen Schulalltag eine bedeutende Rolle zukommt.

## 8.4 Stichprobe

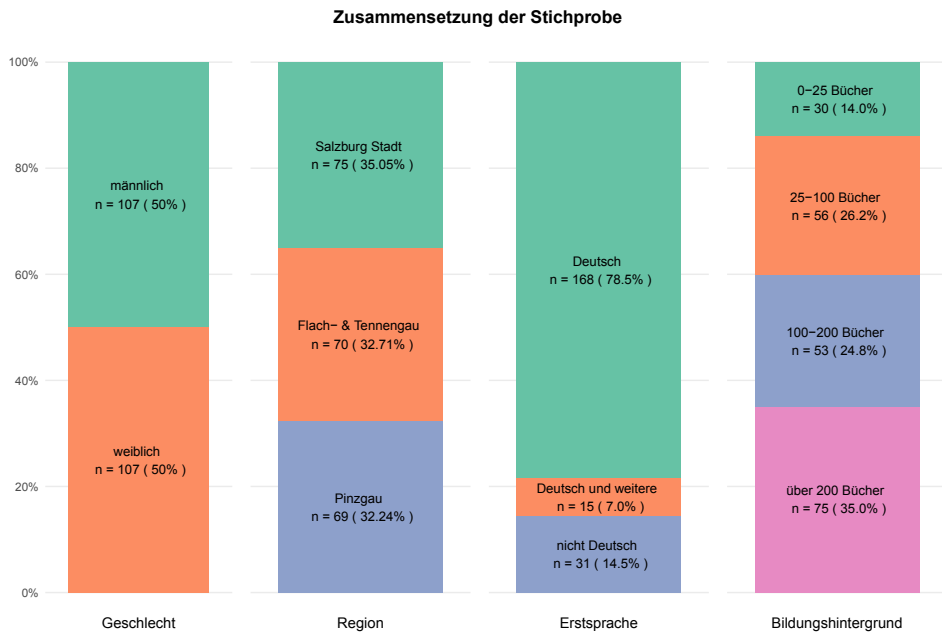
Insgesamt nahmen bei der Untersuchung 214 Schüler:innen mit einem durchschnittlichen Alter von 15,35 Jahren teil. Exakt die Hälfte der Stichprobe gab an, männlich und die andere Hälfte weiblich zu sein. Auch die Verteilung in den drei Regionen stellt sich als sehr ausgewogen dar. So nahmen 75 Jugendliche in der Stadt Salzburg, 70 im Flach- & Tennengau und 69 im Pinzgau an der Untersuchung teil. Große Unterschiede ergeben sich, wenn man die Erstsprache betrachtet. So gaben 168 an, Deutsch, 15 Deutsch und eine weitere und 31 eine andere Erstsprache als Deutsch zu besitzen.<sup>12</sup> Für die Erhebung des Bildungshintergrunds wurde die Frage, die in Abbildung 8.2 wiedergegeben ist, gestellt:

Wie viele Bücher gibt es bei dir zu Hause ungefähr? Zähle NICHT mit: Zeitschriften, Zeitungen und deine Schulbücher.

|  |  |  |  |   |
|--|--|--|--|---|
| O keine oder nur<br>sehr wenige (0-10<br>Bücher) | O genug, um ein<br>Regalbrett zu<br>füllen (11-25<br>Bücher) | O genug, um ein<br>Regal zu füllen (26-<br>100 Bücher) | O genug, um zwei<br>Regale zu füllen<br>(101-200 Bücher) | O genug, um drei oder<br>mehr Regale zu füllen<br>(über 200 Bücher) |
|--|--|--|--|---|

**Abbildung 8.2:** Frage zum Bildungshintergrund

<sup>12</sup> Die Frage im Fragebogen lautete: Was ist/sind deine Muttersprache/n? Es war möglich, mehrere Sprachen einzugeben. Für die statistische Analyse wurden diese Antworten aber zu den drei Gruppen vereinheitlicht.



**Abbildung 8.3:** Sozialdaten der Studienteilnehmer:innen in den Bereichen Geschlecht, Religion, Erstsprache und Bildungshintergrund

Diese Frage stammt aus großen Panelstudien der didaktischen Forschung und hat sich als guter Indikator zur Abfrage des Bildungshintergrunds erwiesen (z. B. Paulus 2009, Bos et al. 2003).<sup>13</sup> Wie bei der Stichprobe von Gymnasialschüler:innen der 10. Schulstufe zu erwarten war, gaben nur sehr wenige die ersten beiden Kategorien, also 0–10 und 11–25 Bücher, an. Um valide Ergebnisse für die statistische Analyse zu erhalten, wurden die beiden Gruppen zu einer einzigen zusammengefügt, die nun 30 Personen umfasst. 25–100 Bücher wurden von 56, 100–200 von 53 und über 200 Bücher von 75 Schüler:innen gewählt. Alle bisher referierten Ergebnisse sind in Abbildung 8.3 zusammengefasst.

## 8.5 Datenerhebung

Die Datenerhebung lief immer nach dem gleichen Schema ab: Die Umfrage wurde auf Tablets durchgeführt. Vor dem Aufbruch an eine Schule wurden die Tablets gewartet, was bedeutet, dass sie geladen und alle Umfragen neu heruntergeladen wurden. Dieser

<sup>13</sup> Dennoch muss man sich die Frage gefallen lassen, ob diese Abfrage heute noch zeitgemäß ist. Durch die Digitalisierung und die Möglichkeit, Medien digital verfügbar zu haben, ist die Menge der physischen Bücher im Haushalt viel weniger aussagekräftig als dies in früheren Zeiten der Fall war. Es würde sich aus meiner Sicht lohnen, sich darüber Gedanken zu machen, wie eine solche Abfrage für die heutige Zeit aussehen könnte.

Schritt randomisierte die einzelnen Blöcke auf jedem der 30 Tablets. Gleichzeitig wurden die Tablets synchronisiert, was dafür notwendig war, die Daten auf den für die Untersuchung angelegten Server zu übertragen. Zusätzlich mussten Teilnahmecodes für die einzelnen Schüler:innen erstellt, Schokolade gekauft und alles reisefertig verpackt werden. Da wir zwei Schulklassen pro Schule erhoben, häufig keine Zeit zwischen den Erhebungen lag und dies auch wegen fehlender Internetverbindung häufig gar nicht möglich gewesen wäre, verzichteten wir darauf, zwischen den Erhebungen der beiden Klassen zu randomisieren. Die Konsequenz davon ist, dass von den ca. 40–50 Schüler:innen einer Erhebungsfahrt 30 eine Umfrage mit voller Randomisierung erhielten und 10–20 eine, die in dieser Form schon einmal durchgeführt wurde. Daher handelt es sich genau genommen um eine Pseudorandomisierung der Umfrage.

Wir fuhren meistens zu zweit mit zwei Koffern zu den Schulen und gingen in oder nach der Pause mit der jeweiligen Lehrkraft in die zu erhebende Klasse. Nachdem wir uns kurz vorgestellt und dabei sehr oberflächlich beschrieben hatten, was der Zweck der Studie sei<sup>14</sup>, teilten wir die Tablets, die Kopfhörer und die Teilnahmecodes aus und leiteten die Schüler:innen technisch an, mit der Umfrage zu beginnen. Wir waren auch immer für Rückfragen oder technische Probleme anwesend, was beides regelmäßig vorkam. Die Umfrage dauerte ca. 20 Minuten und die Schüler:innen, die sie beendet hatten, gaben uns die Tablets ab und bekamen als kleines Dankeschön – wenn sie denn wollten – ein Stück Schokolade. Dabei verliefen Prä- und Posttestung sehr ähnlich. Die Posttestung war etwas kürzer, da die Schüler:innen ihre Sozialdaten nicht mehr angeben mussten, sie durften danach die Zettelchen mit ihren Teilnahmecodes vernichten und wir fragten die Interventionsgruppen nach Abschluss der Testung immer nach ihren Erfahrungen mit den Materialien. Bis zum Ende der Posttestung ließen wir sowohl die Lehrkräfte als auch die Schüler:innen im Unklaren über die Ziele der Befragung, danach beantworteten wir alle Fragen, die uns hierzu gestellt wurden. Am Ende der Untersuchung brachten wir die Koffer mit den Tablets ins Büro der Universität, wo sie dann wieder gewartet wurden.

## 8.6 Quantitative Auswertungsmethoden

Für die Auswertung quantitativer Daten sind Methoden aus der Statistik zentral. Dabei existieren sowohl unterschiedlichste Methoden, die man für Analysen heranziehen kann, als auch unterschiedliche Werkzeuge in Form von Computerprogrammen. Generell lassen sich statistische Auswertungsmethoden in zwei große Teilbereiche zerlegen: den deskriptiven und inferenzstatistischen Teil. In der deskriptiven Statistik werden dabei, wie der Name suggeriert, die gesammelten Daten beschrieben. Dabei sind je nach Struktur der erhobenen Daten – wichtig ist hierbei z. B. das Skalenniveau, das die Daten aufweisen – unterschiedliche Darstellungsweisen möglich, wobei manche mehr und manche weniger dafür geeignet sind, den Informationsgehalt in Form eines Dia-

---

<sup>14</sup> Wir sagten im Großen und Ganzen, dass uns die Meinungen der Schüler:innen zu verschiedenen Sprachformen interessieren und dass sie die Antworten ganz intuitiv geben sollten.

gramms darzustellen. Die hier analysierten Daten wurden in Form von siebenstufigen Likertskalen erhoben (s. Abb. 7.1), was eine gewisse Problematik der Skalenklassifizierung (Ordinal- oder Intervallskala) mit sich bringt. Dieses Problem ist nicht unwichtig und wird daher separat in Abschnitt 8.6 besprochen.

Die zweite Kategorie statistischer Methoden liegt in den inferenzstatistischen Berechnungen. Mit ihnen kann man über die reine Beschreibung der Daten hinweg unter Einhaltung bestimmter Voraussetzungen Schlussfolgerungen für die Gesamtpopulation treffen. Da die Wahl der Stichprobe für inferenzstatistische Aussagen zentral ist und diese in dieser Untersuchung in Form einer Clusterstichprobe vorliegt, muss man dies in den Berechnungen miteinbeziehen.

Die Matched-Guise-Blöcke, die die Grundlage der Analyse bilden, weisen komplexe interne Zusammenhänge auf (s. Abschnitt 8.2). Zur besseren Veranschaulichung werden diese in Abbildung 8.4 dargestellt.

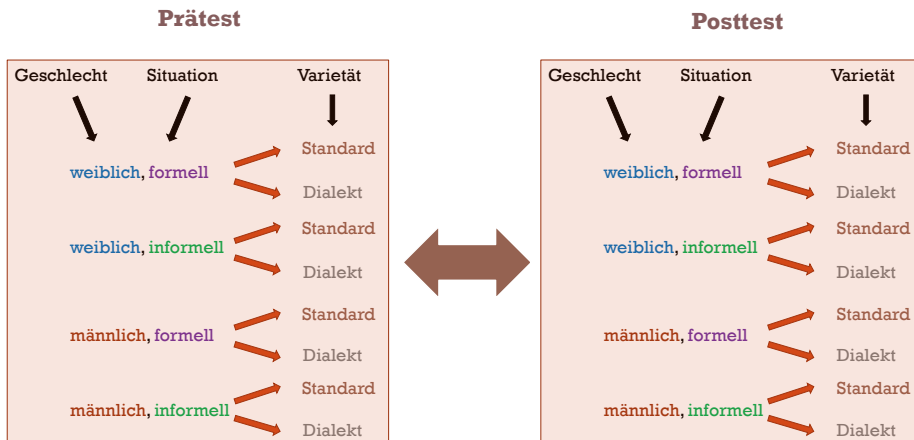


Abbildung 8.4: Testaufbau der Matched-Guise-Blöcke in der Interventionsstudie

Jeder der beiden Matched-Guise-Blöcke wurde von zwei Sprecherinnen und zwei Sprechern eingesprochen. Innerhalb der Geschlechter übernahm eine Person die Rolle in der formellen und eine die in der informellen Situation. Jede Rolle wurde von derselben Person jeweils im Dialekt und in Standarddeutsch eingesprochen. Die Sprachproben der Blöcke wurden ohne Veränderungen sowohl in der Prä- als auch in der Posttestsituation verwendet, einzig die Reihenfolge war durch die Randomisierung zufällig.

### 8.6.1 Deskriptive Statistik

Die Aufgabe der deskriptiven Statistik ist es, die Daten so informativ und einfach wie möglich darzustellen. In den Worten von Tuftle klingt das folgendermaßen:

Graphical excellence is the well-designed presentation of interesting data – a matter of substance, of statistics, and of design [...] [It] consists of complex ideas communicated with

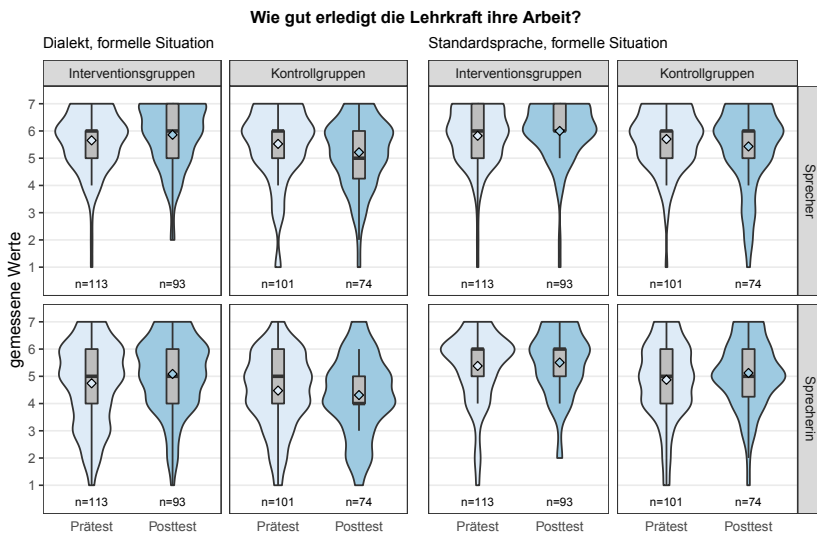
clarity, precision, and efficiency [...] [It] is that which gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space [...], [It] is nearly always multivariate [...]. And graphical excellence requires telling the truth about the data (Tuftle 1983: 51, übernommen aus Healy 2019)

Wenn man diese Ratschläge beherzigen will, merkt man, dass man mit konventionellen Programmen wie Microsoft Excel oder SPSS schnell an die Grenzen der Darstellungsmöglichkeiten stößt. Daher werden die Darstellungen (und in weiterer Folge auch alle anderen Berechnungen) mit dem Programm R (R-Core-Team 2017) erstellt, das mit dem Paket ggplot2 eine Vielzahl an Darstellungsmöglichkeiten für statistische Daten bietet. Mit diesem Paket wurden für die grafische Darstellung der deskriptiven Statistiken in dieser Arbeit primär drei unterschiedliche Diagrammtypen miteinander kombiniert, die die wichtigsten Lageparameter auf einen Blick abbilden.<sup>15</sup> Durch den Violinplot ist es möglich, die Varianz und Schiefe der Datenverteilung abzulesen. Ihm eingebettet befindet sich ein Boxplot, der die beiden Quartile, den Median und mit den Whiskers die eineinhalbfachen Interquartilsabstände abbildet. Zusätzlich wird das arithmetische Mittel mit dem Rautensymbol angezeigt. Mit der Zahl der Personenratings unter den Abbildungen wird den Darstellungen eine zusätzliche nützliche Information hinzugefügt.

Da es völlig unmöglich erscheint, alle einzelnen Variablenaspekte gleichzeitig darzustellen, werden die Daten in mehrere Bereiche unterteilt, die nacheinander analysiert werden. Zum einen wird hierfür die natürliche Trennung zwischen dem Lehrkräfte- und dem Schüler:innen-Block ausgenutzt. Zum anderen wird innerhalb eines Blocks die Analyse in die drei Ebenen, Kompetenzebene, Sympathieebene und Ebene der persönlichen Präferenz, und die beiden Situationen, formelle und informelle Situation, unterteilt. So bleiben die vier Variablen Gruppen, Geschlecht, Varietät und Erhebung übrig, die – ganz nach Tuftle – dem Lesenden eine große und übersichtliche Zahl an Informationen in kurzer Zeit und mit geringem Platzbedarf nahebringen. Abbildung 8.5 zeigt beispielsweise die deskriptive Statistik des Prä-Posttest-Vergleichs der Einschätzungen der Lehrkräfte in formeller Situation auf Kompetenzebene.

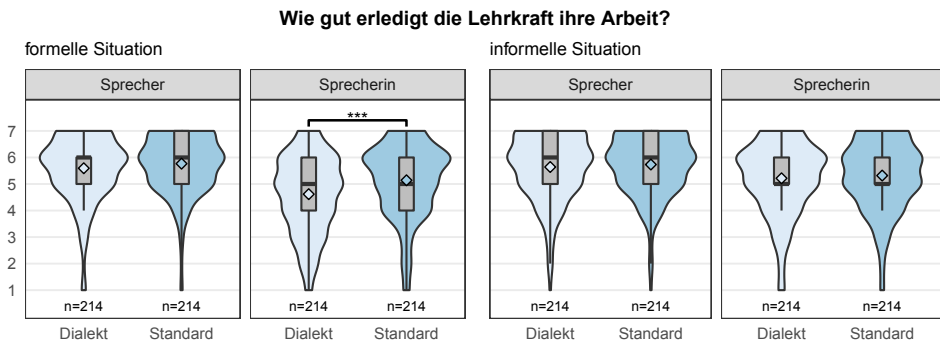
---

15 Ich habe mich hier von den Darstellungen von Maria Kliesch inspirieren lassen, die sie uns bei einem R-Workshop gezeigt hat und die in ähnlicher Form auf ihrer OSF-Seite (<https://osf.io/wcfj3/>) zu finden sind. Die Form der Kombination eines Boxplots, des Means und eines Violinplots wird etwa auch im R-Paket ggstatsplot realisiert.



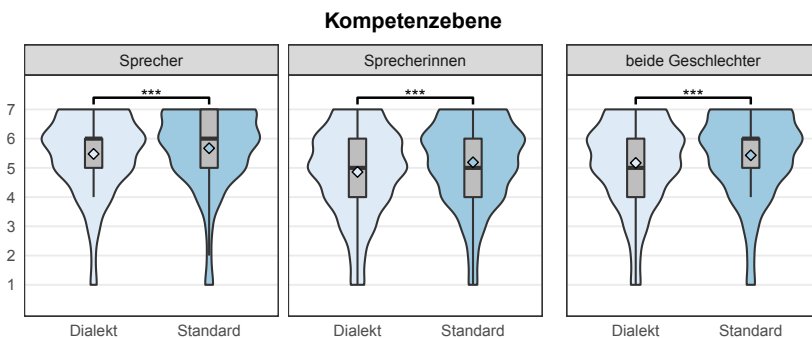
**Abbildung 8.5:** Prätest-Posttest-Vergleich der Einschätzung aller einzelnen Guises bei Frage nach der Kompetenz der Lehrkräfte in formeller Situation

Diese Abbildung stellt nun unterschiedliche Vergleichsmöglichkeiten bereit. Da der Hauptfokus auf dem Prä-Posttest-Vergleich liegt, sind in jedem der acht Kästchen die Erhebungszeitpunkte nebeneinander abgebildet. Die Dialektguises sind in den vier linken, die Standardguises in den vier rechten Kästchen, zwischen denen ein größerer Abstand besteht, dargestellt. Innerhalb der Vierergruppe besteht eine weitere vertikale Trennung zwischen den Gruppen – die Interventionsgruppen befinden sich links, die Kontrollgruppen rechts. Die horizontale Trennung unterteilt die Abbildung zusätzlich in die Geschlechter der Personen, die die Guises einsprachen – Sprecher oben, Sprecherinnen unten. Auf der Y-Achse ist die Skala mit eins bis sieben und nicht mit den Smileys des Fragebogens angegeben, was rein praktische Gründe hat. Diese Zahlen entsprachen der Kodierung der Skala in LimeSurvey und werden auch in den weiterführenden inferenzstatistischen Berechnungen verwendet. Die Zentrierung auf eine Skala von -3 bis +3 wäre ebenfalls möglich gewesen und wurde auch ausprobiert, machte die Interpretation der Ergebnisse aber nur unnötig kompliziert.



**Abbildung 8.6:** Dialekt-Standard-Vergleich der Einschätzung aller einzelnen Guises bei Frage nach der Kompetenz der Lehrkräfte

Zusätzlich werden in der Dissertation auch die Prätestergebnisse zur Darstellung der Ist-Situation der Spracheinstellungen in ähnlicher Form wiedergegeben. Da hier jedoch die Posttestergebnisse wegfallen und Interventions- sowie Kontrollgruppen gemeinsam analysiert werden, fällt die Darstellung etwas einfacher aus, wie in Abbildung 8.6 exemplarisch gezeigt wird. In dieser Abbildung liegt der Hauptfokus auf dem Dialekt-Standard-Vergleich der Varietäten, zusätzlich ist ihr ein Signifikanztest eingebettet. Dieser gehört jedoch schon zu den inferenzstatistischen Berechnungen und wird daher näher in Abschnitt 8.5.2 behandelt.



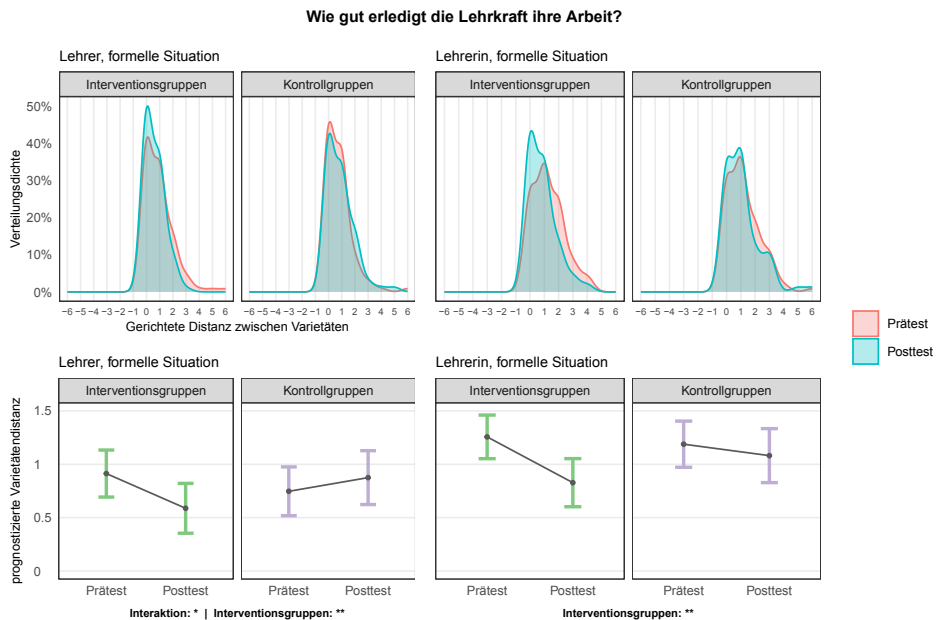
**Abbildung 8.7:** Dialekt-Standard-Vergleich der allgemeinen Einschätzung auf Kompetenzebene

Neben den Darstellungen der einzelnen Guises-Kategorien ist es ebenfalls aufschlussreich, generelle Ergebnisse zu betrachten. Dafür werden die beiden Rollen Lehrkräfte & Schüler:innen sowie die formelle & informelle Situation nicht mehr getrennt voneinander analysiert. Dies stellt sowohl bei den reinen Prätestergebnissen als auch bei den Prä-Posttest-Vergleichen eine Möglichkeit dar, allgemeine Aussagen über das Abstimmungsverhalten der Schüler:innen treffen zu können. In Abbildung 8.7 sind z. B. die Prätestergebnisse der Kompetenzebene abgebildet. Bei dieser Abbildung wird nun die Zählung der ratenden Personen unter den Diagrammen nicht mehr angezeigt.

Dies hat vor allem technische Gründe, da der entsprechende Befehl die Datenpunkte als Personen ansieht, in diesem generellen Ergebnis aber vier Datenpunkte pro Person zusammenkommen und daher der Personencount vervierfacht werden würde. Der Hauptfokus des Diagramms ist wiederum der Dialekt-Standard-Vergleich. In diesem speziellen Fall sind in den ersten beiden Kästchen die Ergebnisse der Sprecher und Sprecherinnen angegeben. Das dritte Kästchen stellt einen weiteren Generalisierungsschritt dar – hier wurden beide Geschlechter gemeinsam analysiert. Ebenfalls sind in diesen Generalisierungen die Ergebnisse von Signifikanztests zur vereinfachten Interpretierbarkeit eingebettet. Eine Aussage wie „Standardsprechende Personen werden im Prätest höchstsignifikant kompetenter eingeschätzt als dialektsprechende“ wird somit möglich.

Für die Beantwortung der zentralen Fragestellung dieser Arbeit, ob sich Sprach Einstellungen verändern lassen, ist noch eine weitere Darstellung der Prä- und Posttestdaten interessant. Betrachten wir dafür noch einmal Abbildung 8.5: Hier lässt sich erkennen, dass die dialektsprechende Lehrerin in den Interventionsgruppen im Posttest positiver als im Prätest bewertet wird, was ein Indiz dafür darstellt, dass sich die Sprach Einstellungen verändert haben. Was diese Abbildung jedoch nicht leisten kann, ist darzustellen, ob sich auch definitiv die Distanz in den Einstellungsbewertungen der einzelnen Schüler:innen innerhalb der Matched Guises der dialekt- und der standard-sprechenden Lehrerin verringert hat. Dies lässt sich mit einem Distanzplot zeigen. Hierfür müssen die Variablenwerte einer Varietät in einer bestimmten Situation von den Variablenwerten der anderen Varietät in der gleichen Situation subtrahiert werden, und zwar für alle Proband:innen. Bleiben wir bei vorliegendem Beispiel: Wenn man auf Kompetenzebene in formeller Situation sowohl in den Prätest- als auch den Posttestdaten die Ergebnisse der standardsprechenden von der dialektsprechenden Lehrerin abzieht, bekommt man die Bewertungsdistanz, die in Abbildung 8.6 auf der rechten, oberen Seite wiedergegeben wird. In der unteren Hälfte befinden sich inferenzstatistische Effektplots, denen wir uns in Abschnitt 8.5.2 widmen werden.

Auf der X-Achse der deskriptiven Plots befindet sich eine Skala von -6 bis +6. Diese fällt so aus, weil es sich um einen gerichteten Distanzplot handelt, der negative wie positive Distanzen enthält. Es wurde ja der standardsprachlichen Variable die dialektsprachliche abgezogen. Wenn also die Standardvariable positiver als die Dialektvariable geratet wurde, liegt das Ergebnis im positiven, wenn es sich umgekehrt verhält, im negativen Bereich. Sehen wir uns dazu den Bereich der Lehrerin in Abbildung 8.8 genauer an.



**Abbildung 8.8:** Prättest-Posttest-Vergleich der Distanz der Dialekt-Standard-Einschätzung bei der Frage nach der Kompetenz der Lehrkraft in formeller Situation

In der Prätistung (rötliche Linie) der Interventionsgruppe liegen die meisten Datenpunkte an Stelle + 1, was bedeutet, dass die Einschätzung der Kompetenz bei Verwendung der Standardvarietät um einen Punkt höher liegt als bei der Dialektvarietät. Die zweitmeisten Datenpunkte liegen auf 0, der Stelle der Gleichbewertung. Die drittmeisten liegen auf Stelle 2, dann kommen die Stellen 3 und 4, bis sich die Linie bei 5 und -1 bei null Datenpunkten einpendelt. Wie es zu erwarten war, wird die standardsprechende Lehrerin von den meisten kompetenter geratet als die dialektssprechende. Das Gegenteil, eine Besserbewertung des dialektsprachlichen Guises, kommt praktisch nicht vor. In der Posttestung (türkisfarbene Linie) verändert sich das Ratingverhalten merklich. Die Stelle der Gleichbewertung legt deutlich zu und die Stellen 2–4 nehmen ab, was einen generellen Rückgang der Bewertungsdistanz zwischen den Varietäten anzeigt.

Beide Formen der Darstellung lassen sich auch für den Vergleich zwischen den Geschlechtern heranziehen. Dieser Vergleich entspricht allerdings nicht mehr ganz der Grundidee von Matched Guise. Da Sprachproben von unterschiedlichen Personen verwendet werden, können auch Faktoren wie die Sympathie der Stimme der Sprechperson oder die unterschiedliche Erfahrung mit Deutschlehrerinnen und -lehrern eine Rolle für unterschiedliche Bewertungsmuster darstellen. Kausale Schlussfolgerungen aus den Ergebnissen abzuleiten, ist dadurch zusätzlich erschwert.

Das allgemeine Ziel der deskriptiven Statistik ist also, die Daten auf unterschiedliche Arten zu gruppieren und zu zählen, um sie miteinander vergleichen zu können –

theoretisch also eine recht simple Angelegenheit, die in der Praxis aber ihre Herausforderungen birgt. Als Grundlage für weiterführende inferenzstatistische Berechnungen ist sie zwingend notwendig, da dadurch eine Überprüfungsmöglichkeit bereitgestellt wird und besser beurteilt werden kann, ob diese Berechnungen plausibel sind.

### 8.6.2 Inferenzstatistik

Die Inferenzstatistik geht über die reine Beschreibung und Darstellung hinaus und hat zum Ziel, von der gezogenen Stichprobe Rückschlüsse auf die Gesamtpopulation zu ziehen. Dafür werden sogenannte Signifikanztests verwendet, die allesamt spezifische Voraussetzungen benötigen, um zu reliablen Ergebnissen kommen zu können. Nehmen wir einmal an, wir wollen das Abstimmungsverhalten zweier Gruppen miteinander vergleichen, z. B. der Interventionsgruppe und der Kontrollgruppe. Dafür müsste man sich eine bestimmte Situation aussuchen wie z. B. die aus Abbildung 8.5. Um einen einfachen Gruppenvergleich berechnen zu können, muss man jedoch die Gruppen noch genauer teilen, damit die unterschiedlichen Abstimmungsergebnisse einer Person nicht gemeinsam analysiert werden. Man müsste sich also noch für eine Varietät entscheiden, dann für ein Geschlecht und einen Testzeitpunkt. Wenn man sich nicht an diese Regel hält und z. B. die Daten des Sprechers und der Sprecherin in einer Gruppe zusammenfügt, verletzt man eine wichtige Regel, auf die man bei inferenzstatistischen Berechnungen achten muss, die Annahme der Unabhängigkeit (independence assumption). Dieses Vorgehen würde durch das „Gesetz der großen Zahlen“ die Typ-1-Fehlerrate in den Berechnungen stark erhöhen, was bedeutet, dass man statistisch signifikante Ergebnisse erhält, obwohl sie in der Grundgesamtheit nicht gegeben sind (s. für nähere Erklärungen B. Winter 2020: Kapitel 14.2).

Wenn nun die Gruppen, die man miteinander vergleichen kann, identifiziert sind, stehen einem unterschiedliche Tests zur Verfügung, die man heranziehen kann. Ein bekannter Test hierfür wäre der T-Test. Hat man einen Test im Blick, muss überprüft werden, ob die jeweiligen Testkriterien erfüllt werden. Bei einem T-Test wären dies folgende (entnommen aus Levshina 2015):

- The subjects have been sampled randomly.
- The data are at least interval-scaled.
- The differences between the pairs of scores (not the scores themselves!) are normally distributed, and/or the sample size is greater than 30.
- The variances in the underlying populations that represent two groups or conditions are equal.

Solch ein Vorgehen führt nun zu mehreren Problemen beim hier vorliegenden Studiensetting. Annahme 1 wird verletzt, da eine Klumpenstichprobe und keine Zufallsstichprobe vorliegt. Die Einhaltung oder Verletzung von Annahme 2 ist unklar, wird aber noch genauer in Kapitel Abschnitt 8.6 besprochen. Annahme 3 ist ein Problem, da die Residuen der vorliegenden Daten nicht normalverteilt sind. Das kommt daher, dass die Bewertungen, speziell zu den Lehrkräften, sich nicht neutral um die Mitte gruppieren, sondern sich im positiven Bereich befinden. Somit liegt Heteroskedastizität speziell in

den negativen Bereichen der Bewertungsskalen vor. Allerdings ist die Stichprobe relativ groß. Daher kann dieses Problem vernachlässigt werden. Wie man mit Annahme 4 umgehen soll, ist ebenfalls unklar. Die Varianzen der Stichproben, die durch den Violinplot angezeigt werden, sind häufig nicht perfekt normalverteilt und sehen auch nicht immer gleich aus. Es ist meiner Meinung nach schwierig zu entscheiden, wann man von ähnlichen und wann von unterschiedlichen Varianzen sprechen soll – der Übergang ist ja fließend.

Die technische Seite stellt jedoch nur eine Seite des Problems dar. Die andere ist, dass allein der Block in Abbildung 8.5 Unmengen an Gruppenvergleichen zulässt, was einerseits in kürzester Zeit zur Unüberschaubarkeit führt, andererseits auch ein weiteres statistisches Problem aufwirft – das Problem der Messwiederholungen. Dafür muss man verstehen, was ein  $p$ -Wert  $< 0.05$ , die klassische Schwelle, an der Signifikanz erreicht wird, eigentlich bedeutet. Wenn es sich um einen realen Effekt handelt, bedeutet es, dass in 19 von 20 Berechnungen die Ergebnisse der Realität entsprechen, sodass die Nullhypothese abgelehnt werden kann. In einem Fall jedoch begeht man einen Messfehler – in diesem Fall wäre das ein Typ-2-Fehler. Wenn man nun multiple Messungen an der gleichen Datengrundlage vornimmt, summieren sich die Messfehler und erhöhen die Chance auf Typ-1- oder Typ-2-Fehler – den Ergebnissen kann man also bei unveränderten  $p$ -Werten weniger trauen.

Was also tun? Die Antwort liegt darin, keine gewöhnlichen Signifikanztests durchzuführen, sondern ein Modell zu konstruieren, das den zugrundeliegenden Daten am ehesten entspricht. B. Winter (2020) empfiehlt sogar, für statistische Berechnungen Modelle Signifikanztests vorzuziehen, da diese flexibler auf die Datengrundlage anpassbar seien. Die hier vorliegende Datengrundlage in Form der Klumpenstichprobe schränkt die Möglichkeit von Modellen noch weiter ein. Diese Form der Stichprobe kann in gemischten linearen Modellen berücksichtigt werden. Diese Modelle funktionieren ähnlich einer klassischen Regressionsanalyse, besitzen jedoch den Vorteil, dass zusätzlich zu den fixen Effekten, die man berechnen will (also Response- und Prädiktorvariablen), Zufallseffekte berücksichtigt werden können. So ein Zufallseffekt kann nun das Clustering der Stichprobe von Schüler:innen in Schulklassen sein. In R existieren zwei Pakete, die für frequentistische Modellberechnungen von gemischten Modellen herangezogen werden können, *nlme* und *lme4*. Ich benutzte Letzteres für die hier durchgeführten Berechnungen (s. für die Anwendung und Funktion des Pakets Bates/Mächler/Bolker 2015). Damit kann zumindest die erste Voraussetzung erfüllt werden.

Im Modellierungsprozess gilt es nun, ein Modell zu finden, welches größere Erklärungskraft als das Nullmodell besitzt. Dafür muss man zuerst ein Nullmodell konstruieren, welches keine fixen Effekte enthält, und dieses dann mit einem Modell mit fixen Effekten vergleichen – soweit die Grundlagen. In der Praxis existieren aber prinzipiell zwei Wege, wie man diese Modellvergleiche angehen kann: der datengeleitete und der hypothesengeleitete Ansatz. Ersterer entspricht dem, dass man sich rein von der Passgenauigkeit und den  $p$ -Werten der unterschiedlichen Modelle leiten lässt; Zweiterer dem, dass man ein hypothesengeleitetes Modell konstruiert und sich weniger auf Passgenau-

igkeit und p-Werte einlässt. Ich wähle hier einen Mittelweg, konstruiere die Modelle also aufgrund der Hypothesen, achte aber zusätzlich auf die Passgenauigkeit und die p-Werte. Im Anschluss ist ein Beispiel abgebildet, wie so eine Modellberechnung in R funktioniert. Dabei werde ich wiederum auf Abbildung 8.5 zurückgreifen und ein Modell der dialektsprechenden Lehrerin entwickeln.

Als Erstes muss man ein Nullmodell berechnen, das die Zufallseffekte enthält. Diese Effekte müssen auch im Modell mit den fixen Effekten vorhanden sein, da sich ansonsten aufgrund der unterschiedlichen Freiheitsgrade Probleme bei der Vergleichbarkeit ergeben würden. Hypothesengeleitet sind folgende Zufallseffekte zu bedenken: Für die *Random Intercepts*, also für Unterschiede auf der Y-Achse der Regressionslinie, bieten sich die Proband:innen an, da man nicht davon ausgehen kann, dass alle Schüler:innen gleich abstimmen. Das Gleiche läge bei den Items vor – da jedoch nur ein Item bei diesem Modell betrachtet wird, sind Zufallseffekte in diesem Fall unnötig. Als Zweites ist der Clusteringeffekt, Schüler:innen in Schulklassen, interessant, da es sein kann, dass die Schulklassen einen Effekt auf das Abstimmungsverhalten haben. Dieser ist dann relevant, wenn sich die Ergebnisse innerhalb einer Klasse ähnlicher sind als zwischen den Klassen. Für die Steigungseffekte der Regressionslinie, also die *Random Slopes*, könnte der Erhebungszeitpunkt eine Rolle spielen, da nicht davon auszugehen ist, dass die Intervention auf alle Schüler:innen die gleichen Auswirkungen hat.

```
Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
method ['lmerModLmerTest']
Formula: Bewertung ~ 1 + (1 | Klasse/TNCode)
Data: LmD

      AIC      BIC   logLik deviance df.resid
  1238    1253    -615    1230      377

Scaled residuals:
   Min       1Q   Median       3Q      Max
-3.645 -0.508  0.140   0.714  1.848

Random effects:
 Groups             Name                Variance Std.Dev.
TNCode:Klasse (Intercept) 0.2641     0.514
Klasse          (Intercept) 0.0754     0.275
Residual                    1.1930     1.092
Number of obs: 381, groups: TNCode:Klasse, 214; Klasse, 12

Fixed effects:
              Estimate Std. Error    df t value      Pr(>|t|)
(Intercept)    5.582      0.104 10.940    53.7 0.000000000000013 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

**Abbildung 8.9:** Nullmodell eines linearen gemischten Modells mit Random Intercept „Proband:innen“ und Random Slope „Klasse“

Abbildung 8.9 zeigt ein Nullmodell, in dem die Rater:innen in den Schulklassen geclustert sind. An der Regressionstabelle lassen sich nun unterschiedliche Werte ablesen. Speziell drei Abteilungen sind interessant: Die Tabellen mit den Zufallseffekten und den fixen Effekten und die Zeile, die nach der Modellformel wiedergegeben ist. Diese

Zeile enthält unterschiedliche Informationskriterien, wie das *Akaike Information Criterion* (AIC) oder das *Bayesian Information Criterion* (BIC), die über die Passform des Modells Auskunft geben. Auch die Zeile unter den Zufallseffekten ist wichtig, da sie anzeigt, ob das Modell die einzelnen Gruppen richtig erkannt hat. In diesem Modell ist das der Fall. Die Nummer der Beobachtungen liegt bei 381, was eine geringere Zahl als im vorliegenden Datensatz ist. Dies erklärt sich folgendermaßen: Um ein Prä-Posttest-Modell rechnen zu können, eliminiert R alle Personen, die entweder beim Prä- oder beim Posttest gefehlt haben. Die Überprüfung mit dem Befehl `drop_na()` zeigt das gleiche Ergebnis. Der Terminus `TNCode:Klasse` bezieht sich auf die Gesamtanzahl der Schüler:innen, die zumindest bei einem der beiden Testzeitpunkte anwesend waren, und `Klasse` bezieht sich auf die Menge der Schulklassen, die bei der Untersuchung mitgemacht haben. Die Werte sind alle korrekt, was bedeutet, dass das Modell die Eingabe der Zufallseffekte richtig erkannt hat. Ein weiterer wichtiger Punkt, der aus Abbildung 8.9 nicht hervorgeht, ist, dass die Modellberechnung keine Fehler zeigt. Dies ist leider keine Seltenheit bei `lmer` und muss auf jeden Fall gelöst werden.<sup>16</sup>

Wenn das Nullmodell steht, muss man es in einem folgenden Schritt mit dem eigentlichen Modell, das die Prädiktoren enthält, vergleichen. Dies kann mit dem Befehl der Varianzanalyse `anova()` realisiert werden, der die Modelle miteinander vergleicht, wie dies in Abbildung 8.10 ersichtlich ist.

```
Models:
LmDnull: Bewertung ~ 1 + (1 | Klasse/TNCode)
LmDmod: Bewertung ~ Erhebung * Gruppen + (1 | Klasse/TNCode)
      npar  AIC   BIC logLik deviance Chisq Df Pr(>Chisq)
LmDnull    4 1238 1253   -615     1230
LmDmod     7 1235 1262   -610     1221  8.89  3      0.031 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

**Abbildung 8.10:** Modellvergleich zwischen Nullmodell und Modell mit fixen Effekten

Der Modellvergleich zeigt am einfachsten am Chi<sup>2</sup>-Test durch  $p < 0.05$ , dass das Modell mit den fixen Effekten dem Nullmodell überlegen ist. Auch die meisten anderen Werte liegen unter denen des Nullmodells, mit Ausnahme des BIC, der leicht darüber liegt. Nachdem die Passgenauigkeit des Modells auf die Daten überprüft wurde, kann man sich der Interpretation des eigentlichen Modells zuwenden, das in Abbildung 8.11 wiedergegeben ist.

<sup>16</sup> Sonderegger (2021) widmet diesen Problemen – zumeist handelt es sich um Singularitätsprobleme – einen eigenen Abschnitt und zeigt verschiedene Lösungswege auf.

```

Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
method ['lmerModLmerTest']
Formula: Bewertung ~ Erhebung * Gruppen + (1 | Klasse/TNCode)
Data: LmD

      AIC      BIC    logLik deviance df.resid
    1235     1262     -610     1221      374

Scaled residuals:
    Min       1Q   Median       3Q      Max
-3.654 -0.472  0.154   0.674  1.845

Random effects:
Groups             Name              Variance Std.Dev.
TNCode:Klasse (Intercept) 0.2817    0.531
Klasse           (Intercept) 0.0437    0.209
Residual                        1.1579    1.076
Number of obs: 381, groups: TNCode:Klasse, 214; Klasse, 12

Fixed effects:
              Estimate Std. Error      df t value
Pr(>|t|)
(Intercept)              5.656      0.142  16.190   39.84
<0.0000000000000002 ***
ErhebungPosttest              0.225      0.152  190.592    1.48
0.14
GruppenKontrollgruppe          -0.132      0.205   17.422   -0.65
0.53
ErhebungPosttest:GruppenKontrollgruppe -0.531      0.226  196.471   -2.35
0.02 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
      (Intr) ErhbnP GrppnK
ErhbnPstts -0.474
GrppnKnttrl -0.694  0.329
ErhbnPstts:GK 0.319 -0.673 -0.470

```

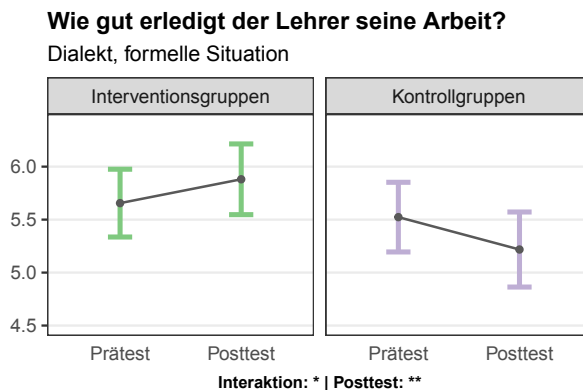
**Abbildung 8.11:** Lineares gemischtes Modell mit Random Intercept „Proband:innen“ und Random Slope „Klasse“ sowie den fixen Effekten Erhebungszeitpunkt (Prä- oder Posttest) und Gruppe (Interventions- oder Kontrollgruppe)

Wie ist nun aber das Gesamtmodell zu interpretieren? Sehen wir uns zuerst den Abschnitt mit den fixen Effekten an. Prinzipiell ist das Modell so konstruiert, dass es den Prä-Posttest-Vergleich mit gleichzeitiger Berücksichtigung (Interaktion) der Interventions- und Kontrollgruppe berechnet. In der kurzen Schreibweise von R wird das als *Erhebung\*Gruppen* realisiert. Das Modell zeigt vier Gruppenwerte, den *Intercept*, dann *ErhebungPosttest*, *GruppenKontrollgruppe* und *ErhebungPosttest:GruppenKontrollgruppe*, bei denen verschiedene Werte in mehreren Spalten angegeben werden.

Für den Anfang sehen wir uns die *Estimate*-Werte an, mit deren Hilfe man die vom Modell prognostizierten Haupteffekte berechnen kann. Den *Intercept*-Wert enthält jedes Modell und prinzipiell zeigt er die Abweichung des Schnittwertes der Regressionslinie auf der Y-Achse zu Null. Da es im Nullmodell nun keine fixen Effekte und damit auch keine Regressionslinie gibt, zeigt er die Gesamtheit der Werte und ihre Abweichung zu Null, was bedeutet, dass bei den dialektspredenden Lehrkräften in formeller Situation die Frage nach der Kompetenz vom Modell auf einen Wert von 5.56 prognostiziert wird. Da das Modell auf einer Skala von 1–7 gerechnet wird, bei der die neutrale

Bewertung<sup>17</sup> bei 4 liegt, ist die Berechnung des p-Werts allerdings nicht aussagekräftig.<sup>18</sup> Da dieses Modell jedoch nur zur Konstruktion des Modells mit den fixen Effekten dient, können wir auf diesen Wert verzichten.

Beim Modell mit den fixen Effekten bedeutet der *Intercept* nun etwas völlig anderes. Er bildet die Referenzgruppe ab, die man aus den anderen Gruppen erschließen muss. Dies ist in unserem Fall der Prätest der Interventionsgruppe, was man daran sehen kann, dass die zweite und dritte Gruppe in der Tabelle die Namen der gegenteiligen Gruppen Posttest und Kontrollgruppe enthalten. Der *Intercept*-Wert bedeutet also, dass das Abstimmungsverhalten der Interventionsgruppe im Prätest in dieser Situation vom Modell mit 5.66 prognostiziert wird. Da diese Gruppe der Referenzwert ist, sind alle anderen Werte von diesem abhängig. Der Wert *ErhebungPosttest* bedeutet also, dass die Interventionsgruppe im Posttest um 0.23 Punkte höher abgestimmt hat als im Prätest (ca. 5.88). Ähnlich ist auch der Wert von *GruppenKontrollgruppe* zu interpretieren. Die Kontrollgruppen haben im Prätest um 0.13 Punkte niedriger abgestimmt als die Interventionsgruppen (ca. 5.53). Am schwierigsten ist der 4. Wert der Interaktion *ErhebungPosttest:GruppenKontrollgruppe* zu lesen. Dieser bedeutet nämlich einen zusätzlichen Unterschied zwischen der Referenzgruppe und dem Posttest der Kontrollgruppe, der mit -0.53 klar negativ ist. Um also das Ergebnis des Posttests der Kontrollgruppe zu berechnen, muss man die Werte *Intercept*, *ErhebungPosttest*, *GruppenKontrollgruppe* und *ErhebungPosttest:GruppenKontrollgruppe* addieren, was ca. 5.22 ergibt.



**Abbildung 8.12:** Darstellung der Haupteffekte des linearen gemischten Modells

Alles in allem ist das Lesen und Interpretieren der fixen Effekte des Modells rein durch die Regressionstabelle möglich, wenn auch verwirrend und nicht gerade intuitiv. Ungleich einfacher wird es, wenn man die Haupteffekte grafisch darstellt, wie in Abbildung 8.12 ersichtlich.

17 Ob der Smiley mit dem geraden Strich als Mund wirklich als Punkt der neutralen Bewertung angesehen werden kann, ist im Grunde unklar.

18 Einen interpretierbaren p-Wert bekommt man, wenn man die Werte auf -3 bis +3 umskaliert.

Die vier Werte entsprechen den dunkelgrauen Punkten, die innerhalb der Interventions- und Kontrollgruppen zur Verdeutlichung mit einer Linie verbunden wurden. Zusätzlich ist die Grafik durch die Konfidenzintervalle erweitert, die sich aus dem Standardfehler (zweiter Wert in der Tabelle der fixen Effekte) berechnen lassen. Das Konfidenzintervall ist dabei ein günstiges Maß, um bei Gruppenvergleichen visuell einschätzen zu können, ob signifikante Unterschiede vorliegen oder eben nicht. Wenn sich die Balken zweier Gruppen nicht oder nur wenig überschneiden, kann man davon ausgehen, dass signifikante Unterschiede vorliegen – in unserem Beispiel wäre dies bei den beiden Gruppen im Posttest der Fall. Die Überprüfung mit einfacher linearer Regression ergibt einen Gruppenunterschied von  $p < 0.001$ , der Blick hat also nicht getäuscht. Was drücken aber nun die  $p$ -Werte, die in der Regressionstabelle unter *ErhebungPosttest:GruppenKontrollgruppe* zu finden sind, aus? „Posttest“ zeigt, dass vom Modell ein hochsignifikanter Unterschied der Posttestwerte beider Gruppen prognostiziert wird – die Konfidenzintervalle überlagern sich auch nur kaum. Die „Interaktion“ ist etwas schwieriger zu interpretieren. Sie gibt an, dass sich die Steigung der Regressionslinie in den Interventionsgruppen signifikant von der Steigungslinie der Kontrollgruppen unterscheidet. Anders gesagt bedeutet das in diesem Fall, dass die Interventionsgruppen im Posttest signifikant positiver abgestimmt haben, wenn man sie mit den Kontrollgruppen vergleicht, was nicht bedeutet, dass der Unterschied zwischen Prä- und Posttest innerhalb der Interventionsgruppen signifikant sein muss. Die Berechnung durch ein gemischtes Modell mit Klasse als Random *Intercept* zeigt auch keine signifikanten Unterschiede innerhalb der Interventionsgruppen – so wie dies die Konfidenzintervalle andeuten. Wie geht man damit nun bei der Interpretation der Ergebnisse um? Es gibt nämlich in vorliegendem Modell ein Problem, das nicht hinreichend gelöst werden kann: die unterschiedlichen Ergebnisse der Kontrollgruppen.

Das vorliegende Untersuchungsdesign besteht ja aus Interventions- und Kontrollgruppen, um sicherzustellen, dass ein messbarer Effekt auf die Intervention rückführbar ist und nicht auf etwaige andere Faktoren, wie es im Schulsetting z. B. andere Unterrichtsmaterialien/Unterrichtsinhalte darstellen würden. Der wünschenswerte Fall wäre nun, dass sich die Prä- und Posttestergebnisse der Kontrollgruppen ähneln, da diese ja keinen Unterricht mit den Interventionsmaterialien erhalten. Wie ist nun jedoch der auffällige Abfall der Bewertungen zwischen Prä- und Posttestung, wie sie in Abbildung 8.12 oder auch in Abbildung 8.5 erkennbar ist, erklärbar?

Hierfür gibt es zumindest zwei Erklärungen: Einerseits könnte dies auf den Effekt der sozialen Erwünschtheit zurückzuführen sein (s. z. B. Krumpal 2013). Speziell bei der Lehrkräftebewertung – aber auch bei der Bewertung der Schüler:innen – ist es denkbar, dass das Verlangen nach sozialer Zustimmung die Antworten in der ersten Testung positiver ausfallen lässt, als sie eigentlich waren. Dies würde ganz gut die auffällige allgemeine Tendenz der Schlechterbewertung der Kontrollgruppen im Posttest erklären (s. Kapitel 11). Andererseits könnte auch das abnehmende Interesse an der Durchführung eine Rolle spielen. Die Kontrollgruppen führten zwei Mal eine Testung durch, die 20–30 Minuten lang dauerte. Beim ersten Mal mag das ja noch durch das Novum der Situation, den Besuch von Universitätspersonal und eine Testung, die voll-

digital auf Tablets durchgeführt wurde, aufregend gewesen sein. Beim zweiten Mal war dieser Effekt bei vielen sicher schon verfliegen. Dass die Schlechterbewertung im Posttest der Kontrollgruppen auf andere Effekte wie andere Unterrichtsinhalte oder die Altersdifferenz der Proband:innen von vier bis fünf Monaten zurückgeführt werden kann, ist meiner Meinung nach unwahrscheinlich.

Welche Gründe auch ursächlich sind, für die Interpretation der Modelle hat dies zur Konsequenz, dass man die Prätestergebnisse generell infrage stellen muss, da diese offensichtlich verfälschenden Einflüssen unterliegen. Die einfachste Lösung für dieses Problem wäre, sich auf die Posttestergebnisse beider Versuchsgruppen zu fokussieren. In Abbildung 8.10 differieren auch speziell die Posttestergebnisse. Die Konfidenzintervalle der Prätests überlappen größtenteils, sowohl ein einfaches lineares Modell wie auch ein gemischtes mit Klasse als *Random Intercept* prognostizieren keine signifikanten Unterschiede. Wenn nun also die unbeeinflussten Prätestergebnisse beider Gruppen eher den Posttestergebnissen der Kontrollgruppen entsprechen würden, würde das auch bedeuten, dass man in den Interventionsgruppen sehr wohl eine signifikante Veränderung durch die Intervention vom Modell prognostiziert bekäme, wenn die Testergebnisse des Prätests eben unbeeinflusst und daher niedriger gewesen wären.

Es ist nicht ganz einfach, mit dieser Situation umzugehen. Überall dort, wo bei den Kontrollgruppen eine deutliche Schlechterbewertung von Prätest zu Posttest zu beobachten ist, kann aber davon ausgegangen werden, dass auch die Angaben der Interventionsgruppen im Prätest zu hoch ausfallen. Durch diese Situation sieht man jedoch auch, wie wichtig die Betrachtung der deskriptiven Daten ist und wie wenig Aussagekraft ein isoliert herausgegriffener statistischer Test potenziell hat. Im Beispiel von Abbildung 8.10 besitzt der statistische Signifikanzwert der Interaktion also folgende Bedeutung: Es wird ein signifikanter Unterschied prognostiziert, wenn man die Gruppen, Interventions- und Kontrollgruppen, sowie die Testzeitpunkte, Prä-Posttest, gleichzeitig modelliert. Die Nullhypothese einer Nichtveränderung (der Einstellungen zur Kompetenz des dialektsprechenden Lehrers in der formellen Situation) durch die Intervention kann also in diesem Setting abgelehnt werden. Auf den angegebenen Grad der Steigungslinien darf man sich jedoch aus diesen Gründen nicht verlassen. Die Steigung in den Interventionsgruppen ist wahrscheinlich steiler, die der Kontrollgruppen flacher, exakt feststellbar ist dies aus der Datengrundlage heraus meiner Meinung nach aber nicht.<sup>19</sup>

Die Spracheinstellungen des Prätests sind also nicht eindeutig darzustellen. Aber auch die Effektstärke ist etwas, was in dieser Studie nicht sinnvoll in Prozentzahlen ausgedrückt werden kann. An der Intervention haben sechs Schulklassen mit ca. 100 Schüler:innen teilgenommen. Die zehn Unterrichtseinheiten wurden von den regulären Deutschlehrkräften durchgeführt, die wir im Vorfeld darauf geschult hatten. Was

---

<sup>19</sup> Man könnte sich schon gewisse statistische Szenarien überlegen, z. B. könnte man das arithmetische Mittel des generellen Bewertungsunterschieds innerhalb der Testzeitpunkte der Kontrollgruppen berechnen und dieses dann den Prätestergebnissen abziehen. Das würde die Unterschiede in den Testzeitpunkten nivellieren, würde allerdings die vorliegenden Daten verfälschen. Ich habe mich gegen eine generelle Modifikation der Daten entschieden und werde auf die vorliegende Sachlage bei Bedarf in der Diskussion der Ergebnisse verweisen.

wir dadurch jedoch nicht kontrollieren konnten, ist, ob alle Module wirklich durchgeführt wurden und auch mit welchem Elan dies geschah. Hier gibt es sicherlich Variation unter den Lehrkräften. Sechs Personen sind nun aber zu wenig, um eine Prognose auf die Gesamt-Lehrerschaft geben zu können, daher ist es unklar, ob unsere Lehrkräfte sich besonders oder eher wenig bemüht haben. Der zweite Punkt, der die Effektstärke der Intervention schlecht einschätzbar macht, liegt in der Coronapandemie. Die Intervention wurde im Wintersemester 2021/22 durchgeführt, in dem es in Österreich sehr hohe Coronazahlen gab und die Schulen kurz davor waren, geschlossen zu werden. Die Lehrkräfte führten die Intervention zwar durch, allerdings stellte die Coronasituation einerseits eine Belastung dar, andererseits waren dadurch auch viele Fehlzeiten auf Seiten der Schüler:innen vorhanden. Meiner Einschätzung nach führen diese Umstände mit hoher Wahrscheinlichkeit zu einer schwächeren Effektstärke in der Studie, als sie durch die Materialien eigentlich erreichbar ist. Dies ist auch der Grund, warum ich in dieser Arbeit darauf verzichte, prozentuelle Ergebnisse zu präsentieren, da sie das Gesamtbild eher verzerren würden. Die visuelle Analyse ist bei dieser Ausgangslage viel besser dafür geeignet, Unterschiede zu erkennen und Ergebnisse darzustellen.

Abschließend in diesem Abschnitt möchte ich noch kurz auf die *Random Effects* eingehen. Sie stellen den zentralen Unterschied zwischen linearen Modellen und gemischten linearen Modellen dar und erfüllen eine praktische Funktion. Die Prognosen, die durch inferenzstatistische Methoden für die Gesamtbevölkerung angestellt werden können, folgen zumeist dem zentralen Grenzwertsatz. Zumindest im frequentistischen Setting werden hierzu überwiegend Verfahren angewendet, die sich einer Form des Bootstrappings bedienen (wie z. B. die Maximum Likelihood Estimation). Durch die *Random Effects* ist es nun möglich, das Bootstrapping manuell zu beeinflussen, da die *Random Intercepts* und *Random Slopes* im Grunde Bootstrapping-Kategorien entsprechen (für die Funktionsweise von gemischten linearen Modellen s. Oskolkov 2020a; für die händische Nachrechnung s. Oskolkov 2020b). Dies eröffnet ein ganz neues Feld an differenzierten Möglichkeiten, Hypothesen zu testen. Da ich in dieser Studie die *Random Effects* jedoch nur dazu einsetze, um Clusteringeffekte zu kontrollieren und Abhängigkeiten der Datenpunkte innerhalb der Stichprobe aufzulösen, ist eine weitere Erklärung der Funktionsweise nicht vonnöten. Bei weiterführendem Interesse sei an B. Winter (2020) als auch Sonderegger (2021) verwiesen, die hierzu detaillierte Erklärungen liefern.

### 8.6.3 Einzelne Schritte der Auswertung

Die interne Komplexität des Matched-Guise-Designs (s. Abbildung 8.4) stellt in der Analyse der Daten eine große Herausforderung dar, da es unzählige Möglichkeiten gibt, die Daten zu gruppieren. Für die Arbeit war es mir aber wichtig, ein einheitliches Vorgehen sowohl in der Analyse als auch in der Darstellung der Ergebnisse zu finden, um dem Leser/der Leserin die Arbeit so intuitiv wie möglich zugänglich zu machen. Die Analysekategorien (s. Kapitel 9 und 11), denen in dieser Arbeit nachgegangen wird, ergeben sich aus den Hypothesen und der daraus folgenden Erstellung des Frage-

bogens, wobei nicht alle möglichen Sozialdaten analysiert wurden – das hätte den Rahmen dieser Arbeit gesprengt. Zu Beginn der Arbeit wurden die beiden Datensätze aus der Prätestung und der Posttestung in einer Exceltabelle zusammengeführt. Es wurde darauf geachtet, so wenige Eingriffe wie möglich an den Rohdaten vorzunehmen, um größtmögliche Transparenz zu gewährleisten. Die einzige weitere Veränderung am Rohdatensatz, die nicht das Zusammenführen der Erhebungen betrifft, lag darin, manche Spaltennamen umzubenennen, um noch klarer zu sehen, um welches Item es sich handelt.

Der Datensatz wurde dann in R eingelesen und vor der Analyse wurden bestimmte Modifizierungen an ihm vorgenommen. Es wurde z. B. ein langer Datensatz erstellt, bei dem nicht wie in einem breiten Datensatz jede Person eine einzelne Zeile und viele Spalten für jeden Datenpunkt benötigt, sondern es viele Zeilen für die gleiche Person gibt, da alle relevanten Datenpunkte für die Analyse (in diesem Fall waren das die durch das Matched-Guise-Verfahren gewonnenen Spracheinstellungen) sich in einer einzelnen Spalte befinden. Dadurch wird es dann möglich, eigene Spalten für die oben genannten Kategorien zu erstellen (wie z. B. das Geschlecht der Sprechenden) und mit diesem neuen Datensatz die Ergebnisse einerseits grafisch ansprechend darzustellen, andererseits ist ein solcher Datensatz auch nötig, um gemischte lineare Modelle berechnen zu können.

Nachdem die Daten für die Analyse aufbereitet waren, war es nötig, sich einen Überblick über die Datenlage zu verschaffen. Daher wurde ein Programm erstellt, welches eine Analysekategorie so umfassend wie möglich darstellen kann, um auf einen Blick erkennen zu können, in welchen Bereichen sich Veränderungen zeigen und wo nicht. In diesen vielen Einzelanalysen, die in einem Diagramm zusammengeführt wurden, wurden auch automatisierte Signifikanztests eingebettet. Ich habe mich, da man nicht von normalverteilten Daten ausgehen kann, und wegen der Clusterstichproben für Kruskal-Wallis-Rangsummentests entschieden, bei den reinen Prätestergebnissen für unabhängige und bei den Prä-Posttest-Vergleichen für abhängige Stichproben. Durch diese erste explorative Analyse war es dann möglich, den Umfang der Analyseplots für diese Arbeit zu verringern, alles andere hätte den Rahmen immens gesprengt. Es werden also nur Analyseergebnisse präsentiert, in denen Unterschiede bzw. Veränderungen zumindest in der Nähe der Signifikanzgrenze liegen. Die Bereiche, die solche Veränderungen zeigten, wurden dann neu geplottet und mittels eines gemischten Modells analysiert. Nur bei bestimmten Prätestergebnissen, bei denen dies auch gut argumentierbar ist, wurden gleich die Kruskal-Wallis-Testergebnisse verwendet. Trotz dieser Reduktion der Ergebnisse sind es ca. 200 Diagramme geworden, die das Herzstück dieser Arbeit bilden.

## 8.7 Ordinal- oder Intervalldaten? Das Skalenproblem

Wenn Daten in der Form von Likert-Skalen erhoben und weiterverarbeitet werden, stolpert man in der statistischen Literatur regelmäßig über eine bestimmte Diskussion. Aus statistischer Sicht ist es nämlich recht umstritten, welcher Skalenkategorie Daten einer solchen Skala zugeordnet werden müssen<sup>20</sup>. Es stehen zwei Kategorisierungsmöglichkeiten zur Auswahl: die Ordinal- und die Intervallskala. Beide Datenskalen haben gemeinsam, dass eine natürliche Reihenfolge der einzelnen Werte identifiziert werden kann. Der Streitpunkt liegt darin, ob die Abstände zwischen den einzelnen Werten quantifiziert werden können, also als mathematisch gleiche Abstände gelten können, oder ob dies nicht möglich ist. Schulnoten wären ein typisches Beispiel einer Ordinalskala. Auch wenn in Österreich fünf Noten mit den mathematischen Zahlen Eins bis Fünf gegeben werden und dies zuerst einmal so wirkt, als wären die Abstände gleich, ist der Unterschied zwischen einem Vierer und einem Fünfer im Zeugnis nicht gleichzusetzen mit dem Unterschied zwischen einem Dreier und einem Vierer. Ob sich die Lage bei einer siebenteiligen Likertskala, die mit Smileys indexiert wurde und schon rein grafisch gleiche Abstände in den Bewertungskategorien aufzeigt, ebenso verhält, kann, wie ich finde, bezweifelt werden. Was also tun?

Eine Möglichkeit wäre, auf Nummer sicher zu gehen und ordinalskalierte Daten voranzusetzen, allerdings wird dadurch die Bandbreite der statistischen Tests, die man dann noch verwenden kann, stark eingeschränkt. Man könnte auch die Skala als Intervallskala einordnen, wie es häufig in der Linguistik und anderen Forschungsgegenständen geschieht. Dann wäre es möglich, gemischte Modelle, die mittlerweile den Goldstandard in den Berechnungen von Clusterstichproben darstellen, zu rechnen, allerdings bleibt der fahle Geschmack übrig, dass die Berechnungen eventuell nicht den realen Gegebenheiten entsprechen.

Daher habe ich mich dazu entschieden, zwei unterschiedliche Modelle mit der gleichen Datengrundlage zu rechnen, bei denen eines von einer zugrundeliegenden Ordinal- und eines von einer Intervallskala ausgeht. Im Anschluss vergleiche ich sie und wenn die Modelle ähnliche Ergebnisse liefern, sind die Modellberechnungen, die von intervallskalierten Daten ausgehen, validiert.<sup>21</sup>

In R existieren Pakete für beide Modelltypen. Die eine Modellfamilie ist in Abschnitt 8.5.2 schon des Öfteren vorgekommen. Es handelt sich eben um die gemischten linearen Modelle, wie z. B. in Abbildung 8.12 ersichtlich. Daneben existiert mit dem Paket *npard* eine Möglichkeit, nichtparametrische longitudinale Daten, die zumindest ordinalskaliert sein müssen, berechnen zu können. Dieses Modell kann Clusteringeffekte nicht miteinbeziehen und ist auch, was Interaktionen betrifft, nur eingeschränkt anwendbar.<sup>22</sup> Da das Modell in Abbildung 8.12 nur eine Interaktion enthält, kann es

<sup>20</sup> Einen, wenn auch recht einseitigen Überblick zu diesem Problem und statistischen Mythen, die sich um Likertskalen ranken, liefern Carifio/Perla (2007).

<sup>21</sup> Eine weitere Versicherung ist, dass man überprüft, ob die Modellplots den deskriptiven Plots entsprechen. Wenn zwischen diesen keine großen Abweichungen existieren, ist es plausibel, dass das Modell verlässliche Ergebnisse liefert.

<sup>22</sup> In einem *npard*-Modell kann maximal eine Interaktion modelliert werden, die lmer-Modelle sind hier unbeschränkt erweiterbar – ein weiterer entscheidender Vorteil dieses Modelltyps.

jedoch nachgerechnet werden. Die Ergebnisse des *npardL*D-Modells sind in Abbildung 8.13 aufgeführt.

```
Call:
Bewertung ~ Erhebung * Gruppen

Relative Treatment Effect (RTE):

Wald-Type Statistic (WTS):
      Statistic df p-value
Gruppen      5.9815  1 0.01446
Erhebung      0.0842  1 0.77163
Gruppen:Erhebung  7.7237  1 0.00545

ANOVA-Type Statistic (ATS):
      Statistic df p-value
Gruppen      5.9815  1 0.01446
Erhebung      0.0842  1 0.77163
Gruppen:Erhebung  7.7237  1 0.00545

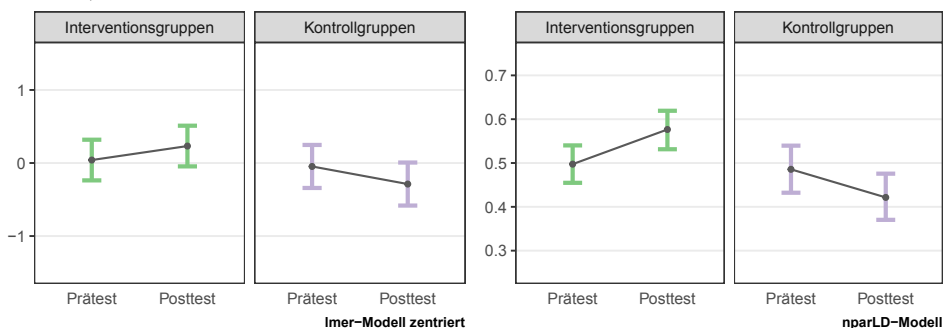
Modified ANOVA-Type Statistic for the Whole-Plot Factors:
      Statistic df1 df2 p-value
Gruppen      5.98  1 160 0.0155
```

**Abbildung 8.13:** Ergebnistabelle des *npardL*D-Modells

Das Modell zeigt zwei unterschiedliche Arten der statistischen Berechnungen. In der Dokumentation des Modells wird dazu geraten, den *ANOVA-Typ* dem *Wald-Typ* vorzuziehen. Da die Ergebnisse in dieser Berechnung jedoch identisch sind, müssen wir uns damit nicht näher befassen.

#### Wie gut erledigt der Lehrer seine Arbeit?

Dialekt, formelle Situation



**Abbildung 8.14:** Vergleich zwischen lmer- und *npardL*D-Modell

Die Modelltabelle ist anders zu interpretieren als die lmer-Tabelle. Sowohl Gruppen als auch Erhebung sind nicht in Relation zur Interaktion zu interpretieren. Es handelt sich hier um die absoluten Unterschiede. Die Signifikanz von Gruppen mit  $p < 0.05$  bedeutet, dass zwischen der Interventions- und Kontrollgruppe ein signifikanter Unterschied vorliegt, ebenso ist der Unterschied der Interaktion mit  $p < 0.01$  hochsignifikant. Beim Erhebungszeitpunkt sind keine Unterschiede erkennbar. Die Interpretation wird wie-

derum einfacher, wenn das Modell geplottet wird. In Abbildung 8.14 sind das *lmer*- sowie das *nparLD*-Modell zur besseren Vergleichbarkeit nebeneinander abgebildet.

Für die bessere Vergleichbarkeit wurde das *lmer*-Modell zentriert und neu skaliert. Das arithmetische Mittel liegt nun bei 0 und die Skalierung der Y-Achse ist nun auf Standardabweichungen umgerechnet. Das *nparLD*-Modell ist auf einer Skala von 0–1 um 0.5 zentriert<sup>23</sup>. Die Limits der Skala sind von 0.25 bis 0.75 gestellt und beim *lmer*-Modell auf 1.5 Standardabweichungen, was in etwa der gleichen Skalierung entspricht (innerhalb von drei Standardabweichungen befinden sich bei normalverteilten Daten 99.7 % des Datenbereichs).

Die visuelle Diagnose zeigt, dass sich die Modelle prinzipiell ähneln (ebenso wie der deskriptive Darstellung unter Abbildung 8.10). Daneben sind aber auch Unterschiede in den Modellen erkennbar. Vor allem sind die Steigungen der Regressionslinien im *nparLD*-Modell steiler. Dies müsste zu einem kleineren p-Wert führen, was man z. B. im Interaktionswert in Abbildung 8.11 auch sieht. Im Großen und Ganzen bedeutet dies, dass die *lmer*-Modelle ähnliche Ergebnisse erbringen, aber eher konservativ berechnet werden und daher zu Typ-1-Fehlern neigen. Daher macht dieser Umstand eine etwaige Bonferroni-Korrektur der Ergebnisse unnötig, da sich dadurch die Typ-1-Fehlerrate zusätzlich weiter erhöhen würde. Die Ergebnisse sind meiner Meinung nach jedoch ähnlich genug, um Likert-Skalen intervallskaliert berechnen zu können. Dezidierte Forschung in diese Richtung wäre jedoch wünschenswert.

## 8.8 Open Science Framework – ein Plädoyer

„Traue keiner Statistik, die du nicht selbst gefälscht hast!“ Wenn ich in einem Gespräch auf den Punkt zu sprechen komme, dass ich in meiner Arbeit statistische Methoden verwende, dann werde ich häufig mit einem solchen oder ähnlichen Zitat konfrontiert. Eine Google-Suche exakt dieses Zitats kommt zu 32.600 Ergebnissen. Dazu möchte ich primär zwei Punkte anmerken: Zum einen kann man den Vorwurf positiv sehen – gefälscht werden ja vor allem Sachen, die einen Wert besitzen. Dies zeigt, dass die Statistik als Methode implizit als wertvoll angesehen wird<sup>24</sup>. Und es ist auch kein Wunder: Die Medien und die öffentliche Meinungsbildung sind voll von Statistiken in allen erdenklichen Qualitätsstufen.

Zum anderen ist der generelle Vorwurf der Fälschbarkeit von Statistiken wahr. Wenn man weiß, was man tut, ist es ein Leichtes, die vorliegenden Daten zu manipulieren, um die Ergebnisse z. B. zu „frisieren“. Speziell wenn die Ergebnisse nicht die erwünschte Signifikanz haben, kann schon einmal ein solcher Gedanke aufkommen. Statistische Daten sind also prinzipiell fälschbar. Was bei diesem Zitat jedoch zumeist ausgelassen wird, ist die Tatsache, dass überall, wo der Mensch involviert ist, gefälscht

23 Ich habe vergeblich versucht, das *lmer*-Modell auf das gleiche Skalenniveau zu bringen, jedoch ist es mir nicht gelungen, gleichzeitig auf die Skala zwischen 0–1 umzukodieren und um den Wert 0.5 zu zentrieren.

24 Gleichzeitig werden die erwartbaren Ergebnisse auch als potenziell wahr angesehen, da es sich ja nicht lohnen würde, eine wissenschaftliche Methode zu fälschen, die unklaren Wahrheitswert besitzt.

und gelogen werden kann. Der Statistik kommt also nur insofern eine besondere Rolle im Spiel der Wahrheitsmanipulation zu, da sie als Methode omnipräsent ist und vielen Personen die statistische Minimalbasis fehlt, um das erkennen zu können.

Es gibt aber ein Mittel gegen potenzielle Manipulation: den transparenten Umgang mit den Daten und den verwendeten Analysemethoden. Dafür gibt es das Open Science Framework, das die Webseite <https://osf.io> betreibt. Auf dieser kann man sich gratis registrieren und seine wissenschaftlichen Ergebnisse (und noch wichtiger: den Weg dahin) der Öffentlichkeit zur Verfügung stellen. Im Idealfall findet man dort zum jeweiligen Projekt den Rohdatensatz und Code-Dokumente, in denen die Datentransformationen und statistischen Berechnungen vorgenommen und kommentiert werden. Dass dies mit Codezeilen passiert, ist wichtig, da sie vollständig nachvollziehbar, replizierbar und damit transparent sind. Je mehr Wissenschaftler:innen sich finden, die nach diesen Kriterien quantitative Forschung betreiben, desto eher werden solche Sätze wie zu Beginn des Abschnittes der Vergangenheit angehören. Idealerweise hat man, was den Einsatz von Methoden betrifft, als Wissenschaftler:in nichts zu verstecken.

Ich möchte daher als gutes Vorbild vorangehen. Da ich alle Auswertungen mit R durchführe, ist es relativ einfach, diese auf OSF hochzuladen – ich mache dies mit R-Markdown, da dieses Datenformat ermöglicht, interaktive PDF-Dokumente herzustellen, die die Analysegrafiken enthalten und in denen die Codezeilen, die dafür verwendet wurden, ein- und ausklappbar sind. Nur die primären Datenumformungen von breitem zu langem Datensatz usw. wurden in R-Scripts durchgeführt, da diese von den einzelnen Markdown-Files eingeladen werden können und man sich dadurch viele Codezeilen am Beginn der Dokumente spart.

Dieses Vorgehen hat drei Vorteile. Zum einen verwende ich OSF als elektronischen Anhang der Dissertation, die dadurch viel kompakter und schlanker werden kann. Zum anderen kann ich viele der technischen Details der Berechnungen und statistischen Tests im Fließtext der Analysekapitel ausklammern. Und zu guter Letzt macht es mein Vorgehen in den Auswertungen transparent und nachvollziehbar. Dieser Grad der Transparenz ist mit schriftlicher Nacherzählung des Analysevorgangs nicht zu erreichen.

Man findet die Code-Dokumente auf meiner OSF-Seite <https://osf.io/7mhwt/>. Zu jedem Kapitel, das statistische Auswertungen enthält, sind einzelne Code-Dokumente vorhanden und einsehbar. Darin sind zusätzlich die explorativen und nichtsignifikanten Ergebnisse enthalten.



## 9 Welche Spracheinstellungen bringen die Proband:innen mit?

Dieses Kapitel steht ganz im Zeichen der deskriptiven und inferenzstatistischen Ergebnisdarstellung ausgewählter Bereiche aus dem Fragebogen. Die Responsevariable bilden dabei die beiden Matched-Guise-Blöcke, die Audiobeispiele von Lehrkräften und Schüler:innen enthalten (s. für das Studiendesign Abschnitt 8.2). Diese Blöcke weisen hohe interne Komplexität auf, die sowohl in der deskriptiven als auch inferenzstatistischen Analyse berücksichtigt werden muss. Ein spezielles Problem sind potenzielle Abhängigkeiten von Datenpunkten innerhalb einzelner Analysekatoren, die bei inferenzstatistischen Analysen zu erhöhter Typ-1-Fehlerrate führen würden. Wie man dieses Problem mit gemischten Modellen lösen kann, wird in Abschnitt 8.5.2 näher ausgeführt.

Die Analyse in den einzelnen Abschnitten dieses Kapitels findet nach dem gleichen Schema statt, um für größere Übersichtlichkeit zu sorgen. Zu Beginn werden detaillierte Ergebnisse des Lehrkräfte- und Schüler:innenblocks in den drei Ebenen, den beiden Situationen und beiden Geschlechtern gegeben. Danach folgt ein Kapitel mit generellen Ergebnissen, in denen nicht mehr zwischen den Rollen (Schüler:innen und Lehrkräften) und den Situationen (formell und informell) unterschieden wird. Am Ende jedes Abschnitts werden die zentralen Erkenntnisse in einem Zwischenfazit zusammengefasst.

Die einzelnen Abschnitte behandeln Unterschiede in folgenden Bereichen:

1. Dialekt und Standarddeutsch
2. Sprecherinnen und Sprecher
3. Raterinnen und Rater
4. Städtische und ländliche Regionen
5. Erstsprache
6. Bildungshintergrund

Innerhalb der Blöcke werden pro Item Fragen auf Kompetenz-, Sympathieebene und Ebene der persönlichen Präferenz gestellt (s. auch Abschnitt 8.2), die auch in dieser Reihenfolge behandelt werden. Die Fragen selbst entsprechen den Überschriften der detaillierten Ergebnisplots. Sowohl für die Kompetenz- als auch die Sympathieebene wurden forschungsgeleitete Hypothesen formuliert (s. Abschnitt 8.1). Die Fragen auf Ebene der persönlichen Präferenz können nach persönlichem Background der Schüler:innen in beiden Varietäten positive oder auch negative Einschätzungen hervorruufen und dienen dazu, den Gesamteindruck der Untersuchung abzurunden.

Auch der Formalitätsgrad der Situationen unterscheidet sich in den Guises. Es wurden formelle und informelle Situationen konstruiert und angesprochen, die in der Analyse ebenfalls eine Rolle spielen.

Da eine gesamte Ergebnisdarstellung über alle Ebenen und Situationen zu viel Raum beanspruchen würde, werden in der Analyse die Bereiche dargestellt, in denen signifikante Unterschiede vorkommen oder sich die Ergebnisse in der Nähe der Signifikanzgrenze befinden. Somit wird die Analyse in jedem Kapitel zwar schematisch gleich angegangen, es kommen jedoch nicht überall alle Ebenen bzw. Situationen vor. Die gesamte Darstellung aller deskriptiven und inferenzstatistischen Berechnungen findet man auf meiner OSF-Seite (<https://osf.io/7mhw7/>) unter den Namen der einzelnen Abschnitte dieses Kapitels.

Um einen ersten, breiten Eindruck der vorhandenen Spracheinstellungen der Schüler:innen bekommen zu können, lohnt es sich, die Prätestdaten sowohl der Interventions- als auch der Kontrollgruppen gemeinsam zu betrachten. Dies ist statistisch gesehen auch problemlos möglich, da hier unterschiedliche Gruppen zusammengekommen werden, jede Person pro Ratingkategorie aber nur einen Datenpunkt aufweist. Damit können die Ergebnisse mit der vorliegenden Forschung (s. Kapitel 4 und 5) verglichen und diese durch neue Daten aktualisiert und ergänzt werden. Spracheinstellungsdaten aus der 10. Schulstufe einer allgemeinbildenden höheren Schule (AHS) sind meiner Kenntnis nach in Österreich bisher noch nicht erhoben worden (s. aber Fuchs/Elspaß 2019 für Daten aus der 10. Schulstufe einer berufsbildenden höheren Schule).

Die gemeinsame Analyse der Interventions- und Kontrollgruppe für die Ausgangslage der Spracheinstellungen birgt wie so viele Entscheidungen einen Vorteil und einen Nachteil. Der Vorteil liegt in der größeren und damit valideren Stichprobe. Der Nachteil liegt darin, dass im Prä-Posttest-Vergleich häufig nur die Interventionsgruppen miteinander verglichen werden, was eventuell zu etwas abweichenden Prätestergebnissen im Prä-Posttest-Vergleich führt.

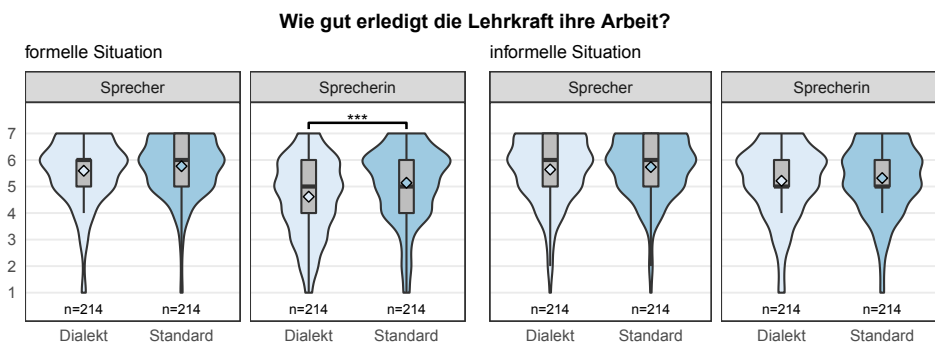
In diesem Kapitel wurden zwecks Übersichtlichkeit die Signifikanztests gleich in den deskriptiven Plots integriert. Die Darstellung wird der statistischen Konvention entsprechend mit Asterisken angegeben, die folgenden Schwellen entsprechen:

$p < .1 \triangle .$  |  $p < .05 \triangle *$  |  $p < .01 \triangle **$  |  $p < .001 \triangle ***$ . Für diese Arbeit bezeichne ich Ergebnisse mit  $p < .05$  als „signifikant“, Ergebnisse mit  $p < .01$  als „hochsignifikant“ und Ergebnisse mit  $p < .001$  als „höchstsignifikant“. Die detaillierten Ergebnisse wurden durch zweiseitige Wilcoxon-Rangsummentests für abhängige Stichproben, die generellen Ergebnisse durch gemischte lineare Modelle mit Proband und Item als Random Intercepts berechnet. Das Clustering in Schulklassen zeigte für die Berechnungen keine Vorteile.

## 9.1 Dialekt und Standarddeutsch

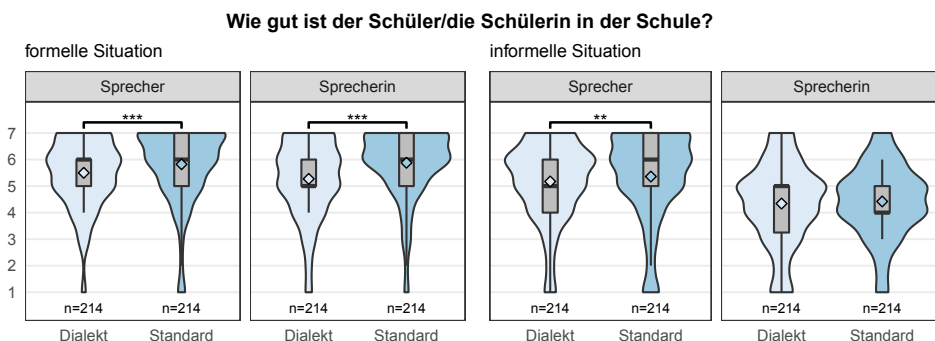
Das Untersuchungsdesign ist um die Guises-Paare Dialekt und Standardsprache, die von den gleichen Sprechern bzw. Sprecherinnen eingesprochen wurden, herum konstruiert. Personenbezogene Unterschiede in den Einschätzungen können somit ausgeschlossen werden. Daher stellt auch der Vergleich der beiden Varietäten den zentralsten Bereich der Studie dar. Alle folgenden Plots zeigen die Einschätzungen auf einer Ebene (Kompetenz-, Sympathieebene und Ebene der persönlichen Präferenz) sowie in einer Rolle (Lehrkräfte und Schüler:innen).

### 9.1.1 Detaillierte Ergebnisse



**Abbildung 9.1:** Dialekt-Standard-Vergleich bei der Frage nach der Kompetenz der Lehrkräfte

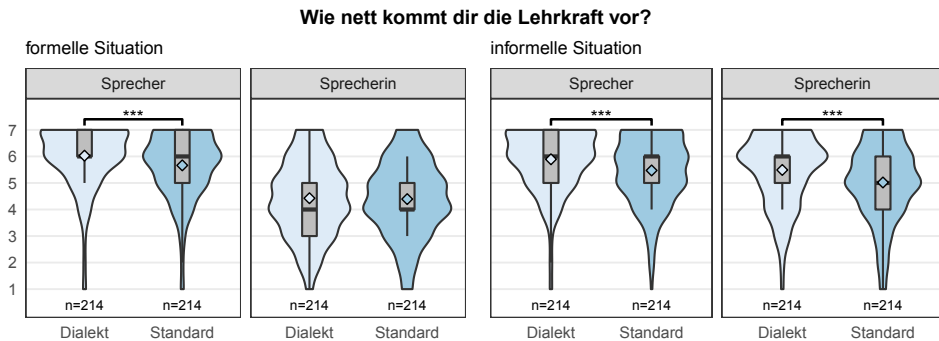
In Abbildung 9.1 sind die Einschätzungen zu den Lehrkräften auf Kompetenzebene abgebildet. Die standardsprechende Lehrerin wird in formeller Situation höchstsignifikant kompetenter eingeschätzt als in dialektaler Sprechweise. Alle anderen Gruppen zeigen keine signifikanten Unterschiede zwischen der Dialekt- und Standardbewertung, wobei die Einschätzung der Standardsprache immer leicht über der des Dialekts liegt.



**Abbildung 9.2:** Dialekt-Standard-Vergleich bei der Frage nach der Kompetenz der Schüler:innen

Bei den Schüler:innen findet man auf Kompetenzebene noch weitere signifikante Unterschiede (s. Abb. 9.2). In formeller Situation wird die standardsprechende Schülerin ebenfalls höchstsignifikant kompetenter eingeschätzt als die dialektsprechende. Dieses Ergebnis findet man nun jedoch auch beim Schüler, der ebenfalls bei standardsprachlicher Varietätenverwendung höchstsignifikant kompetenter eingeschätzt wird. Auch in informeller Situation wird der standardsprechende Schüler hochsignifikant kompetenter eingeschätzt. Bei der Schülerin in informeller Situation findet man zwar keinen Unterschied zwischen den Varietäten, allerdings fällt auf, dass sie deutlich negativer eingeschätzt wird als die Vergleichs-Guises.

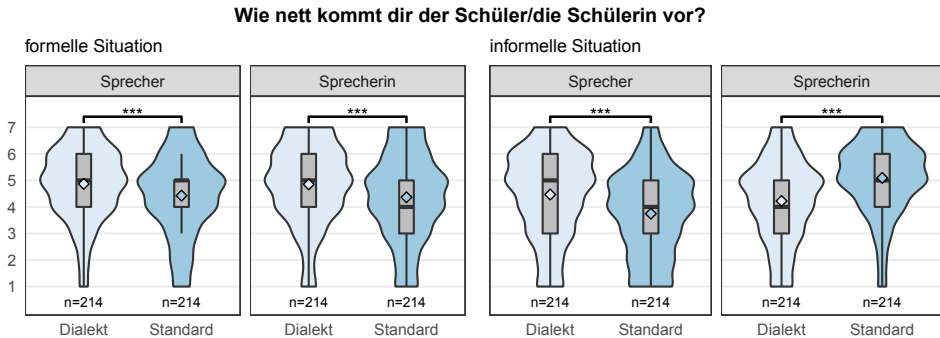
Auf Sympathieebene zeigt sich eine gegenteilige Einschätzung. Der dialektsprechende Lehrer (Abb. 9.3) wird in formeller Situation höchstsignifikant sympathischer eingeschätzt als in standardsprachlicher Sprechweise. Zum Verwechseln ähnlich sieht dessen Ergebnis auch in informeller Situation aus. In informeller Situation wird die dialektsprechende Lehrerin ebenfalls höchstsignifikant sympathischer eingeschätzt als in standardsprachlicher Sprechweise. Anders sieht es bei der Lehrerin in formeller Situation aus. Sie wird zwar in beiden Varietätenverwendungen ähnlich eingeschätzt, die Werte liegen aber deutlich unter denjenigen der anderen Gruppen. In formeller Situation wird die dialektsprechende Lehrerin damit auch nicht sympathischer eingeschätzt als die standardsprechende.



**Abbildung 9.3:** Dialekt-Standard-Vergleich bei der Frage nach der Sympathie der Lehrkräfte

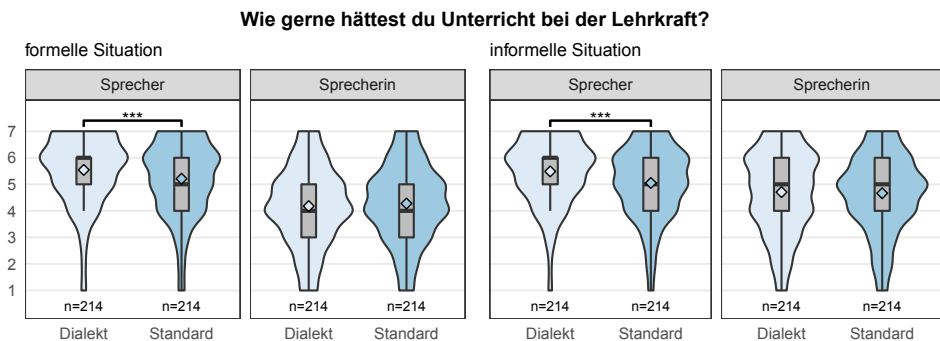
Bei den Schüler:innen findet man Ähnlichkeiten und Unterschiede zu den Lehrkräften auf Sympathieebene (Abb. 9.4). In formeller als auch in informeller Situation wird der dialektsprechende Schüler – wie auch der Lehrer – höchstsignifikant sympathischer eingeschätzt als in standardsprachlicher Sprechweise, wobei die Werte der informellen Situation unter denjenigen der formellen Situation liegen. Im Gegensatz zu den Lehrkräften wird die Schülerin in formeller Situation ähnlich eingeschätzt wie der Schüler – in dialektaler Sprechweise wird sie höchstsignifikant sympathischer eingeschätzt als in standardsprachlicher. Letztere Werte liegen allerdings etwas unter denjenigen des Schülers. Interessant wird es, wenn man sich die Einschätzungen der Schülerin in informeller Situation ansieht. Die standardsprechende Schülerin wird hier

nämlich höchstsignifikant und mit großem Abstand der Werte sympathischer eingeschätzt als die dialektsprechende. Diese Einschätzung bildet einen diametralen Unterschied zu den anderen Werten auf Sympathieebene.



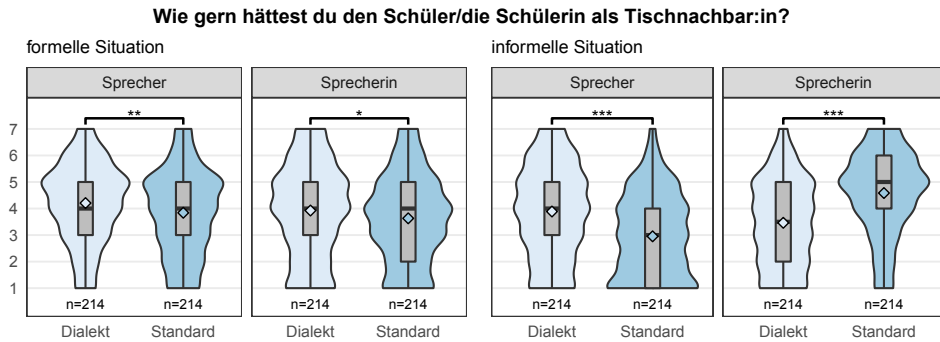
**Abbildung 9.4:** Dialekt-Standard-Vergleich bei der Frage nach der Sympathie der Schüler:innen

Auf Ebene der persönlichen Präferenz ähneln die Werte stark denjenigen der Sympathieebene. In formeller wie informeller Situation wird der dialektsprechende Lehrer höchstsignifikant positiver eingeschätzt als bei Verwendung des Standarddeutschen. Bei der Lehrerin finden sich keine signifikanten Unterschiede, allerdings wird sie deutlich niedriger eingeschätzt als der Lehrer und die Werte der formellen Situation liegen deutlich unter denjenigen der informellen (s. Abb. 9.5).



**Abbildung 9.5:** Dialekt-Standard-Vergleich bei der Frage nach persönlicher Präferenz der Lehrkräfte

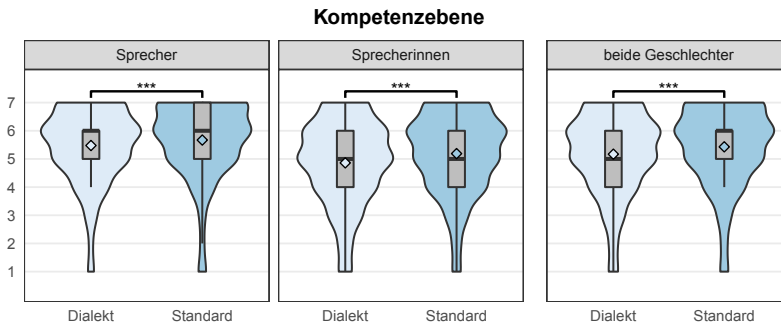
Auch der dialektsprechende Schüler wird in formeller Situation hochsignifikant und in informeller höchstsignifikant positiver eingeschätzt als in standardsprachlicher Sprechweise. In informeller Situation fällt auf, dass der Standard-Guise besonders negative Werte aufweist. Die dialektsprechende Schülerin wird in formeller Situation signifikant positiver eingeschätzt als die standardsprechende. In informeller Situation sind die Ergebnisse wiederum invertiert – die standardsprechende Schülerin wird höchstsignifikant und mit großem Abstand positiver eingeschätzt als die dialektsprechende (s. Abb. 9.6).



**Abbildung 9.6:** Dialekt-Standard-Vergleich bei der Frage nach persönlicher Präferenz der Schüler:innen

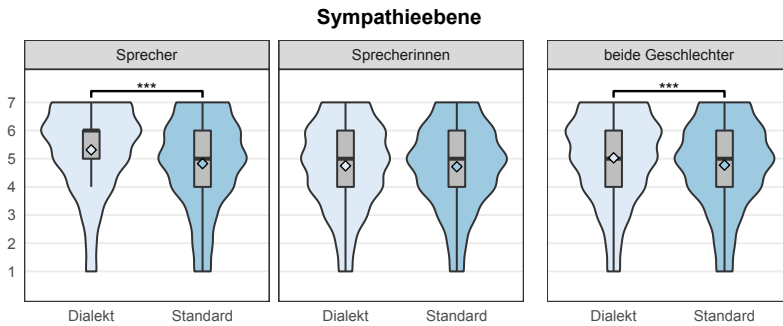
### 9.1.2 Generelle Ergebnisse

In den generellen Ergebnissen werden Lehrkräfte und Schüler:innen sowie formelle und informelle Situation gemeinsam analysiert (s. für nähere Erläuterung des Untersuchungsdesigns Abschnitte 8.5.1 und 8.5.2).



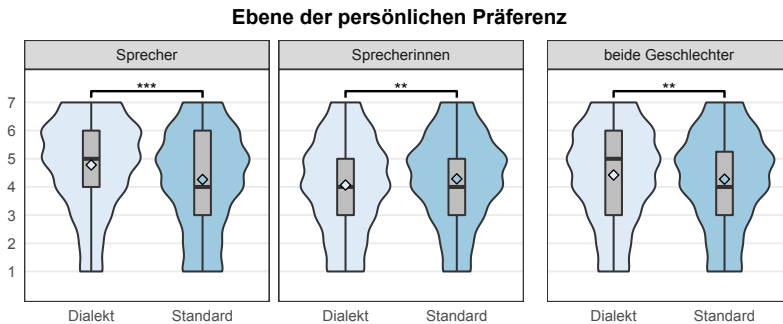
**Abbildung 9.7:** Dialekt-Standard-Vergleich der allgemeinen Einschätzung auf Kompetenzebene

Auf Kompetenzebene lässt sich erkennen, dass sowohl die Standardsprecher als auch die Standardsprecherinnen höchstsignifikant kompetenter eingeschätzt werden als in dialektaler Sprachverwendung, wobei sich die Werte der Sprecherinnen unter denjenigen der Sprecher befinden (s. Abb. 9.7). Es überrascht damit also nicht, dass auch das geschlechterindifferente Ergebnis höchstsignifikante Ergebnisse zugunsten der standardsprachlichen Sprechweise aufweist. Damit kann die Nullhypothese  $H_{1n}$  „Es gibt keinen Unterschied in der Kompetenzeinschätzung von standard- und dialektsprechenden Personen“ eindeutig abgelehnt werden.



**Abbildung 9.8:** Dialekt-Standard-Vergleich der allgemeinen Einschätzung auf Sympathieebene

Auf Sympathieebene (Abb. 9.8) werden die dialektsprechenden Sprecher höchstsignifikant sympathischer eingeschätzt als die standardsprechenden. Bei den Sprecherinnen liegen die Ergebnisse von dialekt- und standardsprachlicher Bewertung dicht beieinander und zeigen keine signifikanten Unterschiede. Dieses Ergebnis ist auch durch das divergierende Einschätzungsmuster der Schülerin in informeller Situation erklärbar. Es nivelliert sozusagen die positiveren Werte für die dialektale Sprechweise der anderen Gruppen. Wenn man die Werte beider Geschlechter gemeinsam analysiert, findet man, dass dialektale Sprechweise höchstsignifikant sympathischer eingeschätzt wird als standardsprachliche. Damit kann auch die Nullhypothese  $H_{2n}$  „Es gibt keinen Unterschied in der Sympathieeinschätzung von dialekt- und standardsprechenden Personen“ eindeutig abgelehnt werden.



**Abbildung 9.9:** Dialekt-Standard-Vergleich der allgemeinen Einschätzung der persönlichen Präferenz

Auf Ebene der persönlichen Präferenz (Abb. 9.9) werden Dialektsprecher wiederum höchstsignifikant positiver eingeschätzt als standardsprechende, wobei der Abstand zwischen den Werten sogar noch größer ist als auf Sympathieebene. Bei den Sprecherinnen findet sich nun das bemerkenswerte Ergebnis, dass standardsprechende Sprecherinnen hochsignifikant positiver eingeschätzt werden als dialektsprechende. Dieses Ergebnis zeigt die starke Beeinflussung der Einschätzung der Schülerinnen in informeller Situation, die nun sogar dafür sorgt, das generelle Ergebnis ebenfalls zugunsten

der positiveren Einschätzung des Standarddeutschen zu drehen. Bei Betrachtung beider Geschlechter wird dialektale Sprechweise hochsignifikant positiver eingeschätzt als standardsprachliche, da die Differenz der Werte bei den Sprechern auch höher ist als bei den Sprecherinnen.

### 9.1.3 Zwischenfazit Dialekt und Standarddeutsch

In der Betrachtung der unterschiedlichen Varietätenbewertungen zeigen sich zumeist erwartbare Ergebnisse. So wird in der Mehrzahl der verglichenen Guises standardsprachliche Sprechweise als kompetenter eingeschätzt und dialektale als sympathischer. Die beiden Nullhypothesen  $H_{1n}$  und  $H_{2n}$  lassen sich also verwerfen. Auffällig sind die divergierenden Ergebnisse für die Schülerin in informeller Situation auf Sympathieebene und Ebene der persönlichen Präferenz. So wird sie entgegen den anderen Werten dieser beiden Ebenen deutlich positiver eingeschätzt, wenn sie die Standardsprache verwendet. Auch die Lehrerin in formeller Situation wird in standardsprachlicher Sprechweise etwas sympathischer und auch allgemein etwas positiver eingeschätzt – es handelt sich wahrscheinlich also nicht um einen personenbezogenen Effekt oder eine Reaktion auf den Satzinhalt.

An diesem divergierenden Bewertungsunterschied innerhalb bestimmter Ebenen und Situationen, die in den generellsten Ergebnissen untergehen, sieht man, wie bedeutsam die Betrachtung und Auswertung von unterschiedlichen Ebenen, Situationen, Geschlechtern usw. ist.

Schüler:innen der 10. Schulstufe Gymnasium schätzen ...

1. standardsprachliche Sprechweise mit sehr hoher Wahrscheinlichkeit als kompetenter,
2. dialektale Sprechweise von Sprechern mit sehr hoher Wahrscheinlichkeit als sympathischer
3. und allgemein die dialektale Sprechweise von Sprechern mit sehr hoher Wahrscheinlichkeit positiver ein.

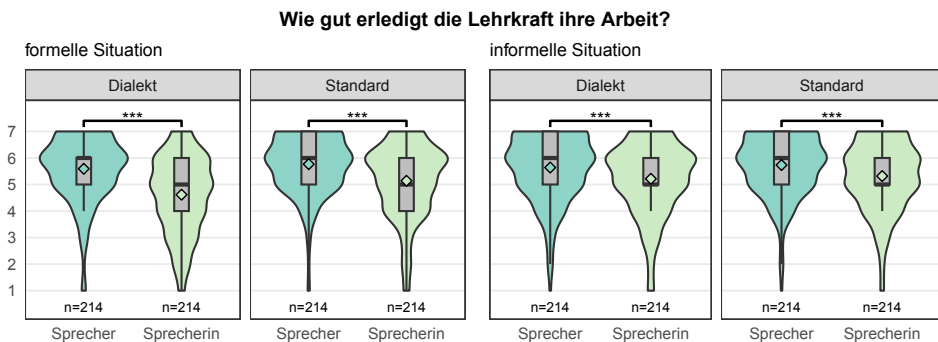
Die Ergebnisse zu den Sprecherinnen auf Sympathieebene und Ebene der persönlichen Präferenz sind dagegen nicht eindeutig. So wird die dialektsprechende Schülerin in formeller Situation auf beiden Ebenen und die dialektsprechende Lehrerin in informeller Situation auf Sympathieebene positiver eingeschätzt; bei der Lehrerin in formeller Situation in beiden Ebenen sowie auf Ebene der persönlichen Präferenz in informeller Situation werden die Varietäten ohne messbare Unterschiede bewertet und die dialektsprechende Schülerin wird auf beiden Ebenen mit großem Unterschied negativer eingeschätzt. Daraus kann gefolgert werden, dass dialektale Sprechweise bei Sprecherinnen nicht automatisch als die sympathischere angesehen wird. Dies scheint stark mit anderen Faktoren wie Situation oder sozialer Rolle in Verbindung zu stehen.

## 9.2 Sprecherinnen und Sprecher

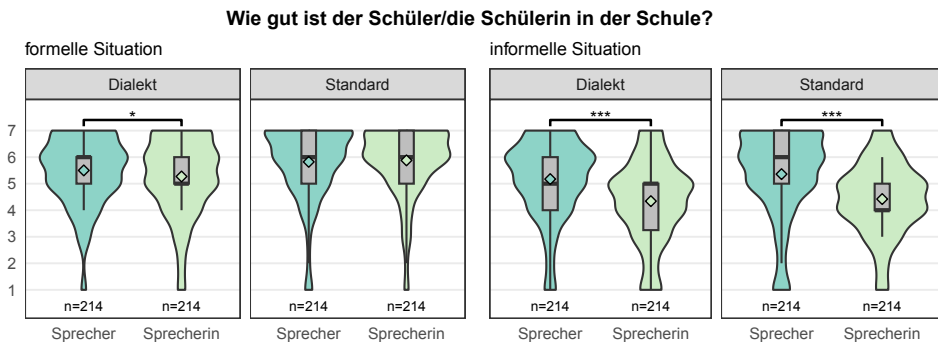
Neben der Gegenüberstellung der „echten“ Guises ist auch die Frage nach unterschiedlichen Spracheinstellungen bezüglich Sprecherinnen und Sprechern beachtenswert. In der Forschung wurden diesbezüglich schon mehrfach Unterschiede festgestellt (s. Abschnitt 4.2).

### 9.2.1 Detaillierte Ergebnisse

Abbildung 9.10 bestätigt dies auch sogleich. In beiden Varietäten sowie Situationen werden Lehrer höchstsignifikant kompetenter eingeschätzt als Lehrerinnen. Am ausgeprägtesten sind die Unterschiede in formeller Situation und dialektaler Sprechweise.



**Abbildung 9.10:** Sprecher-Sprecherin-Vergleich der Einschätzung aller einzelnen Guises bei der Frage nach der Kompetenz der Lehrkräfte

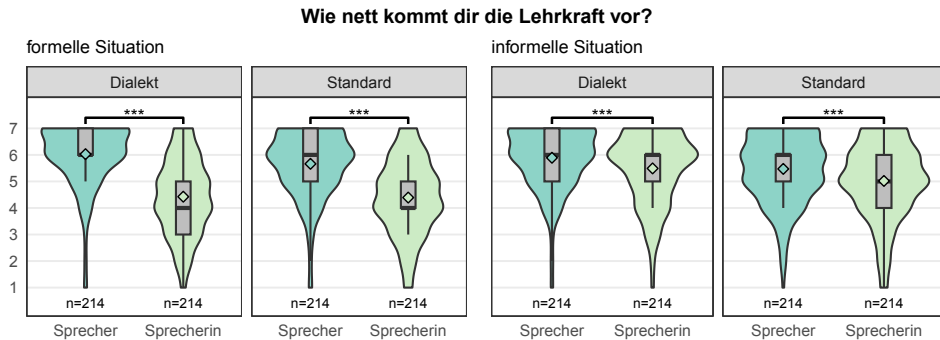


**Abbildung 9.11:** Sprecher-Sprecherin-Vergleich bei der Frage nach der Kompetenz der Schüler:innen

Auch die Schüler werden in informeller Situation in dialektalem und standardsprachlichen Varietätengebrauch höchstsignifikant kompetenter eingeschätzt als die Schülerinnen (Abb. 9.11), allerdings ändert sich die Sachlage in formeller Situation. Hier wird in dialektaler Sprechweise der Schüler signifikant kompetenter eingeschätzt als die Schü-

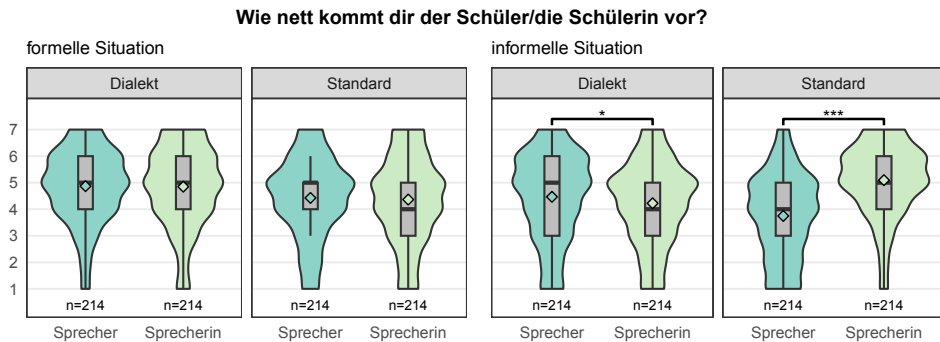
lerin, in standardsprachlicher Sprechweise sind jedoch keine Unterschiede zwischen den Geschlechtern der Sprechenden messbar – die Schülerin schneidet sogar etwas besser ab.

Auf Sympathieebene sieht man bei den Lehrkräften wiederum in beiden Varietätenverwendungen und Situationen, dass der Lehrer im Vergleich zur Lehrerin höchstsignifikant sympathischer eingeschätzt wird (s. Abb. 9.12).



**Abbildung 9.12:** Sprecher-Sprecherin-Vergleich bei der Frage nach der Sympathie der Lehrkräfte in der Prätest-situation

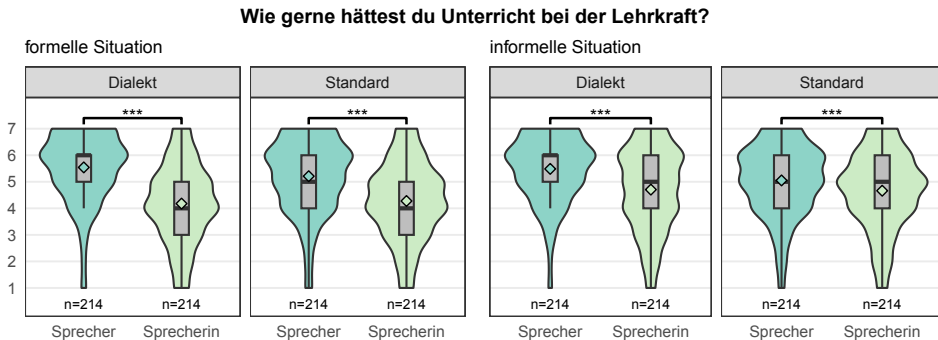
Allerdings sind auch hier die Abstände sehr unterschiedlich. Den größten Abstand findet man in formeller Situation und dialektaler Sprechweise. Aber auch der Unterschied in formeller Situation und standardsprachlicher Sprechweise ist vergleichsweise groß. Geringer werden die Abstände in informeller Situation, wobei standardsprachliche Sprechweise bei beiden Geschlechtern etwas weniger sympathisch eingeschätzt wird.



**Abbildung 9.13:** Sprecher-Sprecherin-Vergleich bei der Frage nach der Sympathie der Schüler:innen

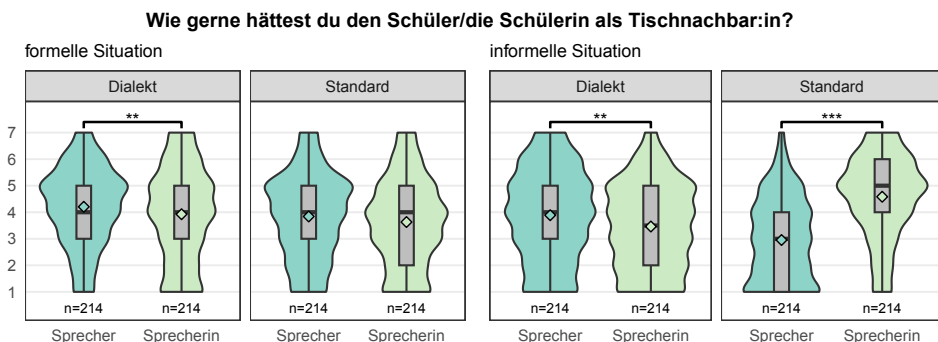
Bei den Schüler:innen ändert sich das Bild (s. Abb. 9.13). In formeller Situation sind in beiden Varietäten keine signifikanten Unterschiede messbar. Es fällt aber auf, dass dialektale Sprechweise bei beiden Geschlechtern sympathischer eingeschätzt wird. In informeller Situation wird der dialektsprechende Schüler signifikant sympathischer

eingeschätzt als die dialektsprechende Schülerin. Auffällig ist nun das Ergebnis, das schon in ähnlicher Form in Abschnitt 9.1 auftrat: In standardsprachlicher Sprechweise wird die Schülerin höchstsignifikant sympathischer eingeschätzt als der Schüler, und zwar mit großem Abstand der Werte. Gleichzeitig ist dieses Ergebnis der Schülerin auch das mit den höchsten Sympathiewerten bei Schülerinnen und Schülern. Dass die höchsten Sympathiewerte bei den Schüler:innen nun nicht bei Dialektverwendung zu finden sind, ist überaus bemerkenswert.



**Abbildung 9.14:** Sprecher-Sprecherin-Vergleich bei der Frage nach persönlicher Präferenz der Lehrkräfte

Auf Ebene der persönlichen Präferenz findet man wiederum, dass der Lehrer über die Varietäten und Situationen hinweg höchstsignifikant positiver eingeschätzt wird als die Lehrerin (Abb. 9.14). Wie schon auf Sympathieebene sind die Abstände der Werte in formeller Situation und dialektaler Sprechweise am größten und standardsprachlicher Sprechweise am zweitgrößten. In informeller Situation nehmen die Unterschiede etwas ab.



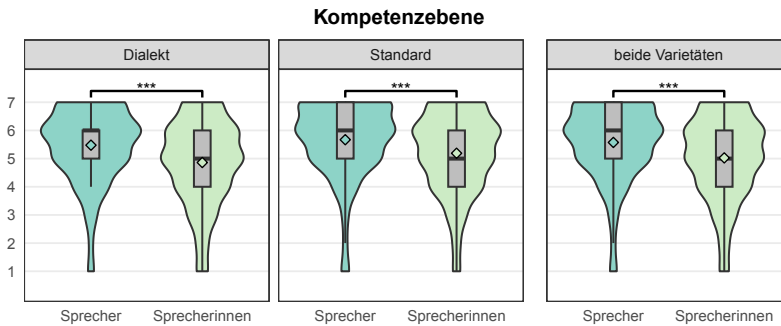
**Abbildung 9.15:** Sprecher-Sprecherin-Vergleich bei der Frage nach persönlicher Präferenz der Schüler:innen

Bei den Schüler:innen (Abb. 9.15) wird der dialektsprechende Schüler sowohl in formeller als auch informeller Situation hochsignifikant positiver bewertet als die Schülerin. In standardsprachlicher Sprechweise sind in formeller Situation keine signifikanten Unterschiede zu finden.

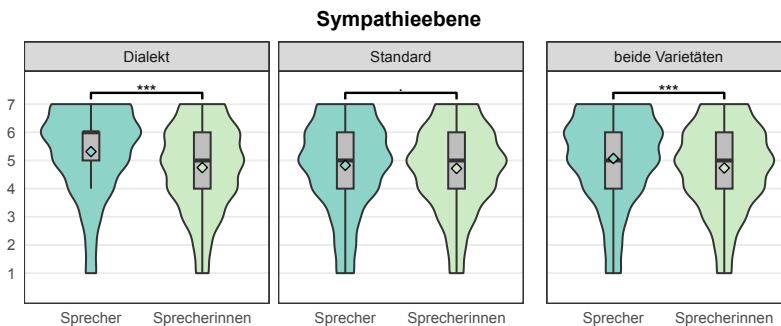
ten Unterschiede zwischen den Geschlechtern der Sprechenden auszumachen – die Werte des Schülers liegen jedoch etwas über denjenigen der Schülerin. In informeller Situation wird die standardsprechende Schülerin – ähnlich der Sympathieebene – höchstsignifikant positiver eingeschätzt als der Schüler. Auch der Abstand der Werte ist vergleichsweise groß. Dem standardsprechenden Schüler werden darüber hinaus sehr niedrige Werte zugeschrieben.

### 9.2.2 Generelle Ergebnisse

Wenn man die Unterscheidung zwischen formeller und informeller Situation sowie den Lehrkräften und Schüler:innen aufhebt, kommt man zu folgenden generellen Ergebnissen: Sprecher werden sowohl in dialektaler als auch standardsprachlicher Sprechweise höchstsignifikant kompetenter eingeschätzt als Sprecherinnen (s. Abb. 9.16). Dieses Ergebnis stellt sich auch ein, wenn beiden Varietäten gemeinsam analysiert werden.



**Abbildung 9.16:** Sprecher-Sprecherinnen-Vergleich der allgemeinen Einschätzung auf Kompetenzebene

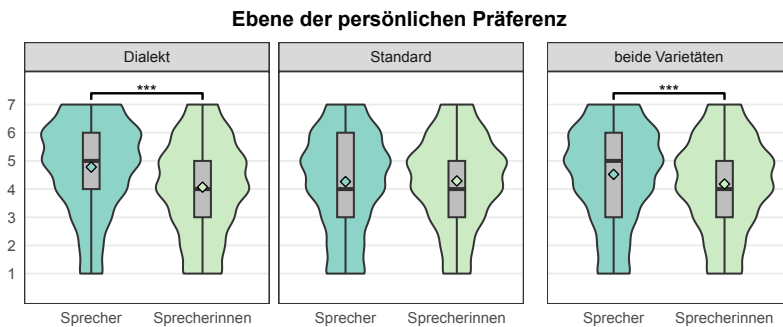


**Abbildung 9.17:** Sprecher-Sprecherinnen-Vergleich der allgemeinen Einschätzung auf Sympathieebene

Auf Sympathieebene sind nicht mehr alle Ergebnisse so deutlich. Dialektsprecher (Abb. 9.17) werden zwar ebenfalls höchstsignifikant sympathischer eingeschätzt als Dialektsprecherinnen. Bei standardsprachlicher Sprechweise werden die Sprecher zwar leicht sympathischer eingeschätzt, die Werte sind jedoch viel näher beieinander und nicht signifikant, wenn auch nahe der Signifikanzgrenze. Wenn man beide Varie-

täten gemeinsam analysiert, schlägt das Ergebnis der dialektalen Sprechweise (und ev. das Gesetz der großen Zahlen) durch. Sprecher werden also generell höchstsignifikant sympathischer eingeschätzt als Sprecherinnen.

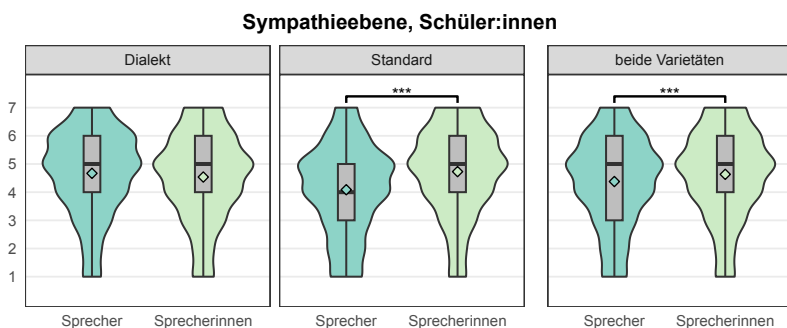
Dialektsprecher werden ebenfalls auf Ebene der persönlichen Präferenz höchstsignifikant positiver eingeschätzt als Dialektsprecherinnen (Abb. 9.18), mit großem Abstand der Werte. Bei Verwendung der Standardsprache sind keine Unterschiede messbar. Hier fällt sogar auf, dass die Sprecherinnen mit ihren Werten etwas über denjenigen der Sprecher liegen. Dieses Ergebnis wird jedoch bei gemeinsamer Analyse beider Varietäten vom Ergebnis der dialektalen Sprechweise geschluckt – Sprecher werden also unabhängig von der Sprechweise auf Ebene der persönlichen Präferenz positiver eingeschätzt als Sprecherinnen.



**Abbildung 9.18:** Sprecher-Sprecherinnen-Vergleich der allgemeinen Einschätzung der persönlichen Präferenz

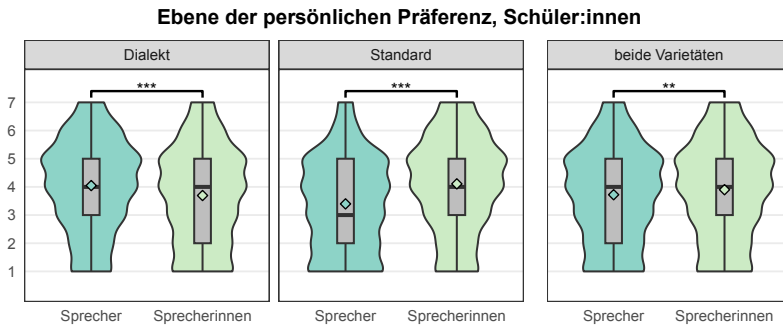
Im Geschlechtervergleich zwischen Sprecherinnen und Sprechern treten teils massive Unterschiede in den Spracheinstellungen der Rater:innen zu Tage. Speziell das Ergebnis bei den Lehrkräften ist erstaunlich, da in allen Kategorien der Lehrer höchstsignifikant positiver eingeschätzt wird als die Lehrerin. Dass es sich hier um keinen reinen Gendereffekt handeln kann, wird ersichtlich, wenn man diese Ergebnisse mit denjenigen der Schüler:innen-Guises vergleicht. Es müssen also noch andere Faktoren eine Rolle spielen. Aus emischer Situationsanalyse sind vor allem zwei Faktoren auffällig: Zum einen gibt es viel mehr Deutschlehrerinnen als Deutschlehrer. Die positivere Bewertung des Lehrers könnte damit auch weniger auf reale Erfahrung, sondern auf ein kognitives Idealbild eines imaginierten Deutschlehrers zurückzuführen sein. In den analysierten Klassen selbst ist das Verhältnis von Lehrerinnen zu Lehrern in den Interventionsgruppen vier zu zwei. Bei den Kontrollgruppen sind die Geschlechter der Deutsch-Lehrkräfte leider unbekannt. Zum anderen fiel in den Erhebungssituationen anekdotisch auf, dass die Deutschlehrerinnen in der Regel standardnäher sprechen als die Deutschlehrer, was mit den Forschungsergebnissen zusammenpassen würde, dass Frauen generell standardnäher sprechen (s. Abschnitt 4.2). Die Kombination beider Faktoren könnte mitverantwortlich für die überdeutlichen Bewertungsunterschiede zwischen der Lehrerin und dem Lehrer sein.

Bei den Schüler:innen-Guises findet man ebenfalls Bewertungsunterschiede. So wird der Schüler zumeist kompetenter beurteilt als die Schülerin, allerdings nicht in formeller Situation und standardsprachlicher Sprechweise. In dieser Situation, die mit Abstand die höchsten Kompetenzwerte einstreicht, wird die Schülerin sogar leicht kompetenter eingeschätzt. Die standardsprachliche Sprechweise scheint für die Schülerin jedoch noch weitere Vorteile auf Sympathieebene und Ebene der persönlichen Präferenz zu haben. Sehen wir uns das im Folgenden noch etwas genauer an. Abbildung 9.19 zeigt die Sympathieeinschätzungen für die Schülerinnen, ohne dass zwischen den Situationen unterschieden wurde. In dialektaler Sprechweise werden Schüler und Schülerinnen ähnlich sympathisch eingeschätzt, tendenziell sind die Schüler etwas im Vorteil. Bei standardsprachlicher Sprechweise werden nun aber die Schülerinnen deutlich sympathischer eingeschätzt als die Schüler. Interessant ist auch, dass die Werte der Schülerinnen sogar diejenigen der Dialekt-Guises sowohl der Schüler als auch der Schülerinnen übertrifft. Das Ergebnis der standardsprechenden Schüler ist dagegen deutlich negativer. Analysiert man nun beide Varietäten gemeinsam, sieht man, dass die Schülerinnen generell sympathischer eingeschätzt werden als die Schüler – das negative Ergebnis bei den Lehrkräften wird hierdurch zumindest teilweise relativiert.



**Abbildung 9.19:** Sprecher-Sprecherinnen-Vergleich der allgemeinen Einschätzung auf Sympathieebene, nur Schüler:innen

Auf Ebene der persönlichen Präferenz werden die dialektsprechenden Schüler höchstsignifikant positiver eingeschätzt als die dialektsprechenden Schülerinnen und die standardsprechenden Schülerinnen höchstsignifikant positiver als die standardsprechenden Schüler. Die Werte der standardsprechenden Schüler sind noch niedriger als auf Sympathieebene – ein standardsprechender Schüler ist somit die am wenigsten präferierte Wahl des Tischnachbars. Die Betrachtung beider Varietäten zeigt wiederum, dass Schülerinnen auf Ebene der persönlichen Präferenz hochsignifikant positiver eingeschätzt werden als Schüler. Die Abstände sind geringer und die Werte liegen generell unter denjenigen der Sympathieebene.



**Abbildung 9.20:** Sprecher-Sprecherinnen-Vergleich der allgemeinen Einschätzung der persönlichen Präferenz, nur Schüler:innen

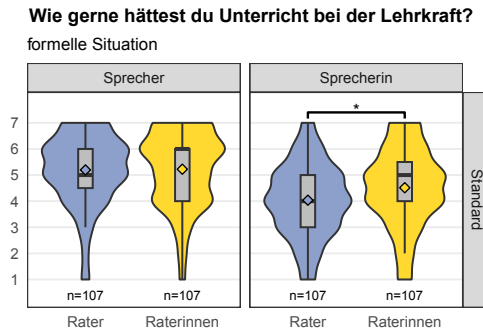
Abschließend ist es wichtig anzumerken, dass nicht mehr, wie bei „echtem“ Matched-Guise in den detaillierten Ergebnissen eine Person, die in unterschiedlichen Varietäten spricht, verglichen wird, sondern unterschiedliche Personen. Einflussfaktoren wie Sympathieeinschätzung der Stimme, Stimmlage, Satzinhalt usw. können bei diesen Vergleichen nicht ausgeschlossen werden.

## 9.3 Raterinnen und Rater

Wenn man Unterschiede zwischen den Geschlechtern betrachten will, ist nicht nur die Frage nach Bewertungsunterschieden, die die Sprechstimmen der Guises betreffen, interessant, sondern auch die nach Unterschieden zwischen den einschätzenden Schülerinnen und Schülern, die, um sie nicht mit den Sprechstimmen der Schüler:innen-Guises zu verwechseln, als Raterinnen und Rater bezeichnet werden. Da die Plots in diesem Kapitel nur die zwei Vergleiche zwischen Ratern und Raterinnen enthalten, wurden aus Platzgründen im Kapitel der generellen Ergebnisse mehrfach zwei Diagramme miteinander kombiniert.

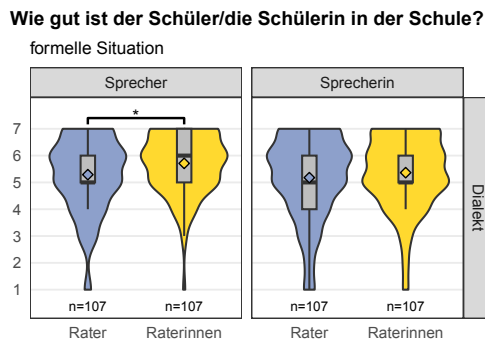
### 9.3.1 Detaillierte Ergebnisse

Bei den Lehrkräften findet man nur ein einzelnes signifikantes Ergebnis, und zwar auf Ebene der persönlichen Präferenz in formeller Situation (s. Abb. 9.21). Hier kann man erkennen, dass die standardsprechende Lehrerin von den Schülerinnen signifikant positiver eingeschätzt wird als von den Schülern. Die deskriptiven Plots zeigen, dass diese Tendenz auch beim standardsprechenden Lehrer vorliegt, wenn auch weniger ausgeprägt und nicht signifikant.



**Abbildung 9.21:** Rater-Raterinnen-Vergleich bei der Frage nach der Sympathie der standardsprechenden Lehrkräfte in formeller Situation

Die Betrachtung der Schüler:innen-Guises offenbart mehrere Unterschiede. In Abbildung 9.22 ist die Einschätzung zu den dialektsprechenden Schüler:innen in formeller Situation abgebildet. Der Schüler wird hier von den Raterinnen signifikant kompetenter eingeschätzt als von den Ratern – ein Ergebnis, das so nicht zu erwarten war. Allerdings ist es ebenfalls so, dass auch die Schülerin von den Raterinnen kompetenter eingeschätzt wird, wenn auch nicht im gleichen Ausmaß und nicht im signifikanten Bereich.

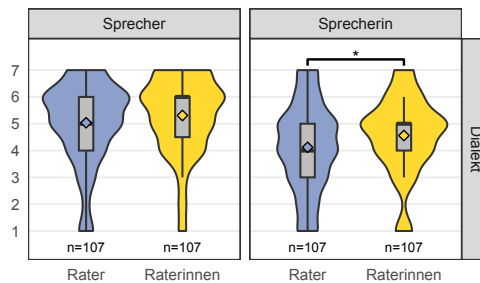


**Abbildung 9.22:** Rater-Raterinnen-Vergleich bei der Frage nach der Kompetenz der dialektsprechenden Schüler:innen in formeller Situation

In informeller Situation und dialektaler Sprachverwendung ist es nun die Schülerin, die von den Raterinnen signifikant kompetenter eingeschätzt wird als von den Ratern (Abb. 9.23). Auch hier ist die Tendenz beim Schüler-Guise ähnlich, wenn auch weniger ausgeprägt. Ebenso fällt auf, dass die Sprecher von Raterinnen wie Ratern die höheren Werte erhalten, also kompetenter eingeschätzt werden.

**Wie gut ist der Schüler/die Schülerin in der Schule?**

informelle Situation

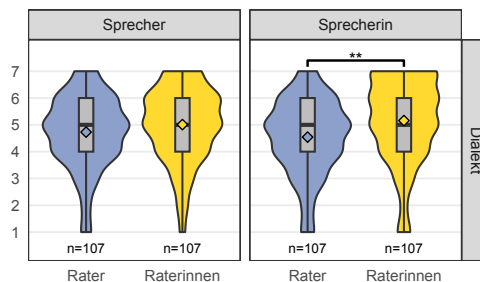


**Abbildung 9.23:** Rater-Raterinnen-Vergleich bei der Frage nach der Kompetenz der dialektprechenden Schüler:innen in informeller Situation

Auf Sympathieebene sind ebenfalls Unterschiede zwischen den ratenden Geschlechtern bemerkbar. In formeller Situation (Abb. 9.24) wird die dialektprechende Schülerin von den Raterinnen hochsignifikant sympathischer eingeschätzt. Die gleiche Einschätzungstendenz ist auch beim Schüler bemerkbar.

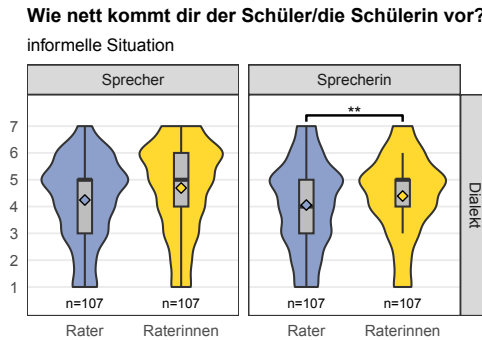
**Wie nett kommt dir der Schüler/die Schülerin vor?**

formelle Situation



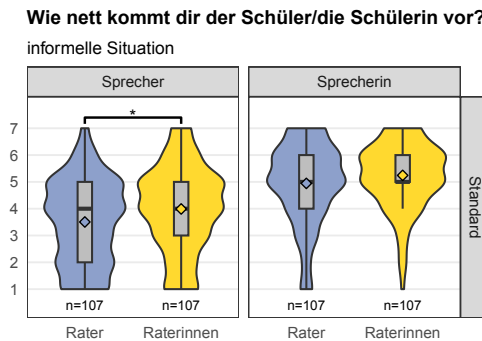
**Abbildung 9.24:** Rater-Raterinnen-Vergleich bei der Frage nach der Sympathie der dialektprechenden Schüler:innen in formeller Situation

In informeller Situation wird nun die dialektprechende Schülerin von den Raterinnen hochsignifikant sympathischer eingeschätzt als von den Ratern – wie schon bei allen vorhergehenden Plots ist die gleiche Tendenz auch beim Schüler-Guise erkennbar (s. Abb. 9.25).



**Abbildung 9.25:** Rater-Raterinnen-Vergleich bei der Frage nach der Sympathie der dialektsprechenden Schüler:innen in informeller Situation

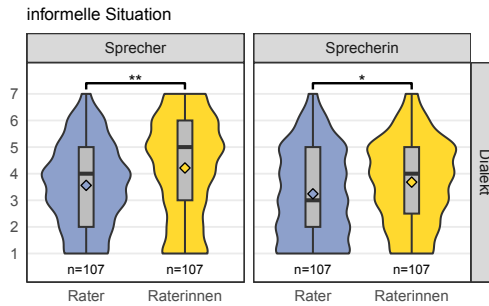
Auf Sympathieebene findet man auch den einzigen signifikanten Unterschied in standardsprachlicher Sprechweise. Wie man in Abbildung 9.26 auf der rechten Hälfte erkennen kann, wird der standardsprechende Schüler von den Raterinnen signifikant sympathischer eingeschätzt als von den Ratern – diese Tendenz findet sich auch beim Schülerinnen-Guise. In dieser Abbildung ist überdies hinaus gut zu erkennen, wie viel sympathischer standardsprechende Schülerinnen im Vergleich zu den Schülern eingeschätzt werden.



**Abbildung 9.26:** Rater-Raterinnen-Vergleich bei der Frage nach der Sympathie der standardsprechenden Schüler:innen in informeller Situation

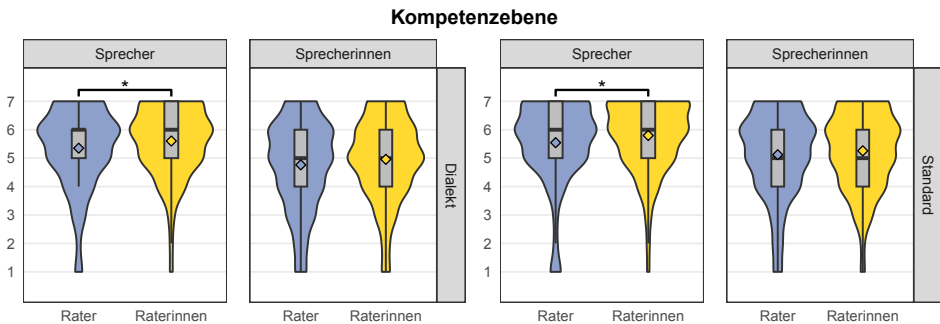
Auf Ebene der persönlichen Präferenz findet man in informeller Situation deutliche Einschätzungsunterschiede zwischen den Rater:innen. Der dialektprechende Schüler wird von den Raterinnen hochsignifikant positiver eingeschätzt als von den Ratern (s. Abb. 9.27). Ebenfalls wird die dialektprechende Schülerin von den Raterinnen signifikant positiver eingeschätzt als von den Ratern.

### Wie gerne hättest du den Schüler/die Schülerin als Tischnachbar:in?



**Abbildung 9.27:** Rater-Raterinnen-Vergleich bei der Frage nach persönlicher Präferenz der dialektsprechenden Schüler:innen in informeller Situation

In den unterschiedlichen Einschätzungen ist vor allem die Tendenz der ratenden Schüler bemerkbar, negativere Einschätzungen abzugeben. Um diesem Eindruck weiter nachgehen zu können, werden die Ergebnisse generalisiert analysiert, indem nicht mehr zwischen der formellen und informellen Situation sowie den Lehrkräften und Schüler:innen unterschieden wird.

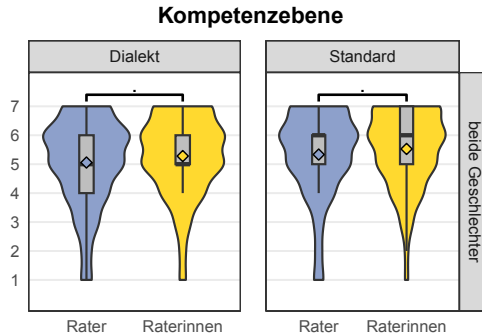


**Abbildung 9.28:** Rater-Raterinnen-Vergleich der allgemeinen Einschätzung auf Kompetenzebene, Dialekt- und Standardguises, getrennt nach Geschlechtern

Auf Kompetenzebene ist zu erkennen, dass der Dialekt Sprecher von den Raterinnen signifikant kompetenter eingeschätzt wird als von den Ratern (s. Abb. 9.28, linke Hälfte). Die gleiche Tendenz ist auch bei den Dialekt Sprecherinnen erkennbar, allerdings weit weniger ausgeprägt und nicht signifikant. Auffällig im Geschlechtervergleich ist, dass die Sprecher-Guises deutlich kompetenter bewertet werden als die der Sprecherinnen, und zwar von beiden Geschlechtern.

Bei standardsprachlicher Sprechweise ähnelt das Bild prinzipiell demjenigen der dialektalen Sprechweise. Auch hier werden die Sprecher von den Raterinnen signifikant kompetenter eingeschätzt als von den Ratern (Abb. 9.28, rechte Hälfte) und die

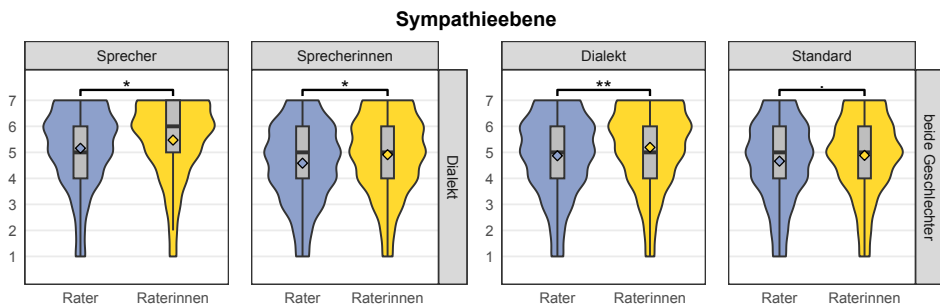
Sprecherinnen-Guises liegen mit ihren Werten deutlich unter denjenigen der Sprecher. Die Werte selbst sind aber in allen Einzelgruppen etwas höher angesiedelt als bei den vergleichbaren Gruppen der dialektsprachlichen Guises.



**Abbildung 9.29:** Rater-Raterinnen-Vergleich der allgemeinen Einschätzung auf Kompetenzebene, Dialekt- und Standardguises, beide Geschlechter

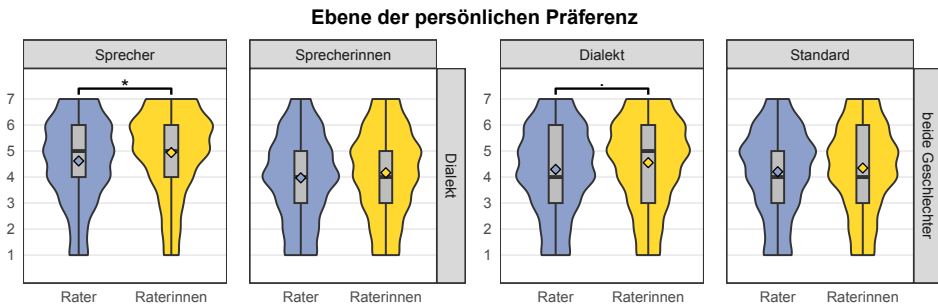
Wenn man nun zusätzlich die Differenzierung zwischen den Geschlechtern aufhebt, erhält man die Werte in Abbildung 9.29. Sowohl dialektale als auch standardsprachliche Sprechweise wird von den Raterinnen im Vergleich zu den Ratern positiver eingeschätzt. Die Unterschiede befinden sich jedoch nur nahe der Signifikanzgrenze. Die Personen mit standardsprachlicher Sprechweise werden von Ratern wie Raterinnen etwas kompetenter eingeschätzt.

Auf Sympathieebene werden in dialektaler Sprechweise sowohl die Sprecher als auch die Sprecherinnen von den Raterinnen signifikant sympathischer eingeschätzt als von den Ratern (s. Abb. 9.30, linke Hälfte). Die Sympathiewerte der Sprecher liegen deutlich über denjenigen der Sprecherinnen. In standardsprachlicher Sprechweise schätzen die Raterinnen zwar ebenfalls etwas positiver ein als die Rater, die Abstände sind jedoch viel geringer und es handelt sich nicht um signifikante Unterschiede.



**Abbildung 9.30:** Rater-Raterinnen-Vergleich der allgemeinen Einschätzung auf Sympathieebene, Dialektguises getrennt nach Geschlechtern auf der linken Seite und beide Geschlechter getrennt nach Sprachverwendung auf der rechten Seite

Die Ergebnisse werden noch deutlicher, wenn man die Geschlechterebene der Guises nicht mehr differenziert (s. Abb. 9.30, rechte Hälfte). Dialektale Sprechweise schätzen die Raterinnen hochsignifikant sympathischer ein als die Rater, standardsprachliche Sprechweise schätzen sie nur etwas sympathischer ein, die Ergebnisse befinden sich hier nur in der Nähe der Signifikanzgrenze. Daher kann auch  $H_{4n}$  (Es gibt keinen Unterschied in der Sympathieeinschätzung standardsprechender Personen zwischen Schülerinnen und Schülern.) nicht abgelehnt werden.



**Abbildung 9.31:** Rater-Raterinnen-Vergleich der allgemeinen Einschätzung der persönlichen Präferenz, Dialektguises getrennt nach Geschlechtern auf der linken Seite und beide Geschlechter getrennt nach Sprachverwendung auf der rechten Seite

Auf Ebene der persönlichen Präferenz setzen sich die Tendenzen, die schon auf Sympathieebene erkennbar waren, fort (s. Abb. 9.31, linke Hälfte). Die dialektsprechenden Sprecher werden von den Raterinnen signifikant positiver eingeschätzt als von den Ratern. Bei den dialektsprechenden Sprecherinnen ist die Tendenz ähnlich, allerdings schwächer ausgeprägt und damit nicht signifikant. Die Werte der Sprecherinnen liegen deutlich unter denjenigen der Sprecher.

Auch in standardsprachlicher Sprechweise sind diese Tendenzen erkennbar, jedoch nicht im Signifikanzbereich. Die rechte Hälfte von Abbildung 9.31 zeigt die generellen geschlechterindifferenten Ergebnisse auf Ebene der persönlichen Präferenz. Dialektale Sprechweise wird von den Raterinnen positiver eingeschätzt, allerdings befindet sich der Unterschied nur in der Nähe der Signifikanzgrenze, was vermutlich mit der großen Streuung der Datenpunkte zu tun hat. In standardsprachlicher Sprechweise stimmen die Raterinnen ebenfalls etwas positiver ab als die Rater. Der Unterschied ist jedoch nur sehr gering und nicht signifikant.

### 9.3.2 Zwischenfazit Raterinnen und Rater

Auch wenn sich die Ergebnisse über die einzelnen Ebenen und Situationen recht ähneln, zeigen sie ein erstaunliches Ergebnis. Zum einen wird deutlich, dass die Raterinnen häufig positiver abstimmen als die Rater – dieser Trend ist in allen Ebenen und beiden Situationen auffindbar. Damit kann auch  $H_{3n}$  (Es gibt keinen Unterschied in der Einschätzung von Spracheinstellungen auf Kompetenz- und Sympathieebene zwischen Schülerinnen und Schülern.) abgelehnt werden. Die Abstände in den Katego-

rien unterscheiden sich zwar, allerdings findet man keine Situation bzw. Ebene, in der die Rater mit ihren Einschätzungen über denjenigen der Raterinnen liegen würden.

Wenn man diesen Einschätzungsbias nun in die Betrachtung mit einfließen lässt, fällt zum anderen auf, dass die Raterinnen und Rater die Sprecherinnen und Sprecher recht ähnlich einschätzen. Die Konsequenz daraus ist, dass die negative Einschätzung mancher Guises – wie z. B., dass der dialektsprechende Lehrer kompetenter eingeschätzt wird als die dialektsprechende Lehrerin – in gleichem Maße von den Raterinnen als auch von den Ratern vorgenommen wird. Diese misogynen Tendenz in der Stichprobe wird also nicht nur von den Männern, sondern auch von den Frauen getragen. Wenn man sich die linke Hälfte von Abbildung 9.28 noch einmal ansieht, ist zu erkennen, dass die Raterinnen die Sprecher sogar signifikant positiver einschätzen als die Rater selbst. Die Raterinnen schätzen das männliche Geschlecht also kompetenter ein als die Rater ihr eigenes. Dadurch wird die Distanz zwischen den dialektsprechenden Sprechern und Sprecherinnen von den Raterinnen sogar vergrößert und damit auch die misogynen Tendenz dialektsprechenden Sprecherinnen gegenüber.

## 9.4 Städtische und ländliche Regionen

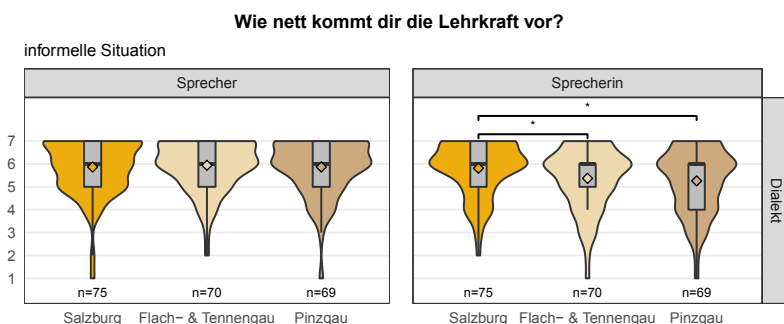
Die Schulorte wurden so gewählt, dass Unterschiede zwischen den Spracheinstellungen der Schüler:innen im städtischen, im ländlichen Raum sowie einem Übergangsbereich analysiert werden können. In den Grafiken dieses Abschnitts sind daher drei Analysekategorien nebeneinander angeordnet, die diesen Regionen entsprechen. Salzburg verweist dabei auf die Stadtschulen, in denen wir erhoben haben. Flach- & Tennengau verweist auf die Schulen in Pendeldistanz zur Stadt Salzburg und Pinzgau auf die Landschulen, die sich eben beide im Pinzgau befinden.<sup>25</sup>

### 9.4.1 Detaillierte Ergebnisse

Bei den Lehrkräften findet man auf Kompetenzebene weder bei den dialektalen noch den standardsprachlichen Guises Unterschiede zwischen den drei Gruppen. Dieses Bild ändert sich auf Sympathieebene. Hier zeigen sich in informeller Situation folgende Unterschiede: Bei Verwendung des Dialekts wird die Lehrerin in der Stadt Salzburg sympathischer bewertet als in Pendeldistanz zur Stadt und außerhalb der Pendeldistanz, wie in Abbildung 9.32 ersichtlich.

---

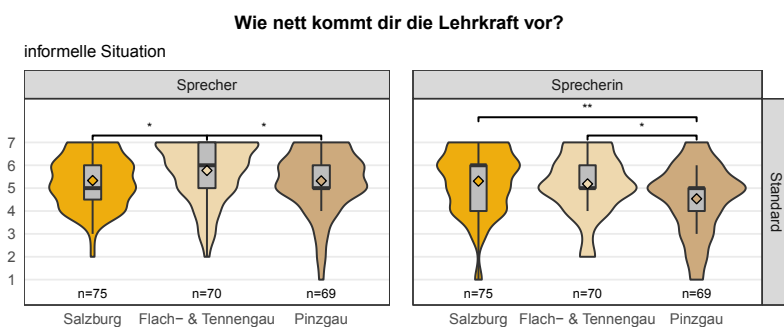
<sup>25</sup> Nähere Informationen zur Schulauswahl findet man in Abschnitt 8.3.



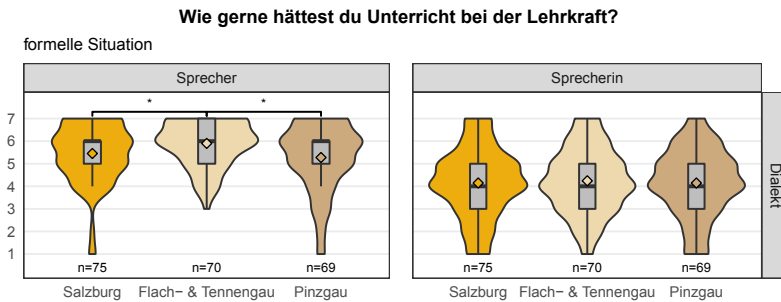
**Abbildung 9.32:** Stadt-Land-Vergleich bei der Frage nach der Sympathie der dialektsprechenden Lehrkräfte in informeller Situation

Beim Lehrer sind hier keine signifikanten Unterschiede messbar, auffällig ist aber, dass sich die Werte des Lehrers und der Lehrerin in den Ratings der Schüler:innen der Stadt-schulen ähneln.

Ähnliche Tendenzen lassen sich auch feststellen, wenn man sich die Verwendung der Standardsprache ansieht (s. Abb. 9.33). Hier sind zwar die Werte generell niedriger als bei Verwendung des Dialekts, das Bewertungsmuster ist bei der Sprecherin jedoch ähnlich. In der Stadt Salzburg wird die standardsprechende Lehrerin deutlich und hochsignifikant sympathischer eingeschätzt als im Pinzgau. Die Schüler:innen in Pendeldistanz schließen sich im Gegensatz zu den Dialekt-Guises in ihrer Einschätzung nun den Stadtschulen an. Beim standardsprechenden Lehrer ist die Situation eine andere. Am sympathischsten wird er von den Schüler:innen in Pendeldistanz bewertet, wobei diese Werte die positivsten Einschätzungen der Lehrerin übersteigen. In der Stadt und am Land fällt die Einschätzung signifikant negativer aus, wobei diese Werte ähnlich hoch sind wie diejenigen der Lehrerin in der Stadt.

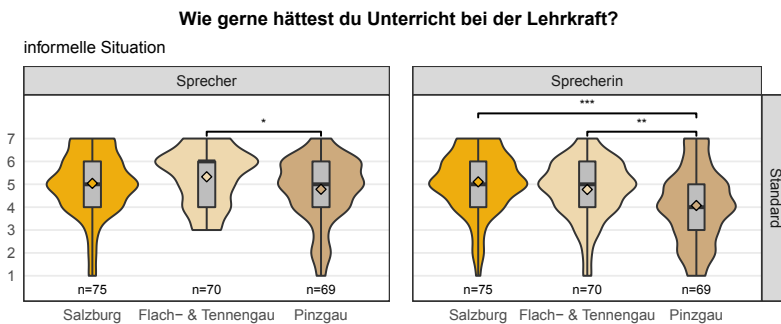


**Abbildung 9.33:** Stadt-Land-Vergleich bei der Frage nach der Sympathie der standardsprechenden Lehrkräfte in informeller Situation



**Abbildung 9.34:** Stadt-Land-Vergleich bei der Frage nach persönlicher Präferenz der dialektsprechenden Lehrkräfte in formeller Situation

Weitere Unterschiede lassen sich auf Ebene der persönlichen Präferenz feststellen. In formeller Situation sind beim dialektsprechenden Lehrer (s. Abb. 9.34) Ähnlichkeiten zu Abbildung 9.33 zu erkennen. Zwischen den Stadtschulen und den Schulen im Pinzgau sind keine Unterschiede feststellbar, die Schüler:innen in Pendeldistanz bewerten den Lehrer signifikant positiver als die beiden anderen Gruppen. Bei der dialektsprechenden Lehrerin sind keine signifikanten Stadt-Land-Unterschiede messbar, allerdings wird sie auffallend negativer eingeschätzt als der dialektsprechende Lehrer.



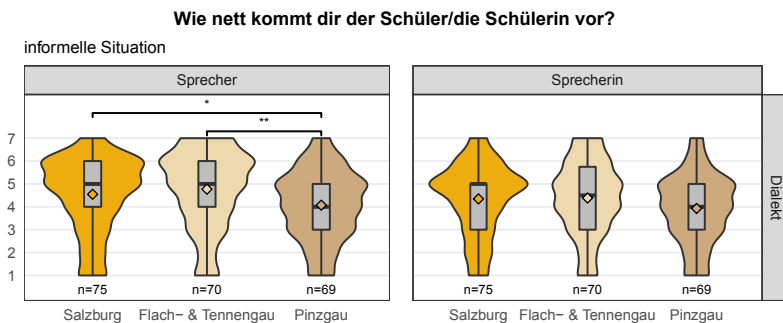
**Abbildung 9.35:** Stadt-Land-Vergleich bei der Frage nach persönlicher Präferenz der standardsprechenden Lehrkräfte in informeller Situation

In informeller Situation und bei Verwendung der Standardsprache findet man bei beiden Lehrkräften Unterschiede (s. Abb. 9.35). Beim Lehrer sind die Tendenzen wieder ähnlich wie in den Abbildungen 9.33 und 9.34. Zwischen Stadt Salzburg und den Schulen in Pendeldistanz sind die Bewertungsunterschiede zwar sichtbar – letztere bewerten etwas positiver – jedoch nicht signifikant messbar. Die Schüler:innen der Pendeldistanz bewerten den standardsprechenden Lehrer signifikant positiver als diejenigen im Pinzgau. Zwischen Stadt Salzburg und Pinzgau ist wiederum kein Unterschied messbar.

Bei der Lehrerin zeigen die Stadtschulen und die Schulen in Pendeldistanz keine messbaren Unterschiede, wobei die Einschätzungen von letzteren etwas negativer aus-

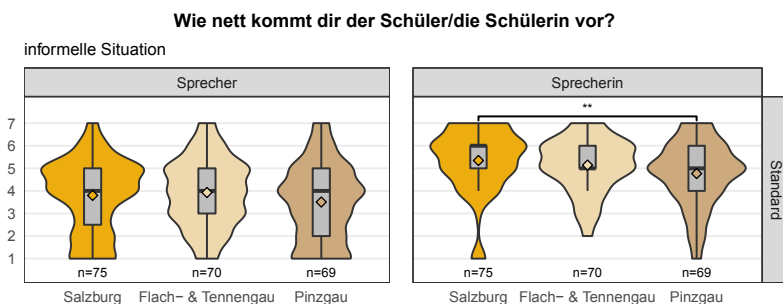
fallen. Im Pinzgau wird sie jedoch mit hoch- und höchstsignifikanten Unterschieden negativer eingeschätzt als von den beiden anderen Gruppen.

Bei den Schüler:innen-Guises sind wie bei den Lehrkräften keine signifikanten Unterschiede auf Kompetenzebene messbar. Auf Sympathieebene finden sich Auffälligkeiten in der informellen Situation.



**Abbildung 9.36:** Stadt-Land-Vergleich bei der Frage nach der Sympathie der dialekt-sprechenden Schüler:innen in informeller Situation

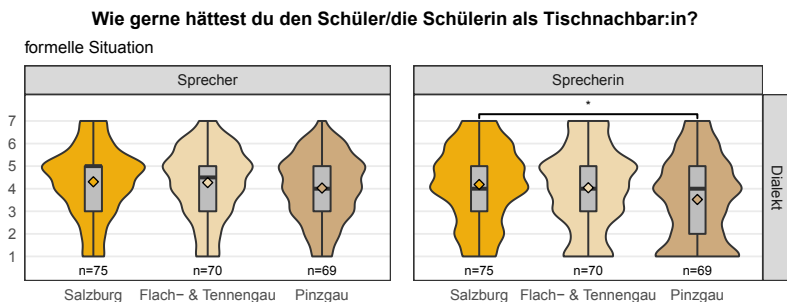
Der dialekt-sprechende Schüler (Abb. 9.36) wird in Stadt Salzburg und in Pendeldistanz mit signifikantem und hochsignifikantem Unterschied sympathischer eingeschätzt als im Pinzgau. Zwischen beiden ersteren Regionen bestehen keine signifikanten Unterschiede. Bei diesen beiden Gruppen kann man auch erkennen, dass sie den dialekt-sprechenden Schüler sympathischer einschätzen als die dialekt-sprechende Schülerin. Im Pinzgau scheint keine solche Differenzierung zu existieren.



**Abbildung 9.37:** Stadt-Land-Vergleich bei der Frage nach der Sympathie der standardsprechenden Schüler:innen in informeller Situation

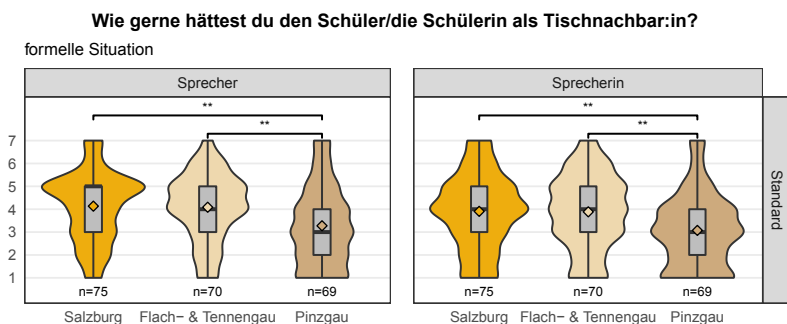
Man findet auch messbare Unterschiede, wenn man die standardsprechende Schülerin auf Sympathieebene betrachtet (s. Abb. 9.37). Hier kann man eine graduelle Abstufung der Bewertung von der positivsten Einschätzung in der Stadt Salzburg zur negativsten im Pinzgau erkennen. Die Unterschiede zwischen diesen beiden Regionen sind auch hochsignifikant. Da die Schulen der Pendeldistanz genau zwischen diesen Einschät-

zungen liegen, sind keine signifikanten Werte messbar. Im Vergleich zum standardsprechenden Schüler fällt auf, dass die Schülerin hier bei Weitem sympathischer eingeschätzt wird.



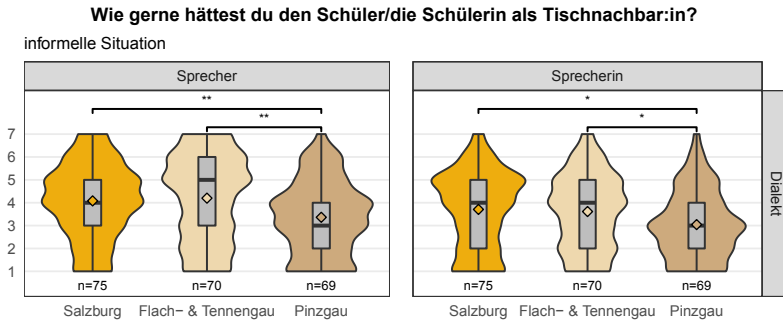
**Abbildung 9.38:** Stadt-Land-Vergleich bei der Frage nach persönlicher Präferenz der dialektsprechenden Schüler:innen in informeller Situation

Die Ebene der persönlichen Präferenz fördert die meisten Unterschiede zutage. Die Einschätzung der dialektsprechenden Schülerin in formeller Situation (Abb. 9.38) ähnelt derjenigen in Abbildung 9.37. Sie wird von den Schüler:innen der Stadtschulen signifikant positiver eingeschätzt als im Pinzgau. Die Schulen in Pendeldistanz liegen wiederum mit ihren Einschätzungen zwischen den beiden anderen Regionen und zeigen daher keine signifikanten Unterschiede.



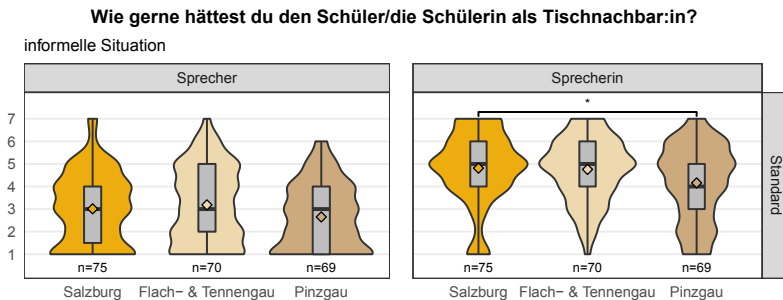
**Abbildung 9.39:** Stadt-Land-Vergleich bei der Frage nach persönlicher Präferenz der standardsprechenden Schüler:innen in formeller Situation

In standardsprachlicher Varietätenverwendung (Abb. 9.39) sind ähnliche Differenzen bei beiden Geschlechtern erkennbar. Zwischen den Stadtschulen und denjenigen in Pendeldistanz sind keine signifikanten Unterschiede erkennbar, in den Schulen im Pinzgau wird sowohl der Schüler als auch die Schülerin hochsignifikant weniger als Tischnachbar:in gewollt.



**Abbildung 9.40:** Stadt-Land-Vergleich bei der Frage nach persönlicher Präferenz der dialektsprechenden Schüler:innen in informeller Situation

Dieses Muster setzt sich in informeller Situation und dialektaler Sprechweise fort. Hier sind ebenfalls bei beiden Geschlechtern Differenzen zwischen den Regionen erkennbar (s. Abb. 9.40). Sowohl beim Schüler als auch bei der Schülerin besteht kein messbarer Unterschied zwischen den Einschätzungen aus der Stadt Salzburg und dem Flach- & Tennengau. Wiederum fällt die Einschätzung im Pinzgau negativer aus – beim Schüler sind die Unterschiede hochsignifikant, bei der Schülerin signifikant. Der dialektsprechende Schüler wird dabei in der Stadt Salzburg und dem Flach- & Tennengau positiver eingeschätzt als die dialektsprechende Schülerin.

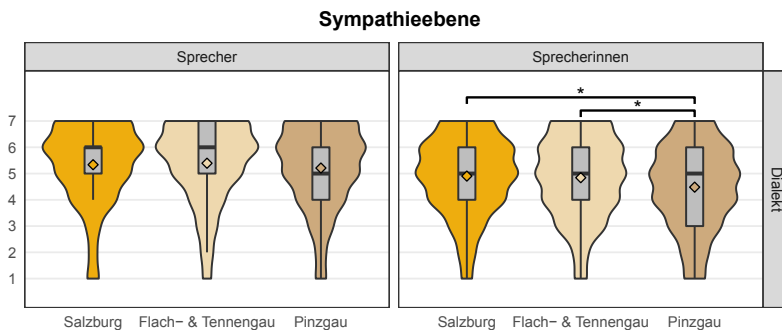


**Abbildung 9.41:** Stadt-Land-Vergleich bei der Frage nach persönlicher Präferenz der dialektsprechenden Schüler:innen in informeller Situation

Bei Verwendung der Standardsprache ist auf Ebene der persönlichen Präferenz bei der Schülerin ein signifikanter Unterschied zwischen der Stadt Salzburg und dem Pinzgau, in dem die Einschätzungen negativer ausfallen, vorhanden (s. Abb. 9.41). Die Schüler:innen in Pendeldistanz liegen mit ihren Einschätzungen leicht unter denjenigen der Stadt (was man vor allem am Violinplot beobachten kann), daher und auch wegen der großen Varianz der Einschätzungen im Pinzgau, liegen die p-Werte zu beiden anderen Regionen nicht im signifikanten Bereich.

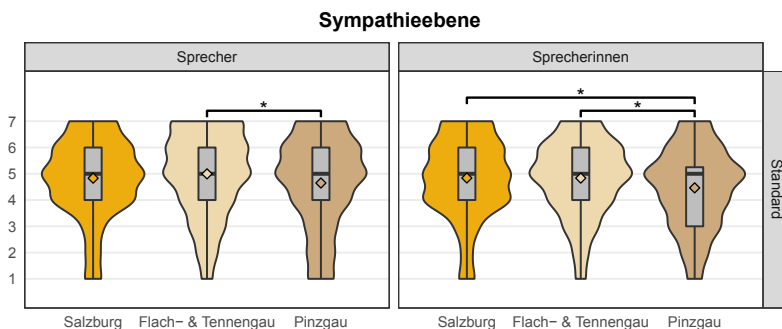
### 9.4.2 Generelle Ergebnisse

Wie schon bei den detaillierten Ergebnissen sind signifikante Unterschiede nur auf Sympathieebene und Ebene der persönlichen Präferenz messbar. Daher kann  $H_{5n}$  „Es existiert kein Unterschied zwischen den Kompetenzeinschätzungen von Schüler:innen am Land, in Pendeldistanz zur Stadt und der Stadt in Bezug auf dialektprechende Personen“ nicht abgelehnt werden.



**Abbildung 9.42:** Stadt-Land-Vergleich der allgemeinen Einschätzung der Sympathie bei dialektaler Sprachverwendung

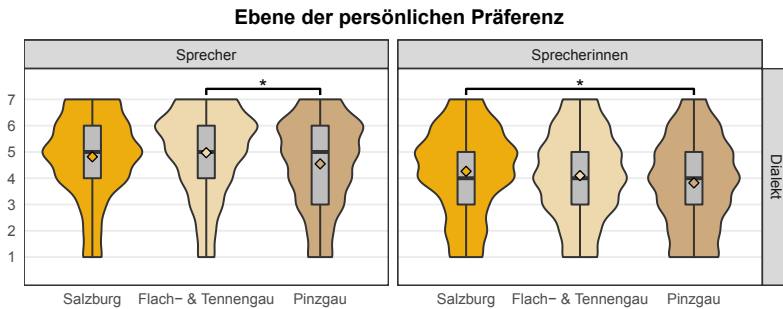
Abbildung 9.42 zeigt, dass die dialektprechenden Sprecherinnen von den Rater:innen in der Stadt Salzburg und dem Flach- & Tennengau signifikant sympathischer eingeschätzt werden als im Pinzgau. Diese Tendenz ist bei den Sprechern ebenfalls erkennbar, jedoch weniger ausgeprägt und auch nicht signifikant.



**Abbildung 9.43:** Stadt-Land-Vergleich der allgemeinen Einschätzung der Sympathie bei standardsprachlicher Sprachverwendung

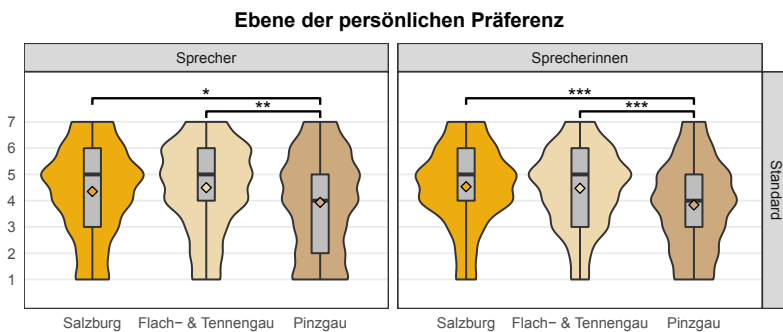
Die Einschätzungen der dialektprechenden Sprecher selbst fallen jedoch durch die Bank höher aus. Auch in standardsprachlicher Sprechweise (Abb. 9.43) werden die Sprecherinnen von den Rater:innen aus der Stadt Salzburg und dem Flach- & Tennengau signifikant positiver eingeschätzt als von denjenigen im Pinzgau. Der Vergleich mit den Sprechern zeigt, dass diese Tendenz hier ebenso existiert und zwischen Flach- & Ten-

nengau und dem Pinzgau auch signifikante Messergebnisse liefert. Damit kann  $H_{6n}$ : „Es existiert kein Unterschied zwischen den Sympathieeinschätzungen von Schüler:innen am Land, in Pendeldistanz zur Stadt und der Stadt in Bezug auf standardsprechende Personen“ abgelehnt werden. Der Unterschied zwischen den Sympathiewerten der Sprecher und Sprecherinnen ist viel schwächer ausgeprägt als bei dialektaler Sprechweise.



**Abbildung 9.44:** Stadt-Land-Vergleich der allgemeinen Einschätzung der persönlichen Präferenz bei dialekt-sprachlicher Sprachverwendung

Auf Ebene der persönlichen Präferenz setzen sich diese Tendenzen in ähnlicher Weise fort. In dialektaler Sprachverwendung (Abb. 9.44) schätzen die Rater:innen der Stadt Salzburg die Sprecherinnen signifikant positiver ein als die Rater:innen aus dem Pinzgau und auch bei den Sprechern findet man ein signifikantes Ergebnis: Sie werden im Flach- & Tennengau signifikant positiver eingeschätzt als im Pinzgau. Die Werte selbst sind für die dialektsprechenden Sprecher wiederum deutlich höher als für die Sprecherinnen.



**Abbildung 9.45:** Stadt-Land-Vergleich der allgemeinen Einschätzung der persönlichen Präferenz bei standard-sprachlicher Sprachverwendung

Die deutlichsten Unterschiede sind auf Ebene der persönlichen Präferenz und standardsprachlicher Varietätenverwendung zu finden (Abb. 9.45), wenn auch die Ten-

denzen gleich bleiben. Die Sprecherinnen werden in der Stadt Salzburg und in Flach- & Tennengau höchstsignifikant positiver eingeschätzt als im Pinzgau. Ebenfalls werden die Sprecher im Flach- & Tennengau hochsignifikant positiver und in der Stadt Salzburg signifikant positiver eingeschätzt als im Pinzgau. Die Unterschiede zwischen Sprechern und Sprecherinnen sind in dieser Varietätenverwendung wieder weniger stark ausgeprägt.

### 9.4.3 Zwischenfazit städtische und ländliche Regionen

Die Analyse der Regionen zeigt nicht in allen Ebenen und Situationen Unterschiede auf. Auf Kompetenzebene sind keine Unterschiede zwischen den Gruppen feststellbar, auf Sympathieebene betreffen diese nur die informellen Situationen und auf Ebene der persönlichen Präferenz sind in beiden Situationen Unterschiede messbar.

Besonders auffällig ist, dass die Rater:innen im Pinzgau in ihren Einschätzungen zumeist negativer sind als die der beiden anderen Regionen. Dies betrifft dabei gleichermaßen die Guises im Dialekt als auch in der Standardsprache. Dass es sich hierbei um eine allgemeine Tendenz handelt, kann jedoch ausgeschlossen werden, da bei 48 analysierten Plots 34 keine signifikanten Unterschiede aufweisen – auf Kompetenzebene z. B. kein einziger.

Die negativeren Einschätzungen müssen also einen anderen Hintergrund haben. Um die Guises konstant zu halten, wurden für alle drei Regionen die gleichen Sprecher:innen verwendet. Diese sprechen einen Dialekt des Salzburger Zentralraums, der sich sprachlich vom Pinzgau abhebt. Es ist also denkbar, dass die Pinzgauer Schüler:innen auf Sympathieebene in informeller Situation die dialektsprechende Lehrerin und den dialektsprechenden Schüler negativer beurteilen, da diese nicht in der vertrauten dialektalen Varietät kommunizieren. Da diese Varietät aus dem Zentralraum stammt und damit vermutlich über verdecktes Prestige verfügt, ist es auch nachvollziehbar, dass dieser Bewertungsunterschied in formeller Situation nicht vorkommt.

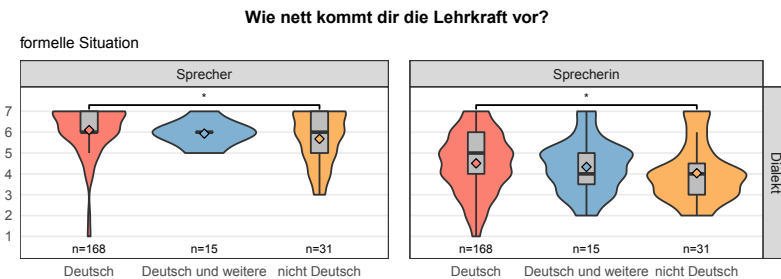
Die Ursache für die negativere Einschätzung von standardsprechenden Schüler:innen und Lehrkräften auf Sympathieebene und Ebene der persönlichen Präferenz kann darin liegen, dass die Standardvarietät im Pinzgau generell weniger in informellen Situationen verwendet wird. In der formelleren wird, ähnlich wie die Varietät aus dem Zentralraum, die Standardsprache offensichtlich mehr toleriert. In den dialektsprachlichen Guises werden die Lehrer im Flach- & Tennengau sowie im Pinzgau sympathischer eingeschätzt als die Lehrerinnen, einzig in der Stadt Salzburg findet man keine derartigen Unterschiede. Bei den Schüler:innen ist diese Tendenz ebenfalls zu erkennen, allerdings findet man hier in Stadt Salzburg und dem Flach- & Tennengau Unterschiede und nicht im Pinzgau. Bei den standardsprachlichen Guises werden Lehrerinnen und Lehrer auf Sympathieebene in etwa ähnlich eingeschätzt (einzig die standardsprechende Lehrerin im Pinzgau erhält deutlich geringere Sympathiewerte), allerdings werden im Gegensatz zu den dialektsprachlichen Guises nun die Schülerinnen in allen Regionen deutlich sympathischer eingeschätzt als die Schüler. Die Ebene der persönlichen Präferenz bestätigt die Tendenzen, die auf Sympathieebene festgestellt wurden, in noch deutlicherem Ausmaß.

## 9.5 Erstsprache

Wenden wir uns nun potenziellen Einstellungsunterschieden von Schüler:innen mit unterschiedlichen Erstsprachen zu. In der Umfrage gab es hierzu die Möglichkeit, eine oder mehrere Erstsprachen anzugeben. Für die quantitative Analyse wurden die Angaben in drei Gruppen eingeteilt: in Personen mit Deutsch als Erstsprache, in Personen mit Deutsch und einer weiteren Erstsprache und in Personen, die eine andere Erstsprache als Deutsch angaben. Die Anzahl der Personen in den drei Gruppen ist sehr unterschiedlich, wie man an den n-Werten unter den Plots ablesen kann. Speziell die Gruppe „Deutsch und weitere“ ist mit 15 Personen zu klein, um verlässliche inferenzstatistische Aussagen treffen zu können, daher sind die Ergebnisse dieser Gruppe mit Vorbehalt zu sehen.<sup>26</sup>

### 9.5.1 Detaillierte Ergebnisse

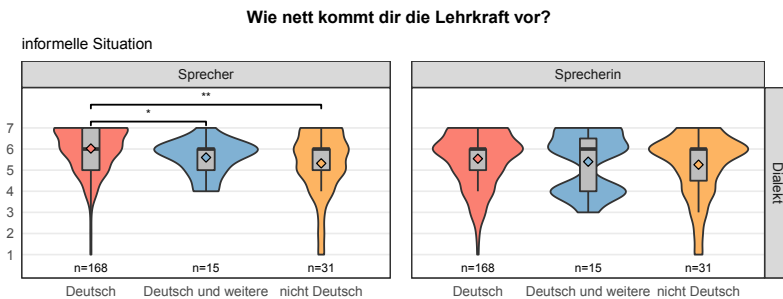
Bei den Lehrkräften sind, ähnlich wie schon beim Stadt-Land-Vergleich, keine Unterschiede auf Kompetenzebene auszumachen. Auf Sympathieebene liegen Unterschiede sowohl in formeller als auch informeller Situation vor.



**Abbildung 9.46:** Erstsprachvergleich bei der Frage nach der Sympathie der dialektsprechenden Lehrkräfte in formeller Situation

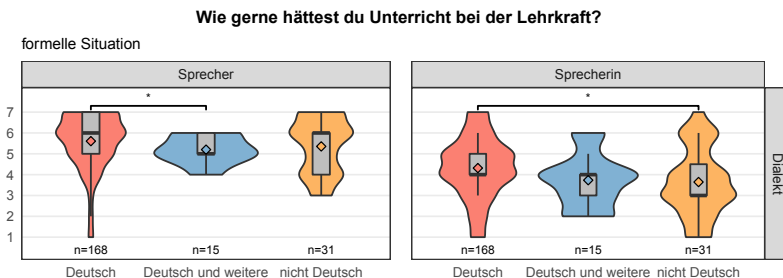
In formeller Situation (s. Abb. 9.46) trifft man auf einen Unterschied in den Einschätzungen zwischen den Personen mit Deutsch als Erstsprache und denjenigen mit einer anderen Erstsprache als Deutsch – Erstere schätzen sowohl den dialektsprechenden Lehrer als auch die dialektsprechende Lehrerin signifikant sympathischer ein. Die Gruppe mit Deutsch und einer weiteren Erstsprache liegt zwischen den beiden Gruppen, daher sind hier keine signifikanten Unterschiede messbar.

<sup>26</sup> Für statistische Tests wird in der Literatur häufig eine Gruppengröße von mindestens  $n = 30$  vorgeschlagen, wie z. B. bei Levshina (2015: 88). Der Grund dafür ist, dass viele statistische Tests von normalverteilten Daten ausgehen. Je größer die Gruppengröße ist, desto leichter sind solche Daten durch Bootstrapping-Methoden zu erreichen. Dass der Wilcoxon-Test zu den nichtparametrischen Tests zählt und daher nicht auf normalverteilte Daten angewiesen ist, ist für die Berechnung zwar ein Vorteil, es bleibt jedoch problematisch, von nur 15 Personen auf eine Grundgesamtheit schließen zu wollen.



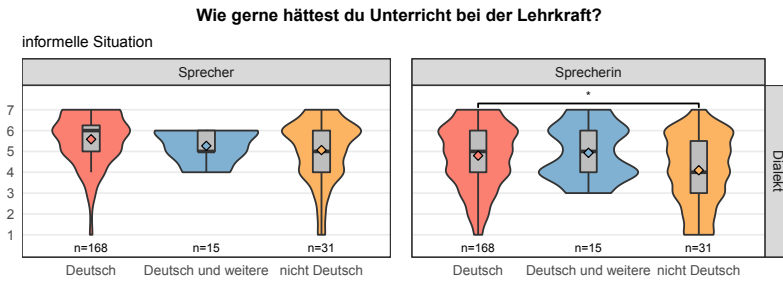
**Abbildung 9.47:** Erstsprachvergleich bei der Frage nach der Sympathie der dialektsprechenden Lehrkräfte in informeller Situation

Ähnlich verhält es sich in informeller Situation beim dialektprechenden Lehrer (s. Abb. 9.47). Dieser wird von den Rater:innen mit Deutsch als Erstsprache signifikant sympathischer eingeschätzt als von der Gruppe mit Deutsch und einer weiteren Erstsprache. Zur Gruppe mit Deutsch als L2 ist der Unterschied sogar hochsignifikant. In standardsprachlicher Varietätenverwendung findet man keine Unterschiede auf Sympathieebene.



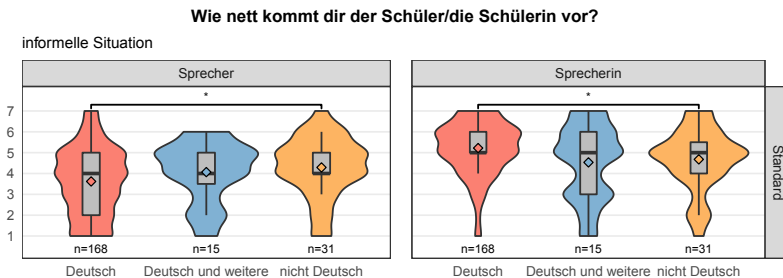
**Abbildung 9.48:** Erstsprachvergleich bei der Frage nach persönlicher Präferenz der dialektsprechenden Lehrkräfte in formeller Situation

Auch auf Ebene der persönlichen Präferenz lassen sich Unterschiede ausmachen, alle wiederum in dialektaler Sprechweise. In formeller Situation wird der dialektprechende Lehrer von den Rater:innen mit Deutsch als Erstsprache signifikant positiver eingeschätzt als von jenen mit Deutsch und einer weiteren Erstsprache. Die dialektprechende Lehrerin wird von Personen mit Deutsch als Erstsprache signifikant positiver eingeschätzt als von jenen mit einer anderen Erstsprache als Deutsch (s. Abb. 9.48). In diesem Plot findet man wiederum das Muster, das schon in Abbildungen 9.46 und 9.47 aufgetreten ist: Die Gruppe „Deutsch“ bewertet am positivsten, danach kommt die Gruppe „Deutsch und weitere“ und am Schluss die Gruppe „nicht Deutsch“.



**Abbildung 9.49:** Erstsprachvergleich bei der Frage nach persönlicher Präferenz der dialektsprechenden Lehrkräfte in informeller Situation

In informeller Situation ist die Lage ebenfalls ähnlich wie in formeller Situation (s. Abb. 9.49). Die Lehrerin wird von den Personen mit Deutsch als Erstsprache signifikant positiver eingeschätzt als von den beiden anderen Gruppen, wobei in der Gruppe „Deutsch und weitere“ eine bimodale Verteilung auffällt. Beim Lehrer zeigen sich zwar keine signifikanten Unterschiede, die Werte der Datenverteilung der einzelnen Gruppen sind jedoch wiederum von links nach rechts abfallend. Der p-Wert zwischen den beiden äußeren Gruppen liegt hier mit 0.07 auch nahe an der Signifikanzgrenze.

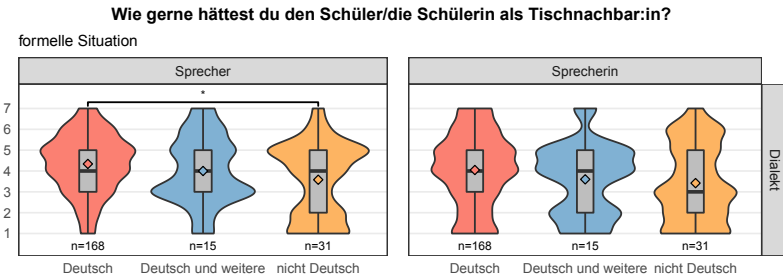


**Abbildung 9.50:** Erstsprachvergleich bei der Frage nach der Sympathie der standardsprechenden Schüler:innen in informeller Situation

Bei den Schüler:innen-Guises liegen nur in zwei Plots Unterschiede vor. Auf Sympathieebene in informeller Situation wird der standardsprechende Schüler von den Rater:innen mit einer anderen Erstsprache als Deutsch signifikant sympathischer eingeschätzt als von denen mit Deutsch als Erstsprache (s. Abb. 9.50). Die Gruppe „Deutsch und weitere“ liegt zwischen diesen beiden Gruppen ohne signifikant messbare Unterschiede.

Diametral anders verhält sich die Sachlage bei der Schülerin. Sie wird von den Deutsch-Erstsprachler:innen signifikant sympathischer eingeschätzt als von den Personen mit einer anderen Erstsprache als Deutsch. Die Personen mit Deutsch und einer weiteren Erstsprache liegen zwar noch etwas unter letzterer Gruppe, allerdings sieht man eine bimodale Verteilung und durch die große Standardabweichung der

Werte ist es plausibel, dass keine signifikanten Unterschiede berechnet werden. Auffällig ist, dass die Einstellungen zu den beiden Geschlechtern bei den Personen mit einer anderen Erstsprache als Deutsch ziemlich ähnlich sind, bei den Deutsch-Erstsprachler:innen jedoch auseinander liegen.



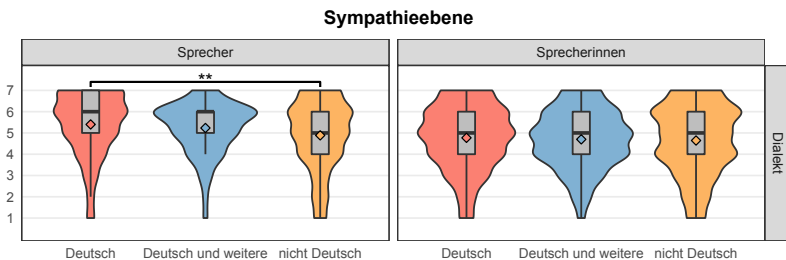
**Abbildung 9.51:** Erstsprachvergleich bei der Frage nach persönlicher Präferenz der dialektstprechenden Schüler:innen in formeller Situation

Der letzte Unterschied zeigt sich auf Ebene der persönlichen Präferenz in formeller Situation und dialektaler Sprechweise (Abb. 9.51). Hier zeigt sich wiederum das schon bekannte Bild, dass die Bewertungen bei den Rater:innen mit Deutsch als Erstsprache am höchsten sind, etwas niedriger bei den Personen mit Deutsch und einer weiteren Erstsprache und am niedrigsten bei denjenigen mit einer anderen Erstsprache als Deutsch. Beim Schüler ist zwischen den äußeren Gruppen auch ein signifikanter Unterschied messbar.

### 9.5.2 Generelle Ergebnisse

Wie schon die detaillierten Ergebnisse keine Unterschiede auf Kompetenzebene erbrachten, ist dies ebenfalls bei generalisierter Berechnung der Fall. Man findet allerdings Ergebnisse auf Sympathieebene und Ebene der persönlichen Präferenz – jeweils nur in dialektaler Sprachverwendung.

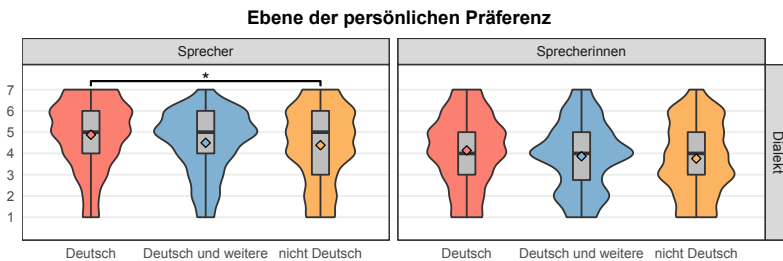
Abbildung 9.52 zeigt nun die generellen Werte der Sympathieebene bei dialektaler Sprechweise für beide Geschlechter.



**Abbildung 9.52:** Erstsprachvergleich bei der Frage nach der Sympathie der dialektstprechenden Lehrkräfte in formeller Situation

Man erkennt, dass die Sprecher von der Personengruppe mit Deutsch als Erstsprache hochsignifikant sympathischer eingeschätzt werden als von der Gruppe mit einer anderen Erstsprache als Deutsch. Bei den Sprecher:innen ist dies nicht der Fall. Hier sieht man, dass alle drei Gruppen sehr ähnlich in ihren Einschätzungen sind. Interessant dabei ist, dass die Gruppe mit einer anderen Erstsprache als Deutsch die Sprecher:innen und Sprecher viel ähnlicher einschätzt als die Gruppe der Deutsch-Erstsprachler:innen. Die Werte der Gruppe mit Deutsch und einer weiteren Erstsprache liegen zwischen den beiden Gruppen. Trotz der offensichtlichen Geschlechterunterschiede kann  $H_{7n}$ : „Es existiert kein Unterschied in der Sympathieeinschätzung von Schüler:innen mit Deutsch als Erstsprache, Deutsch und einer weiteren Erstsprache und einer anderen Erstsprache als Deutsch in Bezug auf dialekt sprechende Personen“ eindeutig abgelehnt werden.

Auf Ebene der persönlichen Präferenz sind ähnliche Tendenzen bei dialektaler Sprachverwendung feststellbar (s. Abb. 9.53). Hier ist es zusätzlich so, dass auch die Sprecher:innen von den Deutsch-Erstsprachler:innen positiver eingeschätzt werden. Die Unterschiede befinden sich jedoch nicht im signifikanten Bereich. Die unterschiedliche Einschätzung der Sprecher fällt ebenfalls nicht ganz so deutlich aus, ist aber dennoch signifikant.



**Abbildung 9.53:** Erstsprachvergleich der allgemeinen Einschätzung der persönlichen Präferenz bei dialektaler Sprechweise

### 9.5.3 Zwischenfazit Erstsprache

Bei der Betrachtung der unterschiedlichen Erstsprachen zeigt sich, dass sie keinen messbaren Einfluss auf die Kompetenzeinschätzungen zeigen. Einstellungsunterschiede sind vor allem auf Sympathieebene und Ebene der persönlichen Präferenz bei dialektaler Sprechweise vorhanden: Rater:innen mit Deutsch als Erstsprache schätzen dialektale Sprechweise sympathischer und allgemein positiver ein als Personen mit einer anderen Erstsprache als Deutsch. Auffällig ist der einzige Bewertungsunterschied, der in standardsprachlicher Sprechweise aufzufinden ist (s. Abb. 9.50). Rater:innen mit Deutsch als Erstsprache schätzen die Schülerin viel sympathischer ein als den Schüler und zeigen hier ein differenziertes Einschätzungsmuster, das bei Personen mit einer anderen Erstsprache als Deutsch nicht vorliegt. Auch bei dialektaler Sprechweise werden die Guises der Sprecher:innen und Sprecher auf Sympathieebene und Ebene

der persönlichen Präferenz von den Rater:innen mit einer anderen Erstsprache als Deutsch ähnlicher eingeschätzt als von den Deutsch-Erstsprachler:innen.

Eine mögliche Erklärung hierfür könnte darin liegen, dass dieser Einstellungsunterschied zu den Geschlechtern bei unterschiedlicher Varietätenverwendung, der schon in Abschnitt 9.1 auffällig war, von Personen mit einer anderen Erstsprache als Deutsch nicht oder nur in abgeschwächter Form erworben wird. Dies würde erklären, warum Spracheinstellungen zu Varietäten des Deutschen bei dieser Gruppe zwischen den Geschlechtern nicht differenziert sind.

## 9.6 Bildungshintergrund

Um den Fragen nach dem Bildungshintergrund nachgehen zu können, wurde am Ende der Umfrage eine Frage nach der Anzahl der Bücher im Elternhaus gestellt, die in Abbildung 9.54 zu sehen ist. Diese Frage soll Rückschlüsse auf den Bildungshintergrund des Elternhauses erlauben, ohne zu direkt dieses sensible Thema anzusprechen. Es musste eine Antwort ausgewählt werden, ein Überspringen der Frage war nicht möglich.

Wie viele Bücher gibt es bei dir zu Hause ungefähr? Zähle NICHT mit: Zeitschriften, Zeitungen und deine Schulbücher.

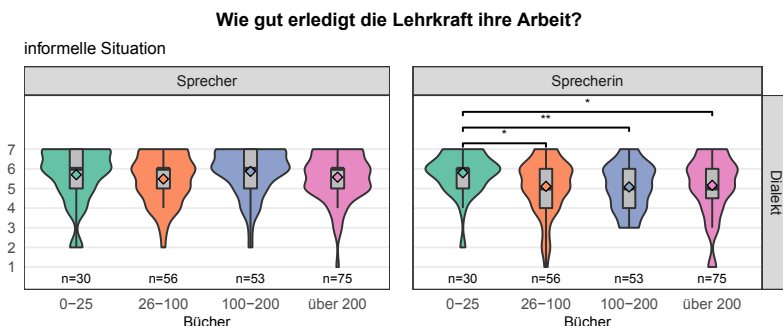
|  |   |   |  |   |
|--|---|---|--|---|
| <input type="radio"/> keine oder nur sehr wenige (0-10 Bücher) | <input type="radio"/> genug, um ein Regalbrett zu füllen (11-25 Bücher) | <input type="radio"/> genug, um ein Regal zu füllen (26-100 Bücher) | <input type="radio"/> genug, um zwei Regale zu füllen (101-200 Bücher) | <input type="radio"/> genug, um drei oder mehr Regale zu füllen (über 200 Bücher) |
|--|---|---|--|---|

**Abbildung 9.54:** Fragen zum Bildungshintergrund

Da die Umfrage in Gymnasialklassen der zehnten Schulstufe durchgeführt wurde, gab es in Relation nur wenige Personen, die die ersten beiden Kategorien ankreuzten. Von den 214 Personen wählten nur neun die Kategorie „0–10“ und 21 die Kategorie „11–25“ Bücher. Da diese Zahlen für statistische Signifikanztests zu gering sind, wurden für die Analyse diese beiden Gruppen zur Kategorie „0–25“ vereint. Wie dies schon in Abschnitt 9.3 teilweise gemacht wurde, werden, um Platz zu sparen, zwei Diagramme unterschiedlicher Kategorien nebeneinander abgebildet.

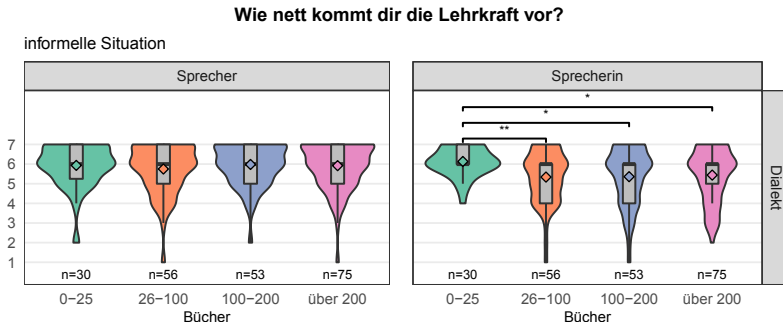
### 9.6.1 Detaillierte Ergebnisse

Auf Kompetenzebene findet man in den meisten Bereichen keine Unterschiede vor. Hier scheinen die einzelnen Gruppen zumeist ähnlicher Meinung zu sein. Die einzige Abweichung von diesem Muster findet man bei der dialektsprechenden Lehrerin in informeller Situation (s. Abb. 9.55).



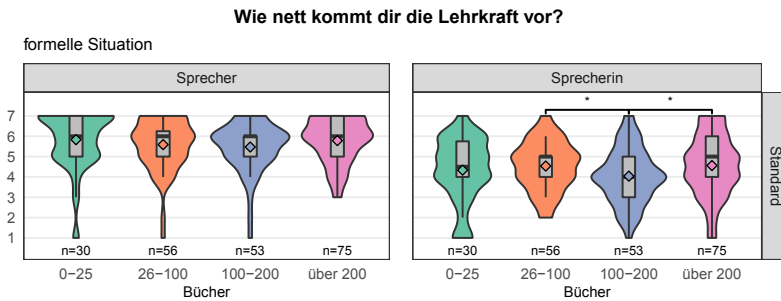
**Abbildung 9.55:** Vergleich des Bildungshintergrunds bei der Frage nach der Kompetenz der dialektsprechenden Lehrkräfte in informeller Situation

Diese wird von der Gruppe „0–25“ am kompetentesten eingeschätzt, mit signifikanten und hochsignifikanten Unterschieden zu den anderen Gruppen, die sich in ihren Einschätzungen wiederum recht ähneln. Beim Lehrer hingegen sieht man sehr ähnliche Einschätzungen der einzelnen Gruppen ohne signifikante Ergebnisse. Interessant ist allerdings, dass die Gruppe mit dem niedrigsten Bildungshintergrund offensichtlich die geringsten Einschätzungsunterschiede zwischen den Geschlechtern aufweist – sie bewertet die dialektssprechende Lehrerin sogar leicht kompetenter als den Lehrer.



**Abbildung 9.56:** Vergleich des Bildungshintergrunds bei der Frage nach der Sympathie der dialektsprechenden Lehrkräfte in informeller Situation

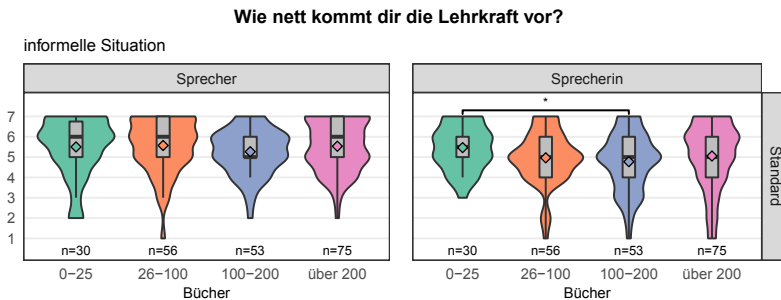
Ein vergleichbares Ergebnis, wenn auch mit etwas höheren Einschätzungswerten, findet man bei der dialektssprechenden Lehrerin auf Sympathieebene in informeller Situation (s. Abb. 9.56). Diese wird von der Gruppe mit dem niedrigsten Bildungshintergrund am sympathischsten eingeschätzt, wiederum mit signifikanten und hochsignifikanten Unterschieden zu den drei anderen Gruppen.



**Abbildung 9.57:** Vergleich des Bildungshintergrunds bei der Frage nach der Sympathie der standardsprechenden Lehrkräfte in formeller Situation

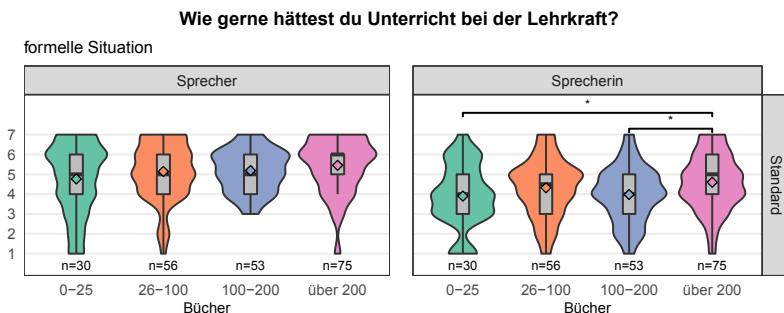
Auch hier sieht man im Vergleich zum Lehrer, dass die Gruppe mit dem niedrigsten Bildungshintergrund in ihren Werten den geringsten Unterschied zwischen den Geschlechtern aufweist – wiederum wird die dialektsprechende Lehrerin sogar sympathischer eingeschätzt als der Lehrer.

Auf Sympathieebene findet man ebenfalls Unterschiede bei der standardsprechenden Lehrerin, und zwar in beiden Situationen. In formeller Situation (s. Abb. 9.57) ist ein divergierendes Verhalten von der Gruppe „100–200“ zu erkennen, die die Lehrerin signifikant weniger sympathisch einschätzt als die beiden sie umgebenden Gruppen. Auffällig ist auch der Unterschied zur Einschätzung des Lehrers, der beträchtlich ist.



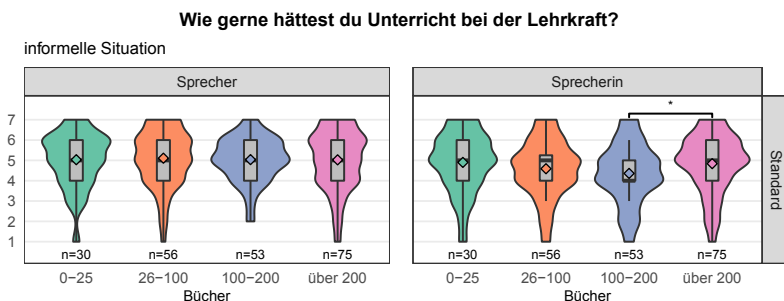
**Abbildung 9.58:** Vergleich des Bildungshintergrunds bei der Frage nach der Sympathie der standardsprechenden Lehrkräfte in informeller Situation

In informeller Situation (s. Abb. 9.58) sticht wiederum die Gruppe mit dem niedrigsten Bildungshintergrund heraus, die die standardsprechende Lehrerin signifikant sympathischer einschätzt als die Gruppe „100–200“. Auch in dieser Situation ist zu erkennen, dass die Gruppe mit dem niedrigsten Bildungshintergrund ähnliche Einschätzungen des Lehrers und der Lehrerin tätigt als die anderen Gruppen. Der Unterschied ist jedoch deutlich kleiner als in Abbildung 9.57.



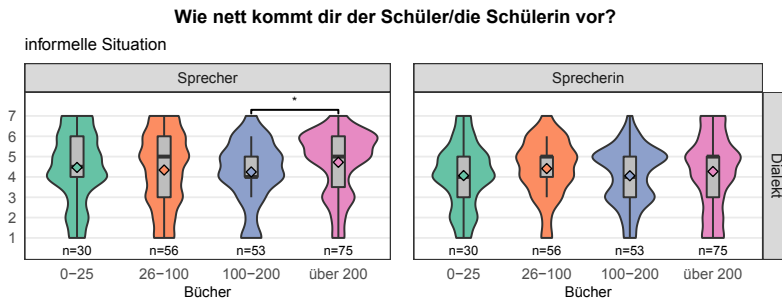
**Abbildung 9.59:** Vergleich des Bildungshintergrunds bei der Frage nach persönlicher Präferenz der standardsprechenden Lehrkräfte in formeller Situation

Auf Ebene der persönlichen Präferenz findet man ebenfalls bei der standardsprechenden Lehrerin in formeller wie informeller Situation Unterschiede. In formeller Situation (s. Abb. 9.59) schätzt die Gruppe mit dem höchsten Bildungshintergrund die Lehrerin signifikant positiver ein als die Gruppen „0–25“ und „100–200“. Interessant ist, dass man folgenden Trend erkennen kann: Je höher der Bildungshintergrund, desto eher wünschen sich die Rater:innen die standardsprechende Lehrerin im Unterricht. Diese Einschätzung wird dadurch noch verstärkt, wenn man sich die Ergebnisse des Lehrers ansieht, bei dem dieser Trend sogar noch deutlicher zu sehen ist.



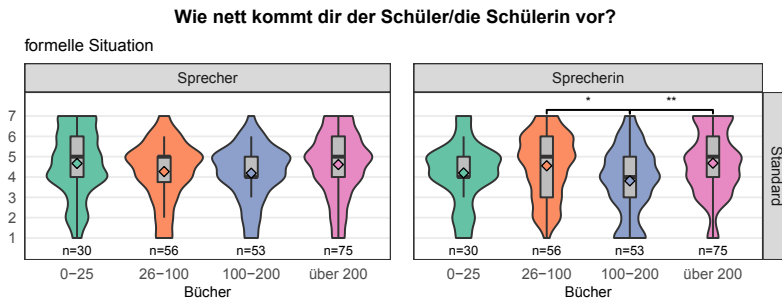
**Abbildung 9.60:** Vergleich des Bildungshintergrunds bei der Frage nach persönlicher Präferenz der standardsprechenden Lehrkräfte in informeller Situation

In informeller Situation findet man ein ähnliches Ergebnis zwischen den Gruppen „100–200“ und „über 200“, allerdings ist ansonsten kein Trend auszumachen. Auch der Blick auf die Ergebnisse des Lehrers verdeutlicht, dass die Gruppen sehr ähnliche Ratingmuster aufweisen. Der Vergleich zeigt außerdem, dass die beiden Gruppen mit dem niedrigsten und höchsten Bildungshintergrund die Geschlechter sehr ähnlich einschätzen, die beiden mittleren Gruppen, und speziell Gruppe „100–200“, schätzen jedoch den Lehrer deutlich positiver ein als die Lehrerin.



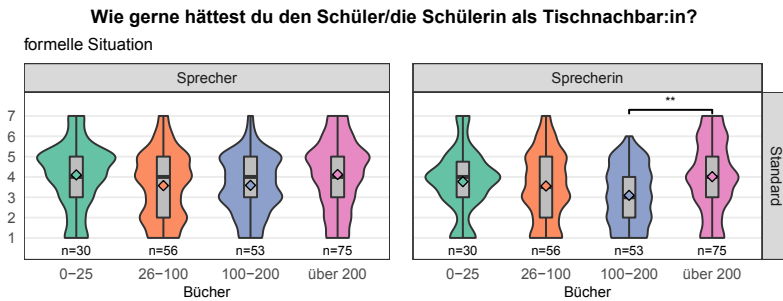
**Abbildung 9.61:** Vergleich des Bildungshintergrunds bei der Frage nach der Sympathie der dialektspredenden Schüler:innen in informeller Situation

Bei den Schüler:innen findet man keine signifikanten Unterschiede auf Kompetenzebene vor. Auf Sympathieebene wird der dialektspredende Schüler in informeller Situation von der Gruppe mit höchstem Bildungshintergrund signifikant sympathischer eingeschätzt als von der Gruppe „100–200“ (Abb. 9.61). Die dialektspredende Schülerin wird im Vergleich ähnlich sympathisch wie der Schüler eingeschätzt, alle Werte liegen zwischen vier und fünf.



**Abbildung 9.62:** Vergleich des Bildungshintergrunds bei der Frage nach der Sympathie der standardspredenden Schüler:innen in formeller Situation

Die standardspredende Schülerin wird von der Gruppe mit dem höchsten Bildungshintergrund und der Gruppe „26–100“ in formeller Situation signifikant bzw. hochsignifikant sympathischer eingeschätzt, als es die Gruppe „100–200“ vornimmt (s. Abb. 9.62). Im Vergleich der Geschlechter sind vor allem die beiden Gruppen „0–25“ und „26–100“ auffällig. Erstere Gruppe schätzt den Sprecher sympathischer ein als die Sprecherin, bei letzterer verhält es sich genau umgekehrt. Die Werte zwischen den Geschlechtern liegen aber nahe beieinander.



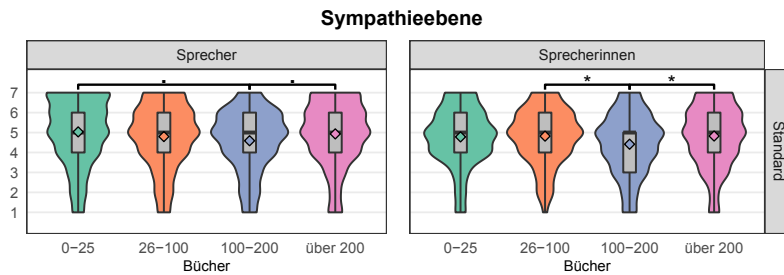
**Abbildung 9.63:** Vergleich des Bildungshintergrunds bei der Frage nach persönlicher Präferenz der standardsprechenden Schüler:innen in formeller Situation

Die Schülerin schneidet auch auf Ebene der persönlichen Präferenz in formeller Situation bei der Gruppe mit höchstem Bildungshintergrund hochsignifikant besser ab als in der Gruppe „100–200“ (Abb. 9.63). Dies soll jedoch nicht darüber hinwegtäuschen, dass die Messwerte in dieser Kategorie generell sehr niedrig ausfallen, was einen deutlichen Unterschied zur Sympathieebene darstellt. Auch der Schüler wird nur gering positiver eingeschätzt, was vor allem an der Gruppe „100–200“ liegt, die den Schüler deutlich positiver einschätzt.

### 9.6.2 Generelle Ergebnisse

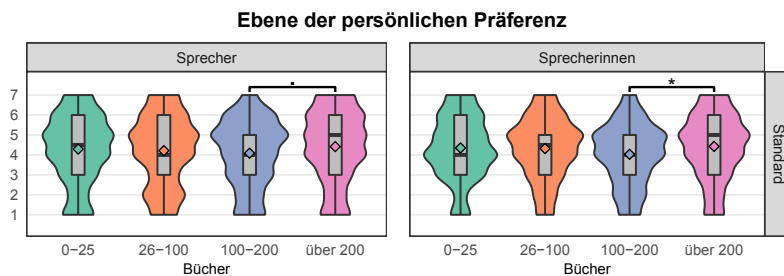
Bisher zeichnete sich noch kein deutliches Bild in Bezug auf Unterschiede zwischen Rater:innen mit verschiedenem Bildungshintergrund ab, eventuell sind die generellen Ergebnisse etwas aufschlussreicher. Wie es sich schon in den detaillierten Ergebnissen angebahnt hat, findet man auf Kompetenzebene keine signifikanten Unterschiede vor – die Einschätzungen sind also über die einzelnen Gruppen hinweg auf dieser Ebene recht konstant. Daher kann auch  $H_{8n}$ : „Es existiert kein Unterschied in der Kompetenzeinschätzung von Schüler:innen mit niedrigem oder hohem Bildungshintergrund in Bezug auf dialektsprechende Personen“ nicht abgelehnt werden.

Auf Sympathieebene zeigen sich bei den standardsprechenden Sprechern Unterschiede, die sich nahe an der Signifikanzgrenze befinden, und bei den standardsprechenden Sprecherinnen welche, die signifikant sind (s. Abb. 9.64). Die Ergebnisse sehen für beide Geschlechter ähnlich aus. Bei den Sprechern lassen sich abnehmende Sympathiewerte mit zunehmendem Bildungshintergrund erkennen. Die Gruppe „über 200“ ist in ihren Einschätzungen allerdings wieder ähnlich hoch wie die Gruppe „0–25“. Bei den Sprecherinnen liegt die Sachlage etwas anders. Hier schätzt die Gruppe „100–200“ die Sprecherinnen deutlich weniger sympathisch ein als die restlichen Gruppen, die recht ähnliche Ergebnisse zeigen. Zu den Gruppen „0–25“ und „über 200“ sind die Unterschiede auch signifikant. Da dieses Ergebnis recht unerwartet ist und keine klar steigende bzw. abfallende Linie von niedrigstem zu höchstem Bildungshintergrund zeigt, ist auch die Frage schwierig, ob man  $H_{9n}$ : „Es existiert kein Unterschied in der Sympathieeinschätzung von Schüler:innen mit niedrigem oder hohem Bildungshintergrund in Bezug auf standardsprechende Personen“ ablehnen kann oder nicht.



**Abbildung 9.64:** Vergleich des Bildungshintergrunds der allgemeinen Einschätzung der Sympathie bei standardsprachlicher Sprechweise

Da signifikante Unterschiede gefunden wurden, kann man sie prinzipiell ablehnen,  $H_9$ : „Schüler:innen mit hohem Bildungshintergrund schätzen standardsprechende Personen sympathischer ein als Schüler:innen mit niedrigem Bildungshintergrund“ wird dadurch jedoch nicht bestätigt.



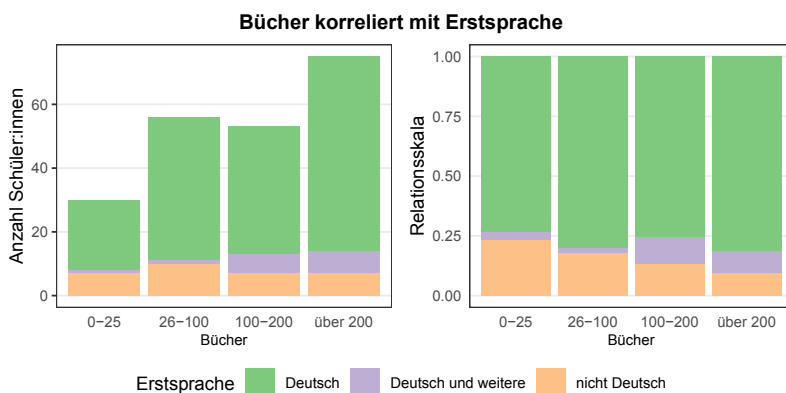
**Abbildung 9.65:** Vergleich des Bildungshintergrunds der allgemeinen Einschätzung der persönlichen Präferenz bei standardsprachlicher Sprechweise

Auf Ebene der persönlichen Präferenz (Abb. 9.65) fällt auf, dass die Einschätzungswerte niedriger ausfallen als auf Sympathieebene. Beim Sprecher sieht man das gleiche Muster, das auf Sympathieebene zu erkennen war. Mit zunehmendem Bildungshintergrund werden die Einschätzungen negativer. Auch hier zeigt die Gruppe mit dem höchsten Bildungshintergrund die höchsten Einschätzungswerte. Der Unterschied zur vorhergehenden Gruppe befindet sich auch in der Nähe der Signifikanz. Bei den Sprecherinnen zeigt sich ebenfalls, dass die Gruppe mit höchstem Bildungshintergrund die höchsten Einschätzungswerte aufweist. Die Gruppe „100–200“ zeigt wie so oft die niedrigsten Werte und der Unterschied dieser beiden Gruppen ist signifikant. Die beiden anderen Gruppen liegen zwischen den Gruppen, wobei auffällig ist, dass die Gruppe „26–100“ die standardsprechende Sprecherin positiver einschätzt als den Sprecher.

### 9.6.3 Zwischenfazit Bildungshintergrund

Die Spracheinstellungen nach Bildungshintergrund zu analysieren, bringt nur zum Teil erwartbare Ergebnisse. Diese sieht man etwa bei der dialekt sprechenden Lehrerin, die sowohl auf Kompetenz- als auch auf Sympathieebene in informeller Situation von den Personen mit niedrigstem Bildungshintergrund am positivsten eingeschätzt wird. Alle anderen Unterschiede beschränken sich auf die Verwendung der Standardsprache. Erwartbar ist das Ergebnis, dass die Personen, die den höchsten Bildungshintergrund angeben, die standardsprechenden Schüler:innen oder Lehrkräfte sehr positiv einschätzen. Eigenwillig dabei ist allerdings, dass signifikante Unterschiede oft zur Gruppe „100–200“, also der Gruppe mit dem zweithöchsten Bildungshintergrund, vorliegen – eine Tendenz, die nicht ganz ins Bild passen will. Ebenfalls ist auffällig, dass die Gruppe mit dem niedrigsten Bildungshintergrund häufig die standardsprechenden Personen ebenso positiv einschätzt wie die Personen, die den höchsten Bildungshintergrund angeben. Es lässt sich also keine lineare Zunahme mit aufsteigendem Bildungshintergrund erkennen, wie es erwartbar gewesen wäre. Dieser Trend ist nur ein einziges Mal in Abbildung 9.59 aufgetreten. In den generellen Ergebnissen sieht man eher den umgekehrten Trend.

Dass die Gruppe „0–25“ in ihrem Abstimmungsverhalten zur Standardsprache abweicht, könnte mit Folgendem zusammenhängen: Personen mit einer anderen Erstsprache als Deutsch schätzen dialektale Sprechweise häufig negativer und standardsprachliche positiver ein, als dies Erstsprachler:innen tun (s. Abschnitt 9.5). Wenn diese Gruppe nun ebenfalls häufiger einen geringen Bildungshintergrund aufweist, könnte dies das Abstimmungsverhalten erklären. Sehen wir uns dafür Abbildung 9.66 an.



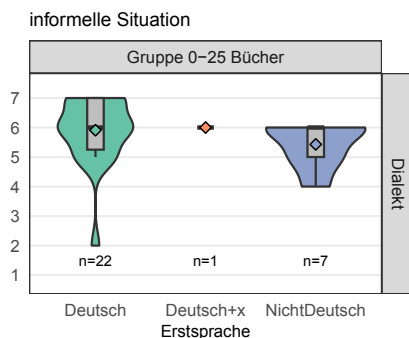
**Abbildung 9.66:** Korrelation Bücher mit Erstsprache in absoluter und relativer Darstellung

In den Säulendiagrammen ist auf der X-Achse der Bildungshintergrund in den schon bekannten Kategorien dargestellt. In den Säulen selbst wurden mit unterschiedlicher Farbgebung die erstsprachlichen Kategorien markiert. Das linke Diagramm zeigt die

absoluten Zählraten, das rechte die relativen, also eine Verhältnisskala zwischen den einzelnen Bildungskategorien.

Sowohl der absolute als auch der relative Plot zeigen, dass es einen negativen Zusammenhang mit geringer Effektstärke zwischen Bildungshintergrund und Personen mit einer anderen Erstsprache als Deutsch zu geben scheint. Ein gemischtes Modell mit Spracheinstellungen als Responsevariable, Teilnehmer:innen und Items als *Random Intercepts* und Bücher sowie Erstsprache als interagierende fixe Effekte (Bücher wurde dabei in eine numerische Variable umgeformt) bestätigt die negative Korrelation der Personen mit einer anderen Erstsprache als Deutsch und dem Bildungshintergrund mit  $p < 0.05$ . Die Gruppe der Personen mit einer anderen Erstsprache als Deutsch hat mit fast 25 % einen nicht unwesentlichen Einfluss auf die Gruppe mit dem niedrigsten Bildungshintergrund. Dies könnte eine Erklärung dafür sein, dass diese Gruppe standardsprechende Personen höher einschätzt, als dies erwartbar wäre. Allerdings erklärt es nicht, warum diese Gruppe die dialekt sprechende Lehrerin auf Kompetenz- und Sympathieebene (s. Abb. 9.55 und 9.56) so positiv einschätzt – die Einschätzungen müssten ja nach vorliegender Forschungslage genau andersherum sein.

#### Wie gut erledigt die Lehrerin ihre Arbeit?



**Abbildung 9.67:** Erstsprachvergleich bei der Frage nach der Kompetenz der dialekt sprechenden Lehrerin in informeller Situation der Gruppe mit dem niedrigsten Bildungshintergrund

Um diesen Umstand zu klären, ist es hilfreich, sich nur die Einschätzungswerte der Gruppe „0–25“ von Abbildung 9.55 anzusehen. Dies ist in Abbildung 9.67 realisiert. Auf der x-Achse wurden die Einschätzungen durch die unterschiedlichen Erstsprachen getrennt. Man erkennt, dass die Gruppe mit Deutsch als L1 deutlich höhere Einschätzungswerte vergibt als die Gruppe mit Deutsch als L2. Die positive Einschätzung der dialekt sprechenden Lehrerin stammt also von ersterer Gruppe.

## 10 Diskussion der erhobenen Spracheinstellungen

Bei der generellen Einschätzung der dialektalen und der standardsprachlichen Guises bestätigen sich auch in dieser Studie die schon bekannten Tendenzen aus der Spracheinstellungsforschung in Österreich, wie sie z. B. Soukup (2009), Bellamy (2012), Fuchs/Elspaß (2019) und Ender (2020) darlegen. Die Ergebnisse zeigen im Wesentlichen, dass standardsprachliche Sprechweise als kompetenter und dialektale als sympathischer eingeschätzt wird – beide Ergebnisse werden mit höchster Signifikanz auf die Gesamtpopulation prognostiziert, was für eine hohe Reliabilität und Replizierbarkeit spricht.

Bei genauerer Analyse dieser übergeordneten Tendenzen sind aber auch Muster zu erkennen, die auf diese Weise noch nicht in der Forschung berichtet wurden. Im Speziellen sind hier die Unterschiede zwischen den Geschlechtern hervorzuheben, von denen es einige zu berichten gibt.

Primär findet man einen erheblichen Genderbias zugunsten der Sprecher. So werden z. B. die Lehrer auf allen Ebenen und Situationen sowohl in dialektaler als auch standardsprachlicher Sprachverwendung höchstsignifikant positiver eingeschätzt. Bei den Schüler:innen ist dies nicht in diesem Ausmaß der Fall. Auf Kompetenzebene liegen beispielsweise die Einschätzungen in formeller Situation und Standardsprache zwischen den Geschlechtern gleichauf – in den anderen drei Ergebnissen dieser Ebene wird jedoch ebenfalls der Schüler kompetenter eingeschätzt.

Auf Kompetenzebene findet man aber auch Ähnlichkeiten zwischen den Geschlechtern. So werden beide bei standardsprachlicher Sprachverwendung kompetenter eingeschätzt. Dies ändert sich, wenn man die anderen Ebenen betrachtet. So werden die Sprecher erwartbar sympathischer geratet, wenn sie Dialekt sprechen, bei den Sprecherinnen liegen die Sympathiewerte beider Varietäten jedoch gleichauf, was so noch nicht in der bisherigen Forschung berichtet wurde. Auf Ebene der persönlichen Präferenz spitzt sich die Situation sogar noch zu. So werden die Sprecher zwar positiver bei dialektaler Sprechweise eingeschätzt, Sprecherinnen aber bei standardsprachlicher. Dieses Bild wird noch zusätzlich verstärkt, wenn man sich die Einschätzungen zur Sprachverwendung zwischen den Geschlechtern ansieht. Die Einschätzungen der Schüler:innen liegen hier auf Sympathieebene und Ebene der persönlichen Präferenz mit großem Abstand über den Werten der Schüler. Dabei übertreffen ihre Werte auch die der dialektsprachlichen Guises! Dies deutet darauf hin, dass das stereotype Bild „Dialektsprechende sind sympathischer“ nur bei den Sprechern gilt, bei Sprecherinnen ist die Lage differenzierter und fraglich, ob man überhaupt von diesem Einschätzungsmuster ausgehen kann. Eine Erklärung hierfür könnte die Ethnolinguistic Identity Theory (Giles/Johnson 1987) liefern: Sprecherinnen scheinen ja generell standardnäher zu kommunizieren (Trudgill 1972; Wardhaugh/Fuller 2015; Brouwer 1989). Dies könnte damit zusammenhängen, dass speziell in urbanen Kontexten Frauen seit mehreren

Jahrzehnten sich um Prestige- und Machtpositionen bemühen, die vorher Männern vorbehalten waren. Um dies zu erreichen, müssen sich Frauen besonders gruppenkonform verhalten. Dieses Verhalten steht jedoch in direkter Verbindung mit prestigehoher Sprachverwendung, da diese mit Kompetenz in Verbindung gebracht wird und sich dies als wichtige Möglichkeit für die Minderheitengruppe erweist, sich als kompetent darzustellen. Dadurch würde Dialektverwendung im urbanen Raum bei Frauen in gesellschaftliche Bereiche zurückgedrängt, in denen dialektale Sprachverwendung stark mit regionaler Identität und „ethnischer Solidarität“ in Verbindung steht, also in die Agrar- und handwerklichen Sektoren. Dadurch wird dialektale Sprachverwendung bei Frauen als noch weniger kompetent eingeschätzt, was wieder dazu führt, dass Frauen, die sich um Prestigeposten bemühen, diese Varietät meiden usw. – ein Teufelskreis also. Dies würde auch eine Erklärung dafür liefern, warum speziell bei den Lehrerinnen die Unterschiede im Rating zwischen den Varietäten so viel stärker ausfallen als bei den Lehrern.

Interessant in diesem Zusammenhang ist auch der Unterschied im Ratingverhalten der Geschlechter. So lassen sich einerseits in allen Analysen bei den Raterinnen höhere Zustimmungswerte ausfindig machen, ein Ergebnis, auf das auch Soukup (2009) und Bellamy (2012) stießen. Neben diesem Ergebnis findet sich andererseits, dass die Raterinnen sowohl in dialektaler als auch standardsprachlicher Sprechweise Sprecher auf Kompetenzebene höher einschätzen als die Rater ihr eigenes Geschlecht. Bei der Einschätzung der Sprecherinnen tritt dieser Ratingunterschied nicht auf. Die Konsequenz davon ist, dass die Raterinnen sogar stärker dafür verantwortlich sind, dass dieser Teufelskreis erhalten bleibt und die Sprachverwendung bei Sprecherinnen eine stärkere Distanz zwischen den Varietäten aufweist als diejenige der Sprecher. Eine Erklärung für dieses Verhalten stellt der *Confirmation Bias* (Klayman 1995) bereit. Dabei handelt es sich um den Effekt, dass Informationen auf die Art und Weise verarbeitet und gespeichert werden, dass sie die eigenen Erwartungen erfüllen. Dieses Ergebnis würde so betrachtet implizit darauf hindeuten, dass Schülerinnen heutzutage immer noch davon ausgehen, dass sie wegen ihres Geschlechts weniger kompetent sind als Schüler.<sup>27</sup>

Zusätzlich widersprechen die Ergebnisse dieser Studie zum Ratingverhalten der Geschlechter denjenigen, die Brouwer (1989) im niederländischen Kontext fand. So werden die standardsprechenden Sprecherinnen in ihrer Studie zwar sympathischer eingeschätzt als die Sprecher, diese Einschätzung geht jedoch sowohl von Ratern als auch von Raterinnen und eben nicht nur von Letzteren aus. Bei der positiveren Einschätzung der dialektsprechenden Sprecher ist das Ergebnis sogar diametral unterschiedlich: Sie werden eben von den Raterinnen noch positiver eingeschätzt als von den Ratern. Ob diese divergierenden Ergebnisse vom unterschiedlichen sprachlichen Hintergrund, der unterschiedlichen Stichprobe oder einem anderen Grund herrühren, kann nicht eruiert werden. Hier wäre mehr differenzierte Geschlechterforschung wünschenswert, die Sprachverwendung in Kombination mit Spracheinstellungen unter die Lupe nimmt.

---

27 Dabei wird jedoch differenziert. So erhält die standardsprechende Schülerin ja die höchsten Kompetenzeinschätzungen, im Erwachsenenalter werden die Lehrer jedoch in beiden Varietätsverwendungen und Situationen kompetenter eingeschätzt.

Die Analysen etwaiger Stadt-Land-Unterschiede erbringen kontraintuitive Ergebnisse. So lautete die Anfangshypothese H<sub>5</sub>, dass in ländlichen Regionen außerhalb der Pendeldistanz dialektale Sprachverwendung positiver auf Kompetenzebene eingeschätzt wird als in der Stadt. Dies konnte nicht bestätigt werden. Auch auf Sympathieebene zeigen sich unerwartete Ergebnisse. So wird am Land dialektale Sprachverwendung sogar am wenigsten sympathisch eingeschätzt. Eine Erklärung hierfür könnte darin liegen, dass die verwendeten Dialektguises in einer Varietät des Salzburger Zentralraums eingesprochen wurden. Diese unterscheidet sich von den Dialekten des Pinzgaus und wird dadurch potenziell als Outgroup-Varietät eingeschätzt. Durch die Differenzmaximierung, die die Social Identity Theory beschreibt (Tajfel/Turner 1979), kommt dann die Schlechterbewertung zustande. Allerdings wird diese Situation für die meisten Schüler:innen eher der Normalfall sein, da viele Lehrkräfte aus diesem Raum stammen bzw. diese Varietät durch das Studium erworben haben. Daher würde ebenfalls der *Confirmation Bias* eine weitere mögliche Erklärung liefern. Der Pinzgauer Dialekt wird durch manche seiner salienten Merkmale im Salzburger Zentralraum sicherlich als stark dialektal wahrgenommen. Diese Distanz zur in Salzburg gesprochenen Standard- und Umgangssprache führt eventuell dazu, dass die Pinzgauer:innen ihre Dialekte als besonders stigmatisiert wahrnehmen, die Stereotype dadurch besonders verinnerlichen und diese Wahrnehmung in einem weiteren Schritt generell auf Dialekte ausdehnen.

Gleichzeitig schätzen die Pinzgauer Rater:innen auch die Standardsprache am wenigsten sympathisch ein. Dies könnte daran liegen, dass diese Varietät im Alltag nur selten angewendet wird und damit auch keine große Rolle in dieser Gegend spielt. Interessant ist überdies, dass es keinen Stadt-Land-Unterschied auf Kompetenzebene der Sprachverwendung zu geben scheint. Die Einstellungen auf dieser Ebene scheinen sehr homogen verteilt zu sein.

Diese gleiche Homogenität der Einschätzung auf Kompetenzebene findet man auch, wenn man sich die Spracheinstellungen der Schüler:innen mit unterschiedlichen Erstsprachen ansieht. Auch hier findet man nur auf Sympathieebene und Ebene der persönlichen Präferenz Unterschiede. Am auffälligsten ist hierbei, dass die Geschlechterunterschiede, die auf diesen beiden Ebenen messbar sind – Dialektsprecher werden sympathischer eingeschätzt als Dialektsprecherinnen – vor allem bei Schülerinnen mit Deutsch als L1 gemacht werden. Die Schülerinnen mit Deutsch als L2 zeigen diese Bewertungsunterschiede nicht auf. Dass sie die Sprecher ähnlich negativ einschätzen wie die Sprecherinnen, führt dazu, dass sie generell niedrigere Einschätzungswerte auf den beiden Ebenen aufweisen. Damit bestätigen sich die Ergebnisse von Baßler/Spietkermann (2001) und Ender/Kasberger/Kaiser (2017), allerdings werden sie bereichert um die Geschlechterkomponente. Ein Desiderat wäre, zu überprüfen, ob die unterschiedlichen Spracheinstellungen mit Dialekt- bzw. standardsprachlicher Kompetenz korrelieren, wie es Ender (2020) gezeigt hat. Dieser Schritt wird in der Dissertation nicht unternommen, ist aber in einer weiteren Publikation in Bearbeitung.

Bei der Betrachtung der unterschiedlichen Bildungshintergründe wird die dialekt-sprechende Lehrerin in informeller Situation von Personen, die den niedrigsten Bil-

dungshintergrund angaben, am kompetentesten und am sympathischsten eingeschätzt. Daneben zeigt sich, dass die Personen, die den höchsten Bildungshintergrund angaben, standardsprechende Personen sehr positiv einschätzen – beides also erwartbare Ergebnisse.

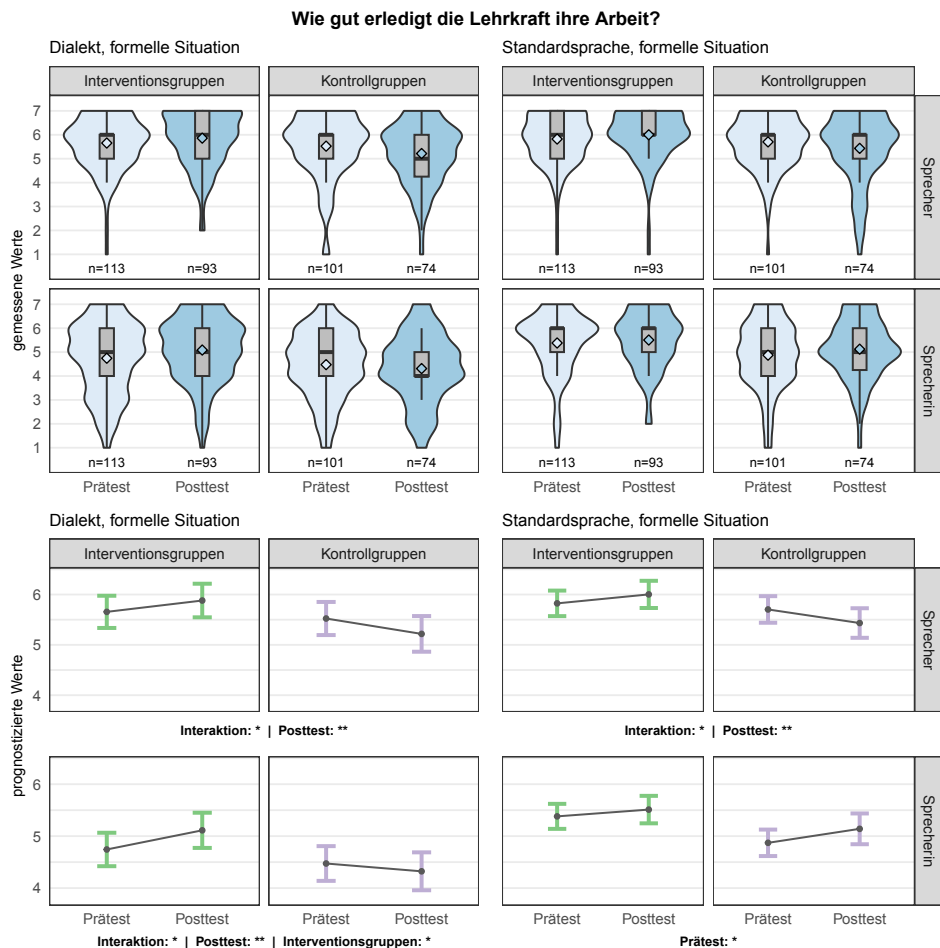
# 11 Lassen sich die Spracheinstellungen verändern?

Kommen wir nun zur Beantwortung der zentralen Fragestellung der Arbeit, den Ergebnissen der Interventionsstudie. Dafür ist es im Vergleich zu den Prätestergebnissen nötig, zwei weitere Variablen in die Datenbeschau miteinzubeziehen: die Gruppenvariable mit den Ausprägungen Interventions- und Kontrollgruppen sowie die Zeitvariable mit den beiden Messzeitpunkten Prätest und Posttest. Dies vermehrt die Menge der Plots, die es benötigt, eine Ebene in einer Situation (z. B. „formelle Situation“ und „Kompetenzebene“) abzubilden. Zusätzlich zu den deskriptiven Plots und den gemischten Modellen werden in Abschnitt 11.2 und 11.3 auch Distanzplots mit den zugehörigen Modellen präsentiert (s. für eine Erklärung der Plots und Modelle Abschnitt 8.5). Generell wird aus Platzgründen wie in Kapitel 9 im Großen und Ganzen darauf verzichtet, nicht signifikante Ergebnisse zu präsentieren – dies geschieht nur, wenn sie von den Tendenzen her ins Bild passen und sich nahe an der Signifikanzgrenze befinden. Eine Gesamtübersicht über die erstellten Plots und Modelle findet man wie bisher auf meiner OSF-Seite <https://osf.io/7mhw7/>.

## 11.1 Allgemeine Prä-Posttest-Unterschiede

Um Veränderungen zwischen Prätest- und Posttestergebnissen mit Einbezug der Interventions- sowie der Kontrollgruppen analysieren zu können, werden die Ergebnisse für Dialekt-Standard-Unterschiede pro Rolle (Schüler:in, Lehrkraft), Situation (formell, informell) sowie Ebene (Kompetenzebene, Sympathieebene, Ebene der persönlichen Präferenz) präsentiert. Insgesamt ergibt das 12 Diagramme, die beide Geschlechter, beide Varietäten, beide Gruppen sowie die Testzeitpunkte beinhalten.

Sehen wir für eine nähere Erklärung der Diagramme Abbildung 11.1 etwas näher an. Im oberen Bereich der Darstellung sind Abstimmungsergebnisse wie üblich in Violin- und Boxplots sowie zusätzlich dem arithmetischen Mittel angegeben. Die beiden unterschiedlich gefärbten Diagramme pro Kästchen stellen die Prä- und Posttestergebnisse innerhalb einer Gruppe, einer Varietät und eines Geschlechtes dar. z. B. beschreibt das Kästchen links oben die Abstimmungsergebnisse der Interventionsgruppen zum dialektalen Sprachgebrauch des Lehrers. Hier ist zu erkennen, dass diese von Prä- zu Posttestzeitpunkt hin etwas gestiegen sind. Rechts daneben findet man die Abstimmungsergebnisse in den gleichen Kategorien der Kontrollgruppen, die zum Posttestergebnis hin niedriger ausfallen. Es ist nun wichtig, diese beiden Kästchen als Analyseinheit zu betrachten, da sie das zugrundeliegende Testdesign abbilden.



**Abbildung 11.1:** Prätest-Posttest-Vergleich bei der Frage nach der Kompetenz der Lehrkräfte in formeller Situation

Pro deskriptiver Abbildung sind also vier solcher Analyseeinheiten vorhanden. Innerhalb dieser Einheiten kann man prinzipiell fünf Fragen stellen:

1. Unterscheiden sich die Ergebnisse der Gruppen zueinander unter Einbezug der beiden Testzeitpunkte (= Interaktion)?
2. Unterscheiden sich Prä- und Posttest der Interventionsgruppen?
3. Unterscheiden sich Prä- und Posttest der Kontrollgruppen?
4. Unterscheiden sich die Prätestergebnisse der beiden Gruppen?
5. Unterscheiden sich die Posttestergebnisse der beiden Gruppen?

Diese Fragen wurden mit linearen gemischten Modellen überprüft (s. für die Einführung in gemischte Modelle Abschnitt 8.5.2). Ein Effektplot der Modelle im selben Schema der deskriptiven Darstellung ist dabei nützlich, um visuell einen Überblick

über die Signifikanz der Ergebnisse erhalten zu können. Daher wurden im unteren Teil der Diagramme die Berechnungsergebnisse von vier Modellen abgebildet, die sich in der gleichen Anordnung befinden wie die Datendarstellung in den deskriptiven Plots.

Für das Studiendesign ist der Vergleich der Interventions- und Kontrollgruppen ausschlaggebend. In statistischen Termini bedeutet dies, dass der Interaktionseffekt der beiden dichotomen Variablen *Erhebung:Gruppen* entscheidende Informationen für die Interpretation der Ergebnisse liefert. Dafür werden im Modell die unterschiedlichen Steigungen der Regressionslinien miteinander verglichen und daraus Signifikanzwerte berechnet. Wie schon in Abschnitt 8.5.2 ausgeführt, ist ja nicht nur aussagekräftig, wenn beide Testzeitpunkte der Interventionsgruppen signifikante Unterschiede zeigen, da speziell die Prätestergebnisse durch Einflussfaktoren wie die soziale Erwünschtheit verzerrt sein können. Die Interaktion stellt damit ein robustes Ergebnis dar, da sie die Abstimmungsergebnisse der beiden Gruppen in Relation zueinander stellt. Es werden darüber hinaus aber auch alle anderen der fünf möglichen Analysekatégorien aufgelistet, wenn sie sich in der Nähe der Signifikanzgrenze befinden oder diese überspringen. Zusätzlich ist bei Berechnung eines gemischten Modells noch der Modellvergleich ausschlaggebend, in dem überprüft wird, ob das Modell einen Vorteil zum Nullmodell aufweist. Diesen „Modelfit“ integriere ich in den Ergebnissen, wenn er keine signifikanten Vorteile zeigt. Ist er also nicht in den Effektplots integriert, zeigte die Berechnung des gemischten linearen Modells signifikante Vorteile zum Nullmodell. In Abbildung 11.2 wäre dies z. B. beim ersten Modell der Fall. Am Effektplot ist gut zu erkennen, dass es zwischen beiden Gruppen keine große Veränderung zwischen Prä- und Posttestzeitpunkt gibt. Da das Nullmodell genau von einer solchen Nichtveränderung ausgeht, ist es also logisch, dass das Modell mit den fixen Effekten keine Vorteile bringt.

### 11.1.1 Detaillierte Ergebnisse

Für die detaillierten Ergebnisse wurden Random-Intercept-Modelle gerechnet, bei denen die Rater:innen in den Klassen verschachtelt wurden. Nur bei manchen Testungen, bei denen die Klassenvariable ein Singularitätsproblem<sup>28</sup> verursachte, wurden einfachere Modelle mit Rater:innen als Random Intercept berechnet.

In formeller Situation (Abb. 11.1) sind auf Kompetenzebene bei den Lehrkräften folgende Veränderungen feststellbar: Der dialektspredende Lehrer wird von den Interventionsgruppen im Posttest in Relation zu den Kontrollgruppen signifikant kompetenter eingeschätzt. Auch die Posttestergebnisse der beiden Gruppen zeigen diese Tendenz: Sie unterscheiden sich hochsignifikant voneinander. Im Vergleich dazu sind die Prätestergebnisse nahezu gleichauf und liefern keine signifikanten Unterschiede.

Ähnlich verhält es sich bei der dialektspredenden Lehrerin. Hier liefert der Vergleich der Gruppen in Relation mit den Testzeitpunkten ebenfalls ein signifikantes Ergebnis und auch die Posttestergebnisse unterscheiden sich hochsignifikant voneinander. Hinzu kommt, dass auch die Interventionsgruppen selbst einen signifikanten

---

28 S. Sonderegger 2021: Kap. 10.4 für eine nähere Beschreibung und Lösungswege für dieses Problem.



In informeller Situation (s. Abb. 11.2) liegen beim dialektspredenden Lehrer keine für die Fragestellung relevanten Unterschiede vor. Es sind zwar sowohl die Prä- als auch die Posttestsituation beider Gruppen signifikant unterschiedlich, das bedeutet allerdings nur ein unterschiedliches Abstimmungsmuster der beiden Gruppen. Eine Wirkung der Intervention ist hier nicht herauszulesen.

Bei der dialektspredenden Lehrerin sieht das Bild anders aus. Die Interventionsgruppen schätzen sie nach der Intervention in Relation zu den Kontrollgruppen signifikant kompetenter ein. Auch die Posttestergebnisse unterscheiden sich höchstsignifikant, was für eine deutliche Wirkung der Intervention spricht.

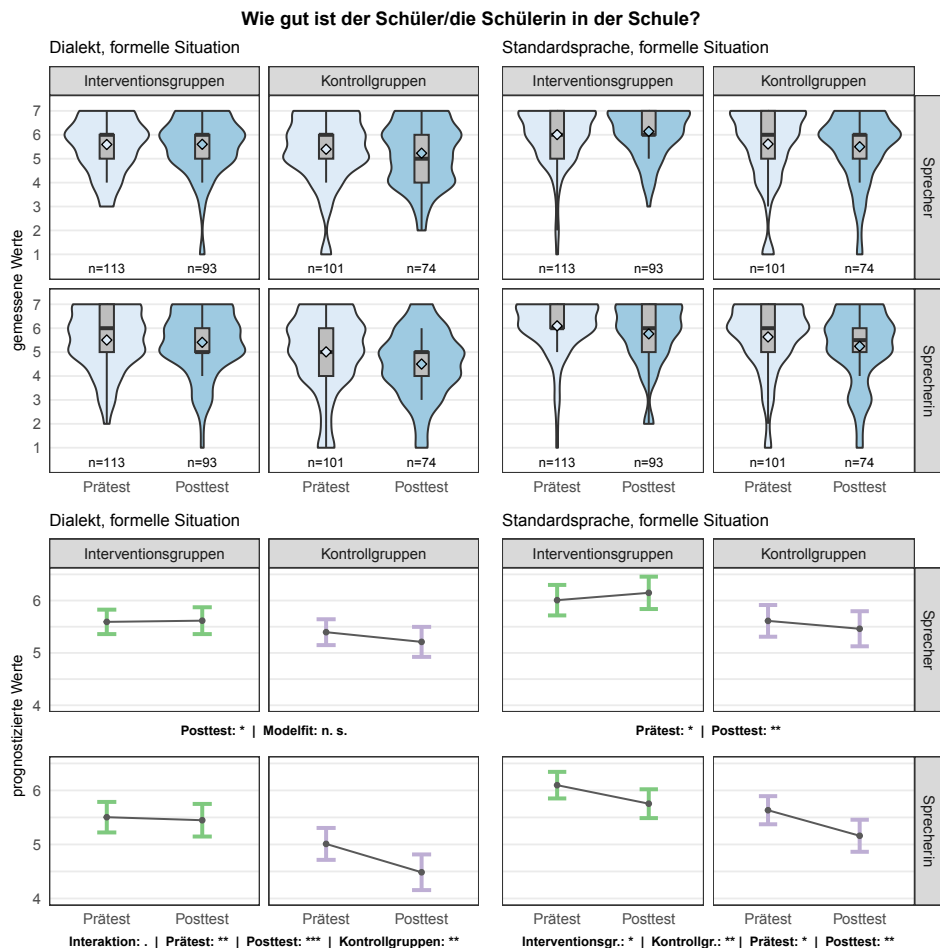
Der standardspredende Lehrer zeigt keine signifikanten Ergebnisse. Bei den Interventionsgruppen ist keine Veränderung zwischen den Testzeitpunkten bemerkbar. Die Kontrollgruppen schätzen ihn im Posttest zwar signifikant negativer ein als die Interventionsgruppen, allerdings zeigt der Modellvergleich keine Vorteile zum Nullmodell – eine Veränderung ist damit nicht herauszulesen.

Die dialektspredende Lehrerin wird von den Interventionsgruppen in Relation zu den Kontrollgruppen im Posttest signifikant kompetenter eingeschätzt. Die Posttestergebnisse unterscheiden sich zwischen den Gruppen sogar höchstsignifikant – eine Wirkung der Intervention in dieser Kategorie ist damit höchstwahrscheinlich. Auch die standardspredende Lehrerin wird im Posttest von den Interventionsgruppen in Relation zu den Kontrollgruppen kompetenter eingeschätzt – die Ergebnisse befinden sich allerdings nur in der Nähe der Signifikanzgrenze. Der Posttestzeitpunkt verstärkt diese Tendenz allerdings, da nun zwischen den Gruppen höchstsignifikante Unterschiede in den Einschätzungen vorliegen.

Bei den Schüler:innen sind die Ergebnisse weniger stark ausgeprägt. In formeller Situation (s. Abb. 11.3) wird der dialektspredende Schüler in der Posttestung von den Interventionsgruppen recht ähnlich und den Kontrollgruppen etwas negativer eingeschätzt. Die Posttestwerte selbst unterscheiden sich auch signifikant voneinander, allerdings zeigt der Modellfit keine signifikanten Werte. Eine Wirkung der Intervention kann also nicht angenommen werden.

Auch bei der dialektspredenden Schülerin kommen unerwartete Werte zum Vorschein. In den Interventionsgruppen zeigt die Kompetenzeinschätzung im Posttest sogar etwas nach unten, eine Tendenz, die bei den Kontrollgruppen noch sehr viel stärker ausgeprägt ist, sodass sich die Prä-Posttest-Werte in Relation der beiden Gruppen sogar in der Nähe der Signifikanzgrenze befinden. Deutlich wird der Unterschied auch darin, dass die Unterschiede von hochsignifikant im Prätest zu höchstsignifikant im Posttest zunehmen.

Der standardsprachliche Schüler wird von den Interventionsgruppen im Prätest signifikant kompetenter eingeschätzt als von den Kontrollgruppen. Dieser Einschätzungsunterschied nimmt im Posttest weiter zu und erreicht hochsignifikante Messwerte.

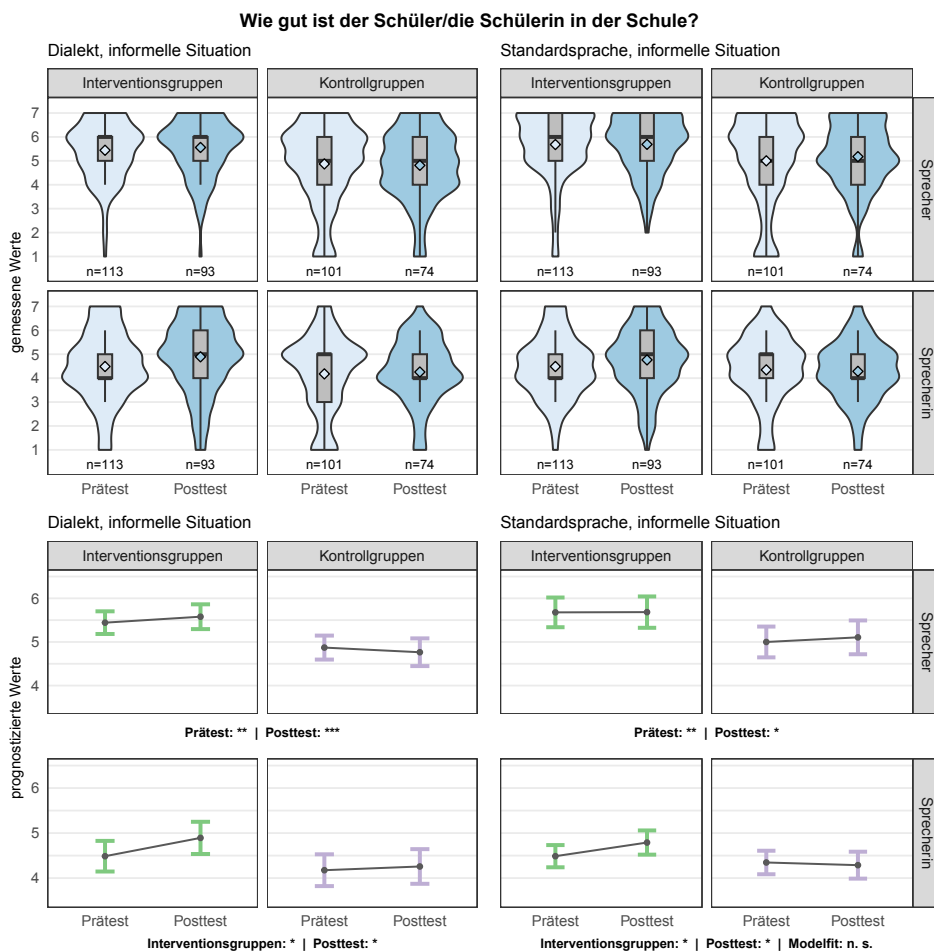


**Abbildung 11.3:** Prätest-Posttest-Vergleich bei der Frage nach der Kompetenz der Schüler:innen in formeller Situation

Die standardsprechende Schülerin wird im Posttest von den Interventionsgruppen signifikant und den Kontrollgruppen hochsignifikant weniger kompetent eingeschätzt – der Abfall der Regressionslinie ist in der Kontrollgruppe steiler. Zwischen den Gruppen besteht ein signifikanter Prä- und hochsignifikanter Posttestunterschied. Das unerwartete Ergebnis kann folgendermaßen interpretiert werden: In den Interventionsgruppen fällt das Prätestergebnis für die standardsprechende Schülerin sehr positiv aus – sie wird höchstkompetent eingeschätzt, mit Werten über sechs auf der siebenteiligen Skala. Die Intervention könnte dazu geführt haben, dass sie im Posttest reflektierter eingeschätzt wurde, was eine leichte Reduktion der Werte bedeutet. In den Kontrollgruppen ist hingegen nur die generelle Tendenz bemerkbar, in der Posttestung negativer abzustimmen als in der Prätestung.

In informeller Situation (s. Abb. 11.4) zeigt sich beim dialektsprechenden Schüler, dass er im Prätest von den Interventionsgruppen hochsignifikant kompetenter eingeschätzt wird als von den Kontrollgruppen. Im Posttest nimmt die Differenz dann weiter zu, sodass die Messergebnisse auch höchste Signifikanz erreichen.

Die dialektsprechende Schülerin wird im Posttest von den Interventionsgruppen signifikant kompetenter eingeschätzt als im Prätest. In den Kontrollgruppen ist dagegen keine große Bewegung bemerkbar. Die Posttestergebnisse beider Gruppen zeigen daher ebenfalls signifikante Unterschiede.

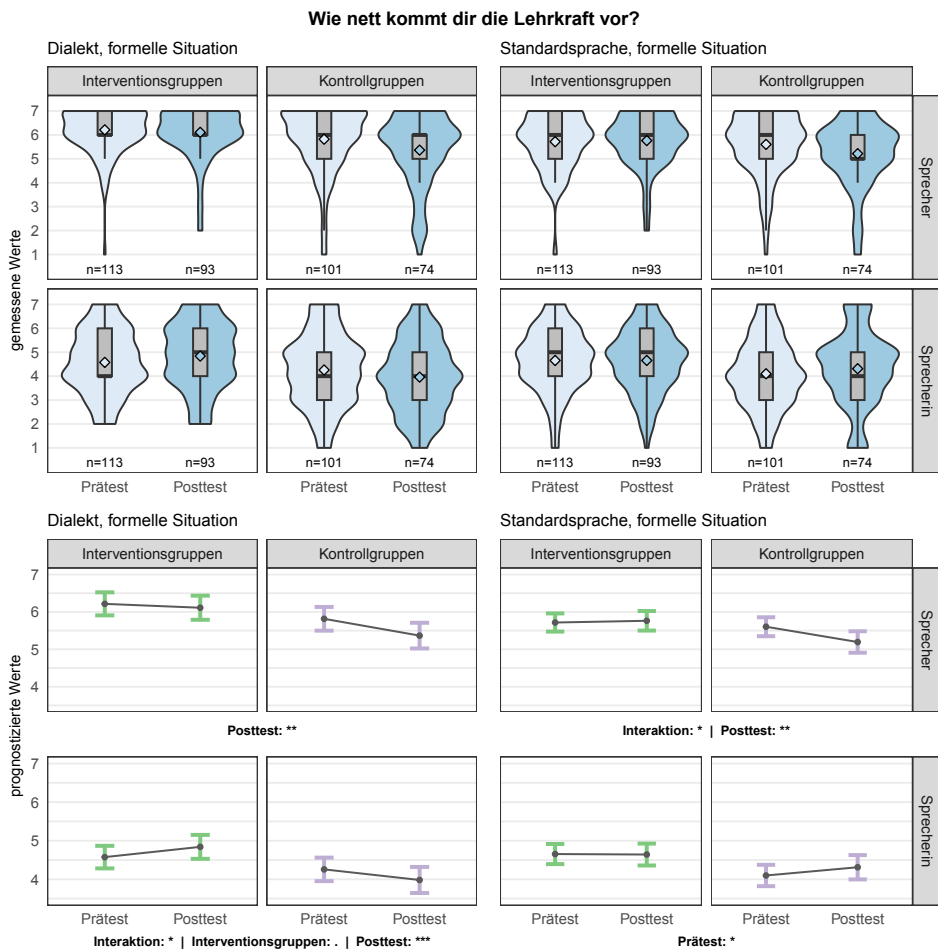


**Abbildung 11.4:** Prätest-Posttest-Vergleich bei der Frage nach der Kompetenz der Schüler:innen in informeller Situation

Der standardsprechende Schüler wird sowohl von Interventions- als auch Kontrollgruppen recht ähnlich bei den beiden Testzeitpunkten eingeschätzt. Zwischen den Gruppen existieren in der Prätestung hochsignifikante und der Posttestung signi-

fikante Unterschiede, die jedoch keine Wirkung der Intervention andeuten, sondern nur aufzeigen, dass die Kontrollgruppen den Schüler weniger kompetent einschätzen.

Auf Sympathieebene sind ebenfalls Auffälligkeiten zwischen den Testzeitpunkten feststellbar. In formeller Situation (s. Abbildung 11.5) schätzen die Interventionsgruppen den dialektsprechenden Lehrer im Posttest hochsignifikant sympathischer ein als die Kontrollgruppen, wobei in den Prätestergebnissen keine signifikanten Unterschiede vorliegen. Dabei bleibt die Einschätzung in den Interventionsgruppen annähernd gleich, jedoch befinden sich die Werte selbst auf sehr hohem Niveau.

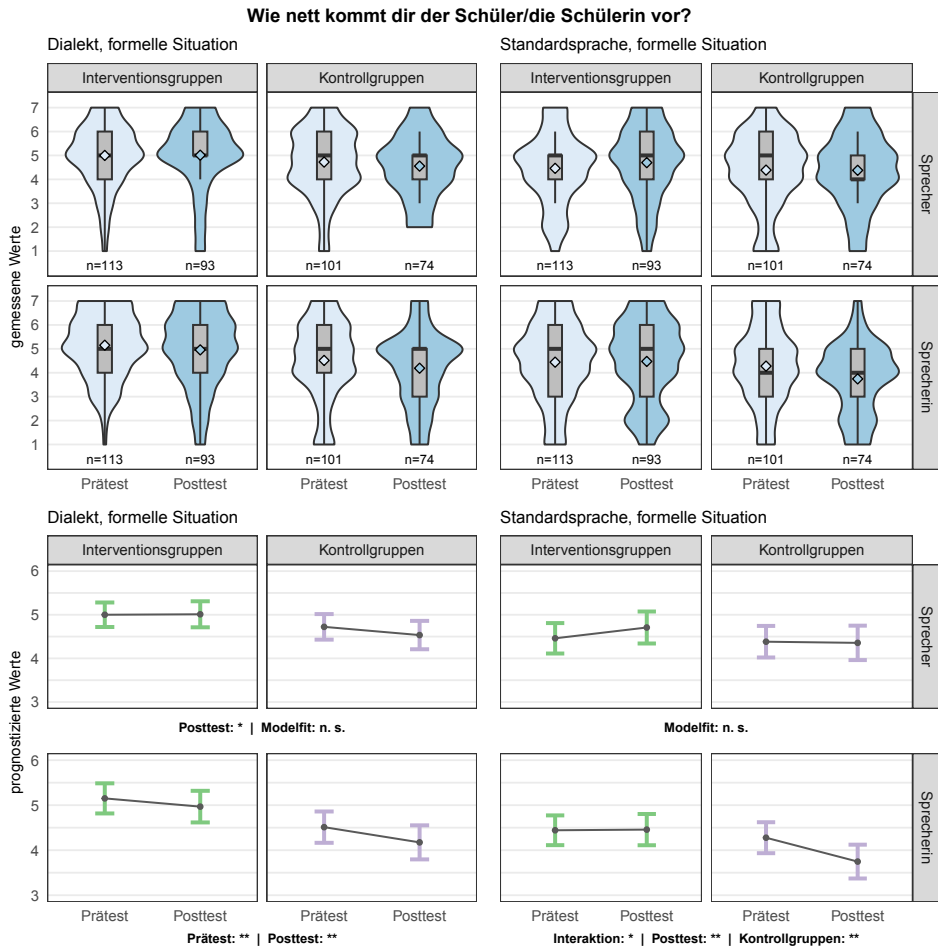


**Abbildung 11.5:** Prätest-Posttest-Vergleich bei der Frage nach der Sympathie der Lehrkräfte in formeller Situation

Die dialektsprechende Lehrerin wird von den Interventionsgruppen im Posttest signifikant sympathischer in Relation zu den Kontrollgruppen eingeschätzt. Auch der Anstieg der Werte der Interventionsgruppen zwischen den Testzeitpunkten liegt in der



In informeller Situation (s. Abb. 11.6) sind praktisch keine Veränderungen beim dialektsprechenden Lehrer zu sehen. Sowohl Prä- als auch Posttest weisen zwar signifikante Unterschiede der beiden Gruppen auf, was für die Untersuchung aber nicht relevant ist.



**Abbildung 11.7:** Prätest-Posttest-Vergleich bei der Frage nach der Sympathie der Schüler:innen in formeller Situation

Die dialektsprechende Lehrerin wird nach erfolgter Intervention in Relation zu den Kontrollgruppen von den Interventionsgruppen hochsignifikant sympathischer eingeschätzt, wenn man den Abstimmungsbias (soziale Erwünschtheit in der Prätestung)

miteinbezieht.<sup>29</sup> In der Prätestsituation schätzen die Interventionsgruppen die Lehrerin signifikant sympathischer ein, in der Posttestsituation steigt der Unterschied auf höchstsignifikantes Niveau. Beide Ergebnisse lassen die Wirksamkeit der Intervention in dieser Kategorie sehr wahrscheinlich erscheinen.

Ähnlich sieht auch das Ergebnis beim standardsprechenden Lehrer aus. Er wird im Posttest von den Interventionsgruppen in Relation zu den Kontrollgruppen hochsignifikant sympathischer eingeschätzt, was sich ebenfalls an den hochsignifikant unterschiedlichen Posttestergebnissen zeigt.

Bei der standardsprechenden Lehrerin ist eine ähnliche Tendenz wie bei der dialektspredenden Lehrerin feststellbar. Sie wird von den Interventionsgruppen im Posttest in Relation mit den Kontrollgruppen etwas sympathischer eingeschätzt. Das Ergebnis befindet sich jedoch nur nahe der Signifikanzgrenze. Der Unterschied der Posttestwerte ist jedoch signifikant.

Bei den Schüler:innen sind die Ergebnisse schwächer ausgeprägt als bei den Lehrkräften. In formeller Situation (Abb.11.7) sind beim standardsprechenden Schüler praktisch keine Unterschiede feststellbar. Die Interventionsgruppen schätzen ihn im Posttest sympathischer ein als die Kontrollgruppen. Der Modellfit ist jedoch nicht signifikant, daher kann nicht von einer Veränderung durch die Intervention ausgegangen werden.

Die dialektspredende Schülerin wird von den Interventionsgruppen sowohl im Prätest als auch im Posttest hochsignifikant sympathischer eingeschätzt. Da beide Gruppenwerte abnehmen, kann auch hier nicht von einer Wirkung der Intervention ausgegangen werden.

Etwas anders sieht es beim standardsprechenden Schüler aus. Die Interventionsgruppen zeigen eine Tendenz zur Besserbewertung im Posttest. Allerdings zeigt der Modellfit keine signifikanten Werte – man kann also ebenfalls nicht von einer Veränderung durch die Intervention ausgehen.

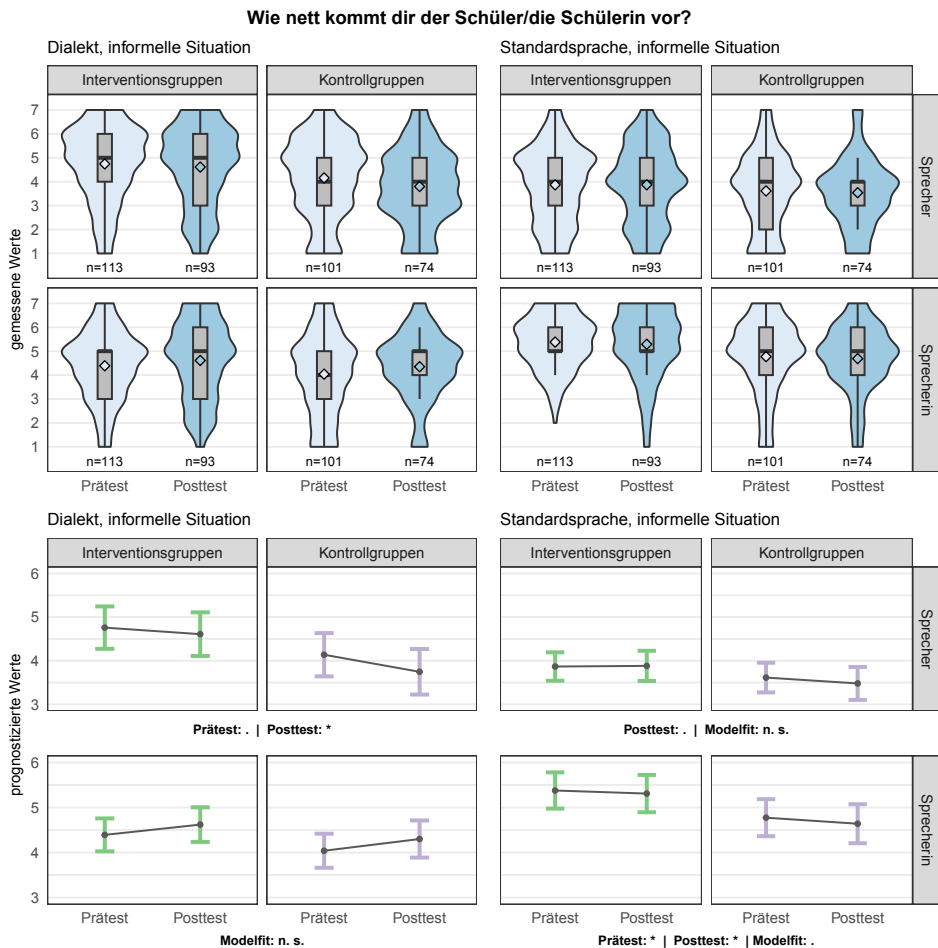
Die standardsprechende Schülerin wird im Posttest von den Interventionsgruppen in Relation zu den Kontrollgruppen signifikant sympathischer eingeschätzt, wobei dies vor allem an den hochsignifikant sinkenden Einschätzungswerten der Kontrollgruppen liegt. Die Posttestergebnisse der beiden Gruppen selbst zeigen ebenfalls hochsignifikante Unterschiede – bei Miteinbeziehung des Social Desirability Bias kann eine Wirkung der Intervention schon angenommen werden.

In informeller Situation (Abb.11.8) findet man beim dialektspredenden Schüler einen Rückgang der Werte sowohl in den Interventions- als auch den Kontrollgruppen – bei Letzteren etwas ausgeprägter. Dabei zeigen die Kontrollgruppen in beiden Testzeitpunkten niedrigere Werte: Im Prätest befindet sich der Unterschied nahe der Signifikanzgrenze, im Posttest ist er dann signifikant. Das Ergebnis ist wenig aussagekräftig.

---

29 Der Effekt der sozialen Erwünschtheit (z.B. Krumpal 2013) könnte dafür verantwortlich sein, dass die Antworten der Rater:innen aus Verlangen nach sozialer Zustimmung in den Prätestungen positiver ausfallen, als sie eigentlich sind. In der zweiten Testung stellte für die Rater:innen das Verfahren der Testung kein Novum mehr dar und war, speziell für die Kontrollgruppen, womöglich langweilig. Dies würde eine Erklärung dafür liefern, warum die Kontrollgruppen in der Posttestung negativer abgestimmt haben und man davon ausgehen könnte, dass auch die Prätestergebnisse der Interventionsgruppen zu hoch sind.

Die dialektsprechende Schülerin wird sowohl von den Interventions- als auch den Kontrollgruppen im Posttest etwas sympathischer eingeschätzt. Die Ergebnisse sind aber nicht signifikant. Gleichzeitig zeigt auch der Modelfit keine signifikanten Werte, daher kann in dieser Kategorie nicht von einer Wirkung der Intervention ausgegangen werden.

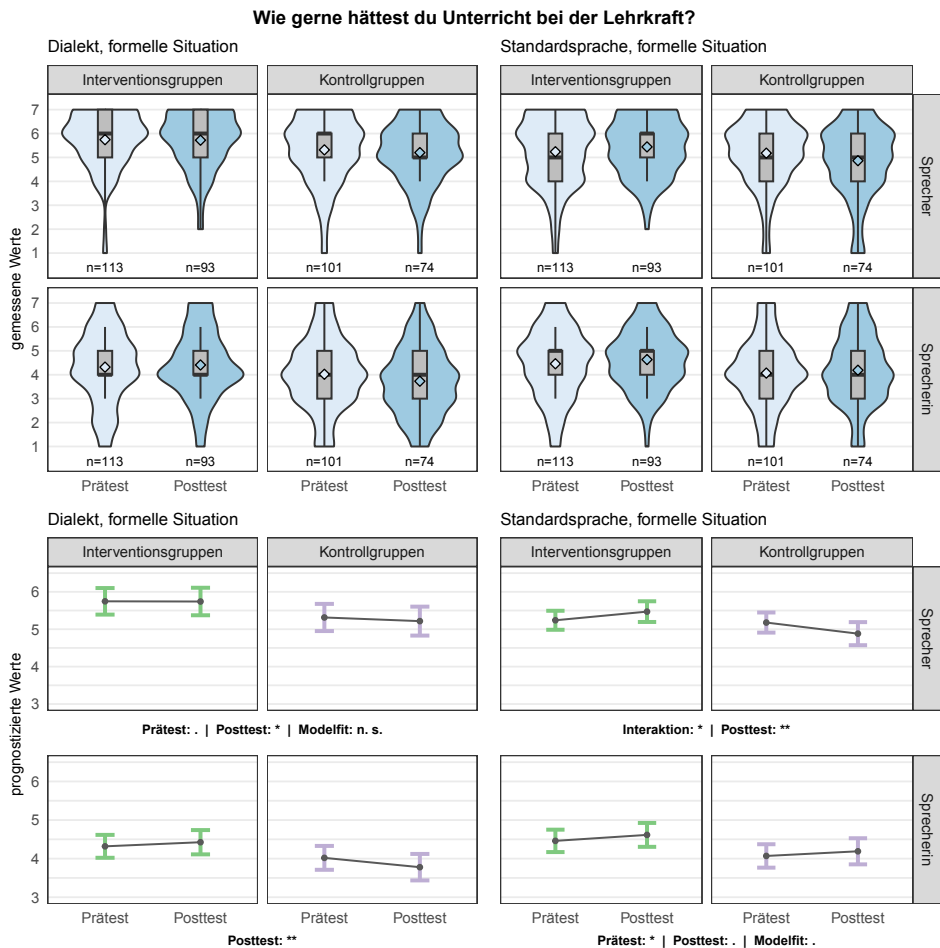


**Abbildung 11.8:** Prätest-Posttest-Vergleich bei der Frage nach der Sympathie der Schüler:innen in informeller Situation

Ähnlich verhält es sich mit den standardsprechenden Schülern. Die Werte zwischen Prä- und Posttestsituation unterscheiden sich in beiden Gruppen nur gering. Die Posttestergebnisse unterscheiden sich zwar – die Werte befinden sich in der Nähe der Signifikanzgrenze – allerdings zeigt der Modelfit wieder keine signifikanten Werte.

Auch bei der standardsprechenden Schülerin findet man analoge Muster. Die Gruppen unterscheiden sich zwar signifikant in ihren Einschätzungen im Prä- und im

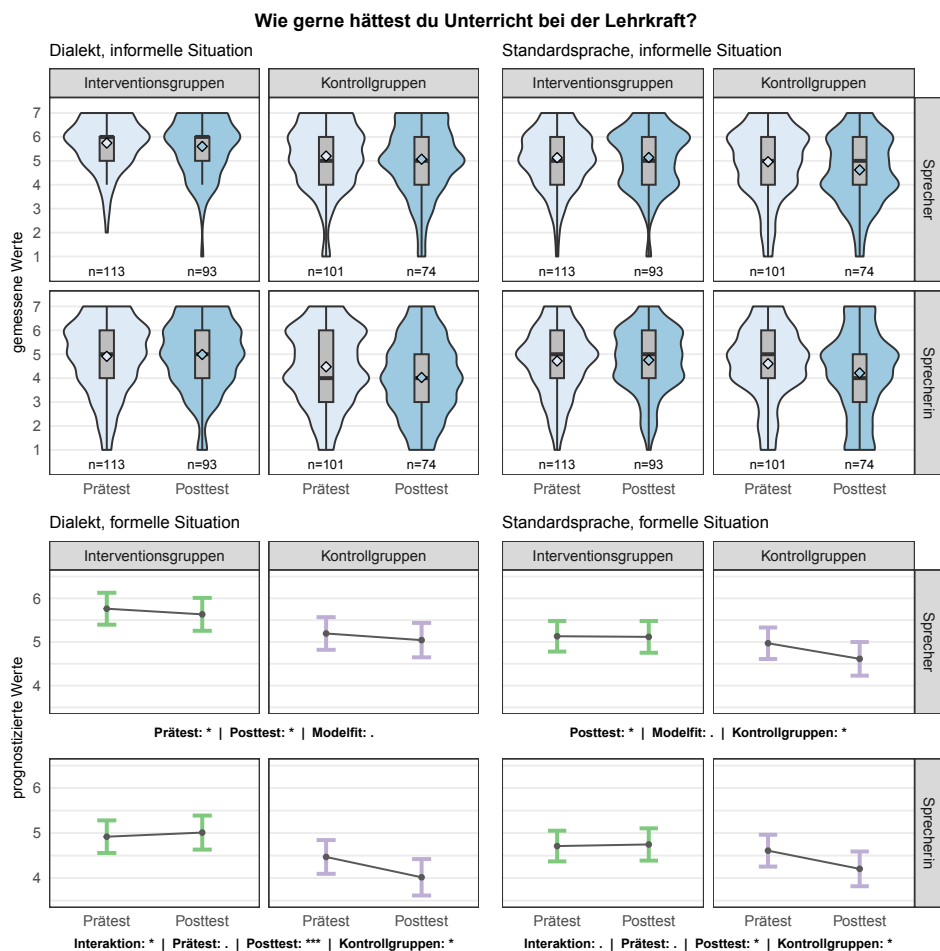
Posttest, zwischen den Testzeitpunkten sind allerdings nur geringe Veränderungen zu sehen. Dies sieht man auch am Modelfit, der sich in der Nähe der Signifikanzgrenze befindet. Alles in allem liefert die Sympathieeinschätzung auf informeller Ebene bei den Schüler:innen keine Ergebnisse, die für die Untersuchung relevant wären.



**Abbildung 11.9:** Prätest-Posttest-Vergleich bei der Frage nach allgemeiner Präferenz für Lehrkräfte in formeller Situation

Auch auf Ebene der persönlichen Präferenz gibt es Prä-Posttestveränderungen zu berichten. In formeller Situation (s. Abb. 11.9) findet man beim dialektsprechenden Lehrer nur geringe Bewegungen zwischen den Testzeitpunkten, was auch mit einem nicht signifikanten Modelfit einhergeht. Wie so oft sieht man aber, dass die Kontrollgruppen in ihren Einschätzungen etwas niedriger liegen – im Prätest sind diese in der Nähe der Signifikanzgrenze, im Posttest sind sie signifikant.

Bei der dialektsprechenden Lehrerin sieht man, dass die Interventionsgruppen im Posttest etwas positiver und die Kontrollgruppen etwas negativer im Vergleich zu den Prätestergebnissen abstimmen. Dies führt zu einem hochsignifikant unterschiedlichen Posttestergebnis, was auf eine Wirkung der Intervention hindeutet. Noch deutlicher ist das Ergebnis des standardsprechenden Lehrers. Er wird von den Interventionsgruppen im Posttest in Relation zu den Kontrollgruppen signifikant positiver eingeschätzt. Ebenso weisen die Posttestergebnisse einen hochsignifikanten Unterschied der Werte auf. Dies macht eine Wirkung der Intervention in dieser Kategorie sehr wahrscheinlich.



**Abbildung 11.10:** Prätest-Posttest-Vergleich bei der Frage nach allgemeiner Präferenz für Lehrkräfte in informeller Situation

Bei der standardsprechenden Lehrerin sieht man in beiden Gruppen eine positivere Einschätzung im Posttest – die Kontrollgruppen liegen mit ihren Prätestwerten signifikant und den Posttestwerten in der Nähe der Signifikanzgrenze unter denjenigen der

Interventionsgruppen. Der Modellfit befindet sich ebenfalls nur in der Nähe der Signifikanzgrenze – eine Wirkung der Intervention kann hier nicht angenommen werden.

In informeller Situation (s. Abb. 11.10) sieht man beim dialektsprechenden Lehrer keine großen Veränderungen – der Modellfit befindet sich daher auch nur in der Nähe der Signifikanzgrenze. Die Interventionsgruppen schätzen ihn sowohl im Prä- als auch Posttest signifikant positiver ein als die Kontrollgruppen – eine Wirkung der Intervention kann nicht angenommen werden.

Die dialektsprechende Lehrerin wird hingegen von den Interventionsgruppen in Relation zu den Kontrollgruppen im Posttest signifikant positiver eingeschätzt. Man sieht zwar, dass die Interventionsgruppen im Prätest die Lehrerin positiver einschätzen, mit Werten in der Nähe der Signifikanz. Im Posttest zeigen die Werte der Gruppen jedoch einen höchstsignifikanten Unterschied, was für eine deutliche Wirkung der Intervention spricht.

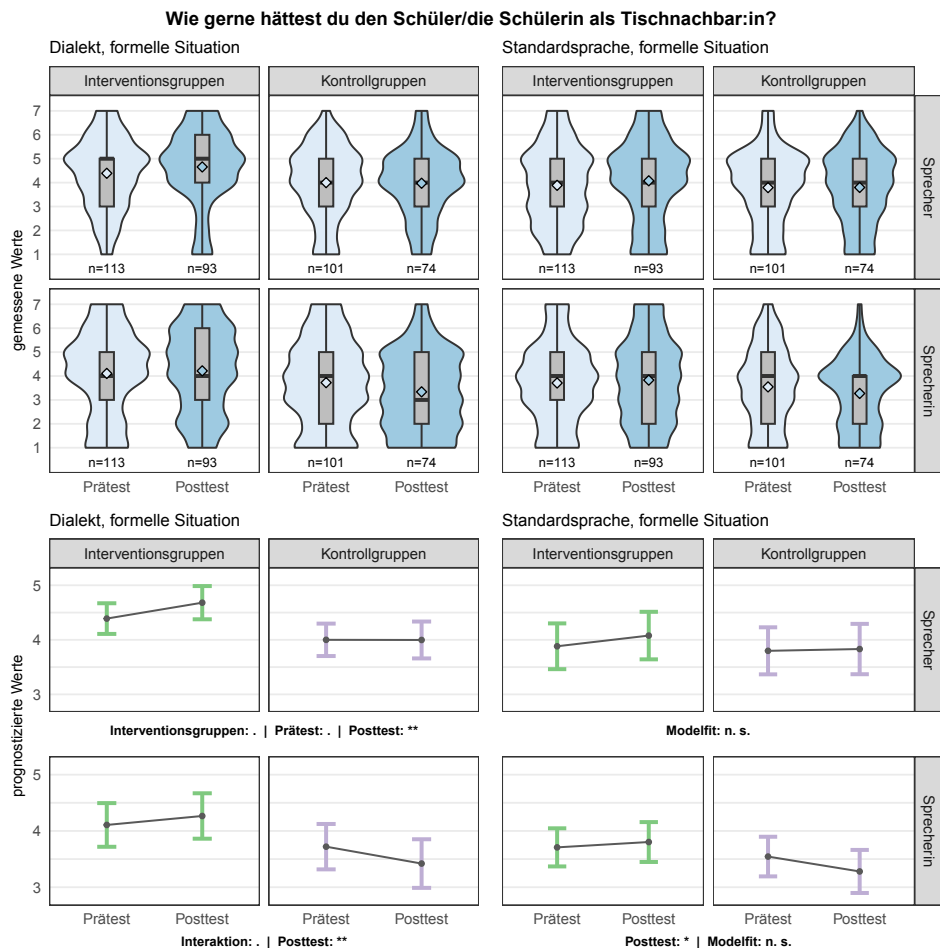
Beim standardsprechenden Lehrer ist praktisch keine Veränderung der Einschätzung in den Interventionsgruppen bemerkbar – der Modellfit befindet sich auch nur in der Nähe der Signifikanzgrenze. Die Posttestergebnisse zwischen Interventions- und Kontrollgruppen offenbaren allerdings einen signifikanten Unterschied.

Ähnlich sieht es bei der standardsprechenden Lehrerin aus. In Relation der Gruppen zueinander wird sie von den Interventionsgruppen im Posttest etwas positiver eingeschätzt – die Werte befinden sich in der Nähe der Signifikanzgrenze. Dies sieht man auch, da sich die Posttestwerte der Gruppen signifikant voneinander unterscheiden, die Prätestwerte sich allerdings nur in der Nähe der Signifikanzgrenze befinden.

Bei den Schüler:innen findet man ebenfalls Veränderungen, speziell den dialektalen Sprachgebrauch betreffend (s. Abb. 11.11). Der dialektsprechende Schüler wird von den Interventionsgruppen zwischen Prä- und Posttest positiver eingeschätzt. Diese Veränderung befindet sich in der Nähe der Signifikanzgrenze. Ebenfalls sieht man, dass die Kontrollgruppen im Prätest im Vergleich zu den Interventionsgruppen etwas negativer abgestimmt haben – die Werte befinden sich ebenfalls in der Nähe der Signifikanzgrenze. Die Posttestergebnisse beider Gruppen unterscheiden sich deutlich mit hoher Signifikanz voneinander – eine Wirkung der Intervention kann also angenommen werden.

Die dialektsprechende Schülerin wird im Posttest von den Interventionsgruppen in Relation zu den Kontrollgruppen positiver eingeschätzt, mit Werten in der Nähe der Signifikanzgrenze. Auch die Posttestergebnisse unterscheiden sich hochsignifikant voneinander – hier kann also ebenfalls von einer Wirkung der Intervention ausgegangen werden.

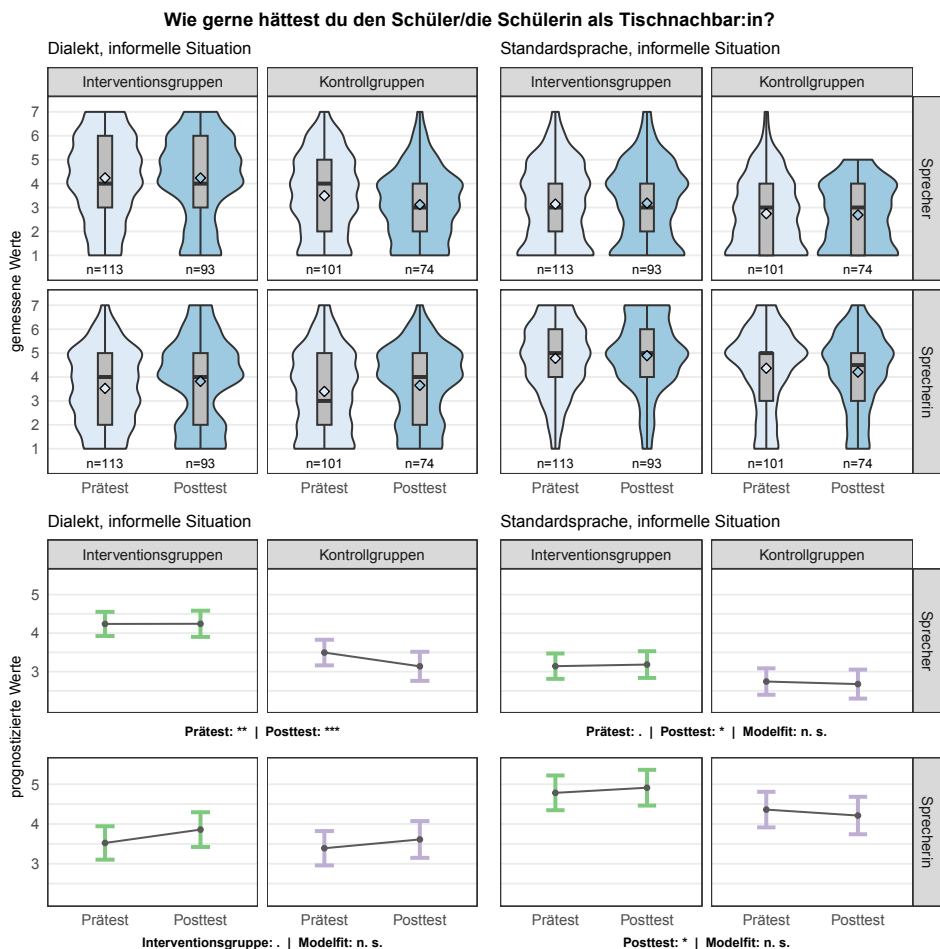
Auf standardsprachlicher Ebene zeigen sich beim Schüler keine Veränderungen. Die Schülerin wird zwar von den Interventionsgruppen etwas positiver im Posttest eingeschätzt und die Posttestergebnisse zwischen den beiden Gruppen unterscheiden sich auch signifikant voneinander, allerdings zeigt der Modellfit keine signifikanten Werte. Daher kann nicht von einer Veränderung durch die Intervention ausgegangen werden.



**Abbildung 11.11:** Prätest-Posttest-Vergleich bei der Frage nach allgemeiner Präferenz für Schüler:innen in formeller Situation

In informeller Situation findet man wiederum nur wenig Aussagekräftiges (s. Abb. 11.12). Der dialektsprechende Schüler wird im Prätest von den Interventionsgruppen hochsignifikant positiver eingeschätzt als von den Kontrollgruppen. Dieser Abstand vergrößert sich zum Posttest hin auf ein höchstsignifikantes Ergebnis.

Die dialektsprechende Schülerin wird von beiden Gruppen im Posttest positiver eingeschätzt – von den Interventionsgruppen etwas stärker, sodass sich die Werte der beiden Zeitpunkte innerhalb dieser Gruppe in der Nähe der Signifikanzgrenze befinden. Der Modelfit lässt aber keine weitere Interpretation der Ergebnisse zu. Ähnlich ist die Sachlage in standardsprachlicher Sprechweise. Sowohl beim Schüler als auch bei der Schülerin zeigt der Modelfit keine signifikanten Werte, daher wird von einer weiteren Interpretation Abstand genommen.

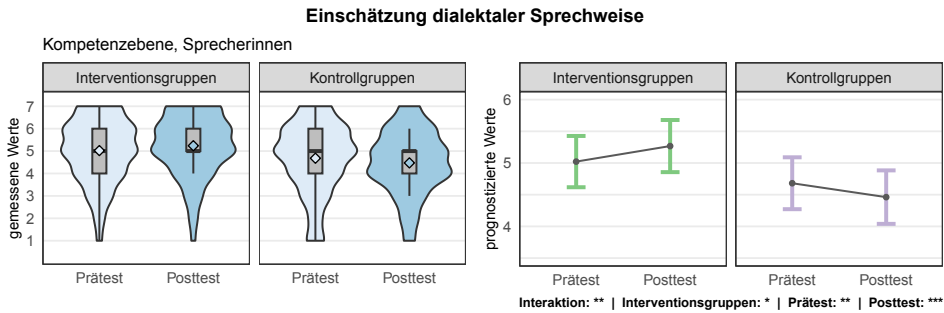


**Abbildung 11.12:** Prä-Posttest-Vergleich bei der Frage nach allgemeiner Präferenz für Schüler:innen in informeller Situation

### 11.1.2 Generelle Ergebnisse

Der Prä-Posttest-Vergleich zeigt Unterschiede auf Sympathie-, Kompetenz- sowie Ebene der persönlichen Präferenz auf. Auf Kompetenzebene wird speziell die dialekt-sprechende Lehrerin sowohl in formeller als auch in informeller Situation nach erfolgter Intervention kompetenter eingeschätzt als zuvor. Beim Lehrer findet man in formeller Situation in beiden Sprechsituationen Veränderungen in diese Richtung und bei der Schülerin in formeller Situation und dialektaler Sprechweise ebenfalls. Wegen der Ähnlichkeiten in den unterschiedlichen Konstellationen ist es einen Versuch wert,

die Ergebnisse zu verallgemeinern, wie dies schon in Kapitel 9 vorgenommen wurde (also über Rollen und Situationen).<sup>30</sup>



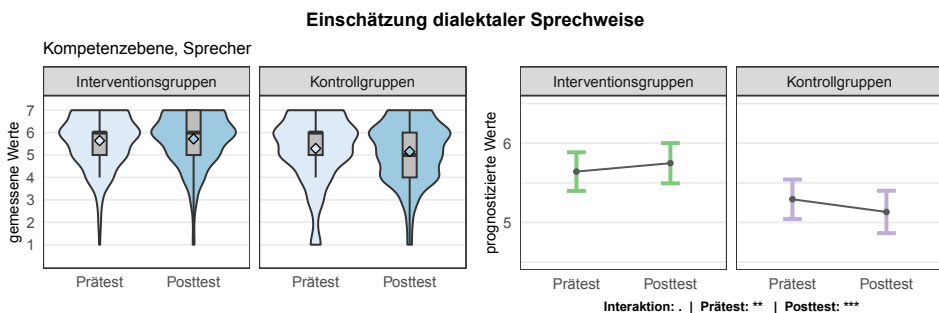
**Abbildung 11.13:** Prätest-Posttest-Vergleich der allgemeinen Einschätzung der Kompetenz von Dialekt Sprecherinnen

Sehen wir uns als Erstes das Ergebnis für die dialekt sprechenden Sprecherinnen an (Abb. 11.13). Sie werden von den Interventionsgruppen im Posttest in Relation zu den Kontrollgruppen höchstsignifikant kompetenter eingeschätzt. Auch die Werte der beiden Testzeitpunkte der Interventionsgruppen zeigen einen signifikanten Unterschied. Zwar liegen die Prätestwerte der Kontrollgruppen hochsignifikant unter denjenigen der Interventionsgruppen, dass sich dieser Unterschied im Posttest jedoch noch vergrößert und höchstsignifikante Werte aufweist, liefert einen weiteren Beleg für die Wirkung der Intervention. Sie kann hier mit sehr hoher Wahrscheinlichkeit angenommen werden.

Bei den dialekt sprechenden Sprechern fällt das Ergebnis, Abbildung 11.14, ähnlich aus. Die Interaktion befindet sich zwar nur in der Nähe der Signifikanzgrenze. Allerdings sieht man wiederum eine Zunahme des Abstands der Werte von hochsignifikant im Prätest zu höchstsignifikant im Posttest.

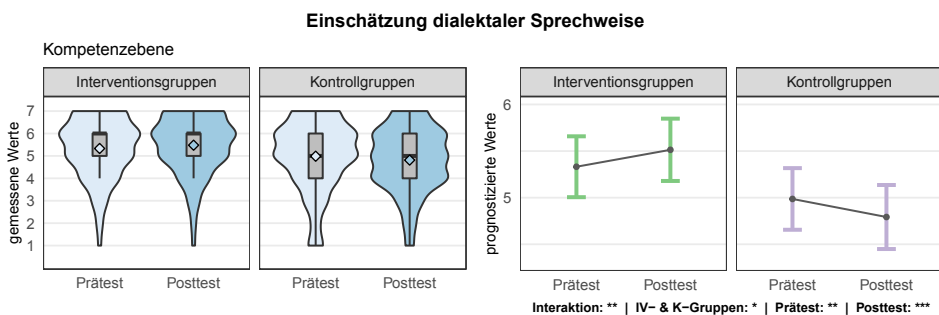
Die Ähnlichkeit dieser beiden Ergebnisse lässt sich noch weiter verallgemeinern, indem auch die Geschlechter zusammengefasst werden.

30 Die Formel für die generelle Berechnung lautet z.B.  $\text{Imer}(\text{Bewertung} \sim \text{Erhebung} * \text{Gruppen} + (\text{Erhebung} | \text{TNCODE}) + (1 | \text{Items}), \text{Dkm}, \text{REML} = \text{T})$ . Dabei handelt es sich um ein Modell mit Crossed Random Effects. Zusätzlich wurde beim Random Intercept TNCODE (das Kürzel steht für Teilnahmekode) die Erhebungssituation als Random Slope realisiert, um für eine höhere Akkuratheit (vor allem geringere Typ-1-Fehlerrate) bei den p-Werten zu sorgen (s. B. Winter 2020). Die Verschachtelung der Rater:innen in den Klassen wurde nicht mehr realisiert, da dies beim Modellfit keine Vorteile brachte. In wenigen Situationen sorgte die Random-Slope-Variable für ein Singularitätsproblem. Dort wurde sie dann weggelassen und ein Crossed-Random-Intercepts-Modell gerechnet.



**Abbildung 11.14:** Prätest-Posttest-Vergleich der allgemeinen Einschätzung der Kompetenz von Dialektsprechern

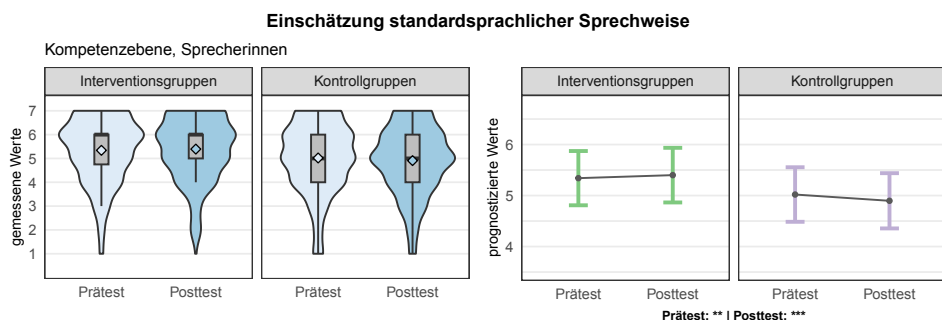
Abbildung 11.15 zeigt nun das allgemeine Modell der Einschätzung der dialektalen Sprechweise auf Kompetenzebene. Die Sprecher:innen werden im Posttest von den Interventionsgruppen in Relation zu den Kontrollgruppen hochsignifikant kompetenter eingeschätzt. Auch die Zunahme der Werte in den Interventionsgruppen zeigt ein signifikantes Ergebnis und zusätzlich sieht man wiederum eine Vergrößerung der Distanz von hochsignifikanten Unterschieden der Gruppen im Prätest zu höchstsignifikanten Unterschieden im Posttest.



**Abbildung 11.15:** Prätest-Posttest-Vergleich der allgemeinen Einschätzung der Kompetenz dialektaler Sprechweise

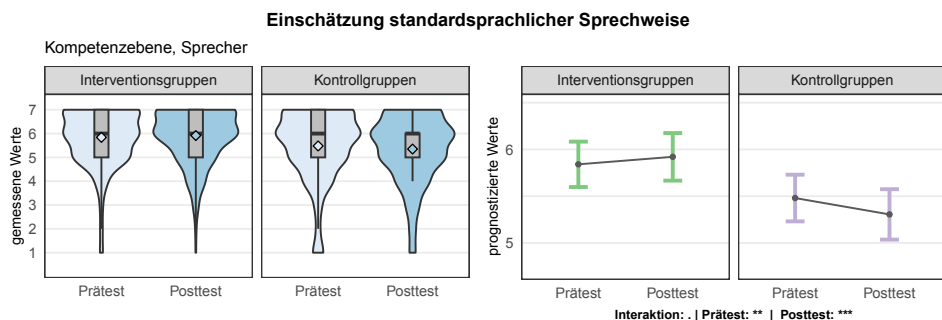
Damit kann auch  $H_{10n}$ : „Schüler:innen schätzen nach erfolgter Intervention dialekt-sprechende Personen auf Kompetenzebene gleich ein wie zuvor.“ eindeutig abgelehnt werden. Personen, die dialektal sprechen, werden also mit hoher Wahrscheinlichkeit nach der Intervention kompetenter eingeschätzt.

Auf standardsprachlicher Ebene fallen die Ergebnisse bescheidener aus. In Abbildung 11.16 kommen leicht positivere Einschätzungen in den Interventionsgruppen und leicht negativere in den Kontrollgruppen zum Vorschein. Dies bestätigt sich auch, da die Prätestergebnisse hochsignifikante und die Posttestergebnisse höchstsignifikante Unterschiede zeigen und sich damit der Abstand vergrößert. Die visuell sichtbare Effektstärke hält sich aber in Grenzen.



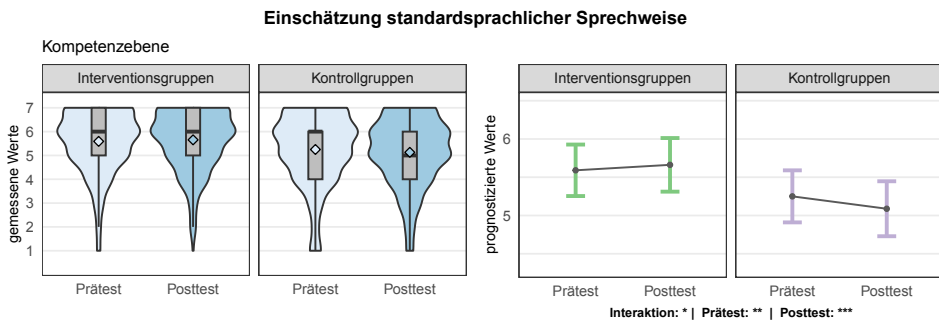
**Abbildung 11.16:** Prätest-Posttest-Vergleich der allgemeinen Einschätzung der Kompetenz von Standardsprecherinnen

Bei den Sprechern (s. Abb. 11.17) zeigt sich ein eindeutigeres Bild. Sie werden von den Interventionsgruppen in Relation zu den Kontrollgruppen im Posttest kompetenter eingeschätzt, wobei sich die Werte hier nur in der Nähe der Signifikanzgrenze befinden. Gleichzeitig findet man wiederum die Zunahme von hochsignifikant unterschiedlichen Prätestwerten zu höchstsignifikant unterschiedlichen Posttestwerten – eine moderate Wirkung der Intervention auf die Einschätzung der standardsprechenden Sprecher kann also prognostiziert werden.



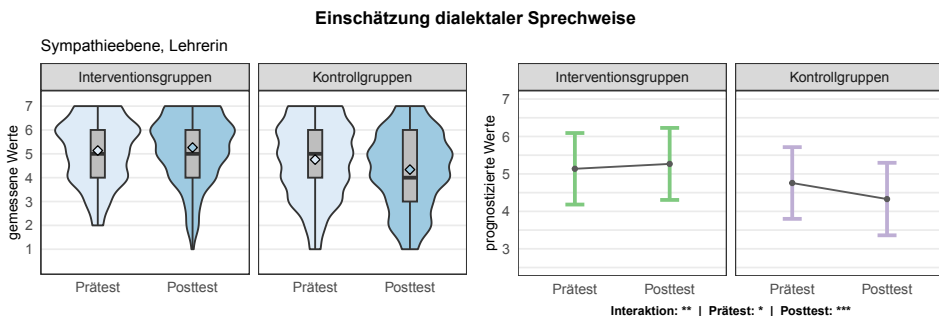
**Abbildung 11.17:** Prätest-Posttest-Vergleich der allgemeinen Einschätzung der Kompetenz von Standardsprechern

In Abbildung 11.18 sind die Werte ohne die Geschlechtertrennung abgebildet. Sie fallen sehr ähnlich aus wie diejenigen der Sprecher mit dem einzigen Unterschied, dass die Interventionsgruppen in Relation zu den Kontrollgruppen im Posttest die standardsprechenden Sprecher signifikant positiver einschätzen. Dieser Anstieg der Signifikanz ist aber wohl dem Gesetz der großen Zahlen geschuldet, da die Steigungslinien sogar flacher als in Abbildung 11.17 und trotzdem die Signifikanzwerte höher sind. Die Intervention scheint aber auch dafür zu sorgen, dass standardsprechende Sprecher:innen etwas kompetenter eingeschätzt werden, wobei das Ergebnis weniger deutlich ausfällt als das der dialektssprechenden.



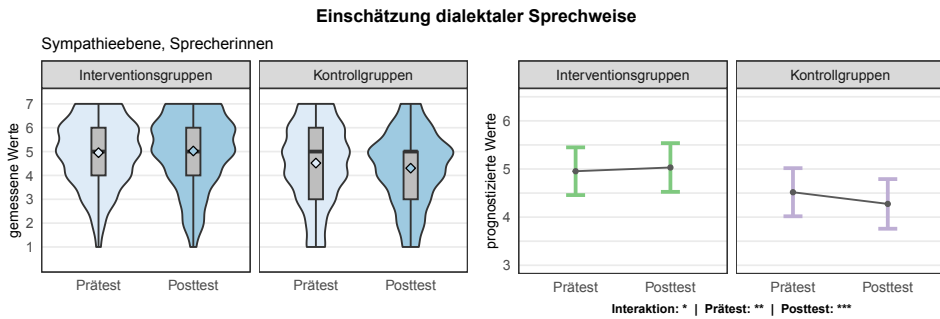
**Abbildung 11.18:** Prätest-Posttest-Vergleich der allgemeinen Einschätzung der Kompetenz von standard-sprachlicher Sprechweise

Auf Sympathieebene kommt bei den Sprecherinnen eine positivere Einschätzung speziell bei der dialekt-sprechenden Lehrerin zum Vorschein. Sie wird von den Interventionsgruppen in Relation zu den Kontrollgruppen im Posttest hochsignifikant sympathischer eingeschätzt. Der signifikante Abstand der Werte beider Gruppen im Prätest vergrößert sich auf höchstsignifikanten Abstand im Posttest (s. Abb. 11.19).



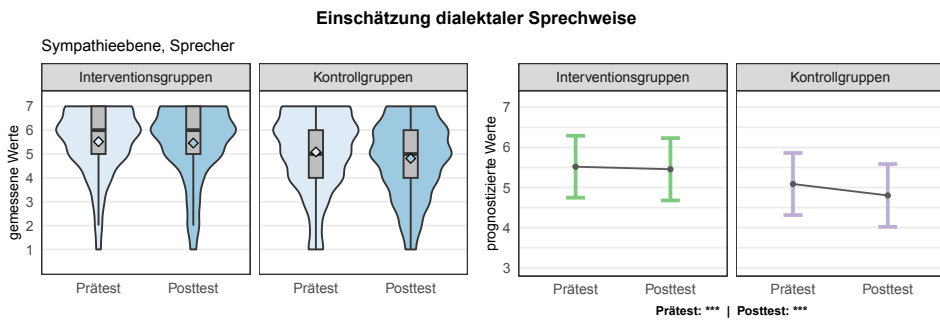
**Abbildung 11.19:** Prätest-Posttest-Vergleich der allgemeinen Einschätzung der Sympathie von dialekt-sprechenden Lehrerinnen

Wenn man die Schülerinnen-Daten in die Berechnung miteinbezieht, zeigen sich ähnliche Ergebnisse. Der Interaktionswert wird signifikant und die Prätestunterschiede hochsignifikant – ein kleiner Rückgang der Differenz (s. Abb. 11.20). Trotzdem ist ein moderater Zugewinn auf Sympathiewerte für die dialekt-sprechenden Sprecherinnen durch die Intervention wahrscheinlich.



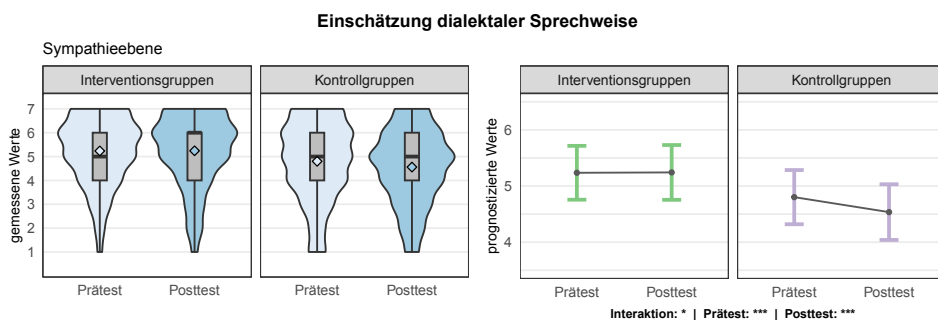
**Abbildung 11.20:** Prätest-Posttest-Vergleich der allgemeinen Einschätzung der Sympathie von Dialekt Sprecherinnen

Bei den Sprechern sieht die Sachlage anders aus (s. Abb. 11.21). Die Gruppen unterscheiden sich zwar sowohl in den Prätest- als auch den Posttestwerten höchstsignifikant voneinander, eine Wirkung der Intervention ist jedoch nicht zu beobachten.



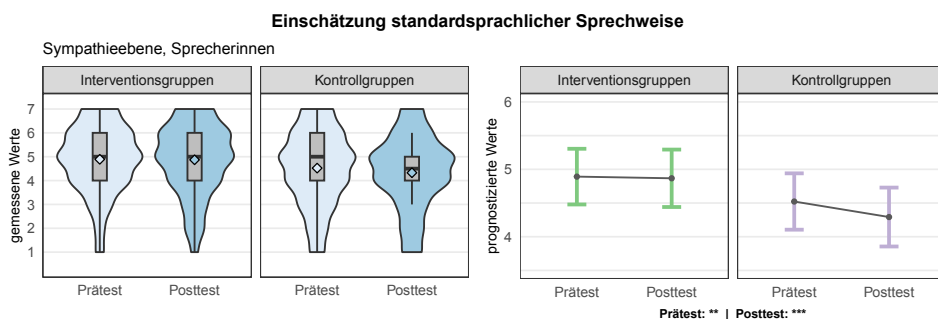
**Abbildung 11.21:** Prätest-Posttest-Vergleich der allgemeinen Einschätzung der Sympathie von Dialekt Sprechern

Das allgemeine Modell, welches beide Geschlechter beinhaltet (Abb. 11.22) zeigt, dass ein signifikanter Unterschied in der Sympathiebewertung nach erfolgter Intervention in Relation der Gruppen zueinander existiert. Ebenso sieht man eine Zunahme des Abstands der hochsignifikanten Prätestunterschiede zu höchstsignifikanten Posttestunterschieden. Dieser Plot verdeutlicht allerdings auch, dass innerhalb der Interventionsgruppen nur eine geringe Veränderung auszumachen ist. Ein Effekt der Intervention innerhalb dieser Gruppen kann also nur mit Einbezug des *Social Desirability Bias* angenommen werden. Meiner Meinung nach ist es dennoch wahrscheinlich, dass die Intervention dazu führt, dass dialektsprechende Personen moderat sympathischer eingeschätzt werden.



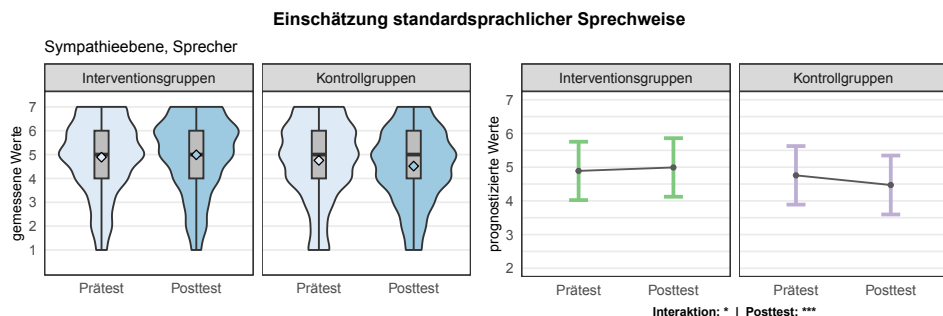
**Abbildung 11.22:** Prätest-Posttest-Vergleich der allgemeinen Einschätzung der Sympathie von dialektaler Sprechweise

Auf Sympathieebene finden sich jedoch nicht nur im dialektalen Sprachgebrauch Unterschiede. Bei den standardsprechenden Sprecherinnen (Abb. 11.23) zeigen sich zwar keine Unterschiede der Zeitpunkte bei den Interventionsgruppen, der Abstand zwischen Prä- und Posttest zwischen den Gruppen nimmt jedoch von hoch- zu höchstsignifikanten Werten zu.



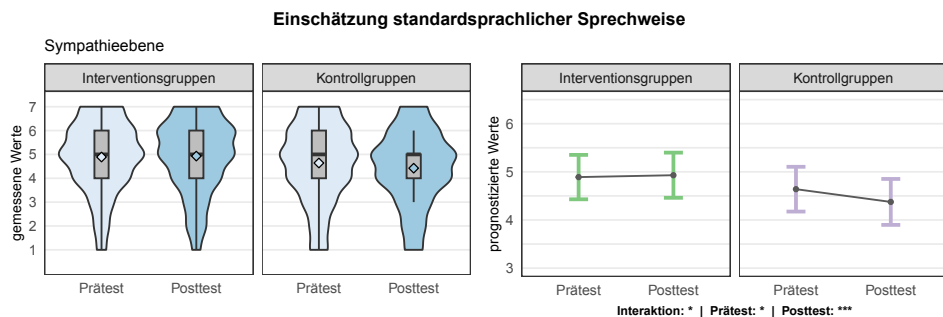
**Abbildung 11.23:** Prätest-Posttest-Vergleich der allgemeinen Einschätzung der Sympathie von Standardsprecherinnen

Etwas eindeutiger ist das Ergebnis der Sprecher (s. Abb. 11.24). Sie werden nach der Intervention in Relation zu den Kontrollgruppen signifikant sympathischer eingeschätzt, allerdings zeigt sich an den Konfidenzintervallen des Modells, dass die Streuung in diesem Fall relativ groß ist. Auch die Posttestwerte beider Gruppen zeigen höchstsignifikant unterschiedliche Werte.



**Abbildung 11.24:** Prättest-Posttest-Vergleich der allgemeinen Einschätzung der Sympathie von Standardsprechern

Wenn man nun die Geschlechter zusammen nimmt, erhält man ein ähnliches Ergebnis, wie es schon bei der allgemeinen Einschätzung der dialektalen Sprechweise der Fall war (s. Abb. 11.25). Die Ergebnisse der Interventionsgruppen unterscheiden sich nur gering voneinander und sind daher nur als Veränderung zu interpretieren, wenn man den Social Desirability Bias in die Betrachtung mit einbezieht. Dennoch kann die Nullhypothese  $H_{1In}$  (Schüler:innen schätzen nach erfolgter Intervention standardsprechende Personen auf Sympathieebene gleich ein wie zuvor.) abgelehnt werden.

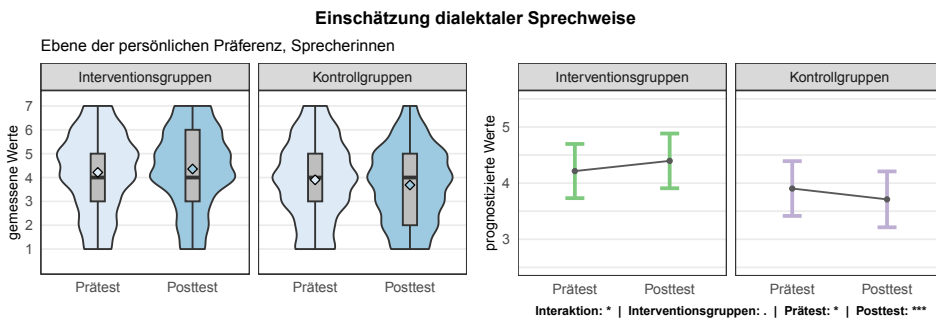


**Abbildung 11.25:** Prättest-Posttest-Vergleich der allgemeinen Einschätzung der Sympathie von standard-sprachlicher Sprechweise

Generell sind die Unterschiede auf Sympathieebene nicht so ausgeprägt, wie man es auf Kompetenzebene gesehen hat. Dies ist auch nachvollziehbar, da es nicht die primäre Intention bei der Erstellung der Unterrichtsmaterialien war, die Sympathie für eine der Varietäten zu verringern oder zu vermehren.<sup>31</sup> Auf Kompetenzebene sieht die Situation anders aus. Darauf ist gezielt in den Unterrichtsmaterialien eingegangen worden, dass Dialekt z. B. grammatische Regeln (Abschnitt 7.2) oder alte Wurzeln besitzt

<sup>31</sup> Wobei die neutral-funktionale Darstellung der Varietäten Dialekt und Standard mit ihren unterschiedlichen Domänen in den Materialien meiner Meinung nach prinzipiell schon dazu geeignet ist, die negativeren Einschätzungen des Standard-deutschen auf Sympathieebene etwas zu nivellieren.

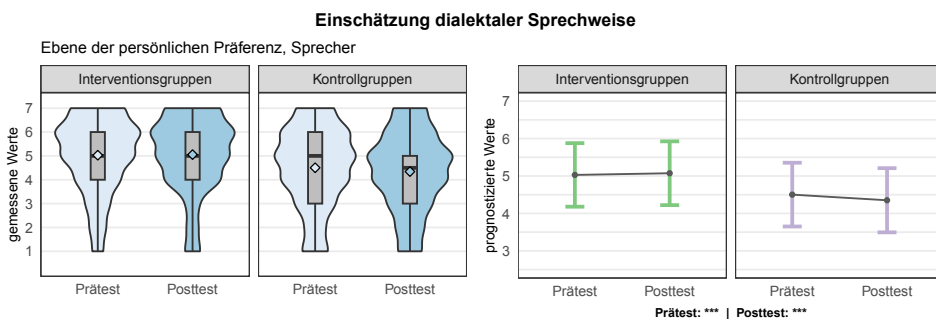
(Abschnitt 7.3). Dieses Wissen sollte stärkeren Einfluss auf die Kompetenzebene haben, da es das Stereotyp Dialekt sei schlampiges Hochdeutsch als Fehlkonzept entlarvt. In den Daten scheint sich dies auch abzubilden. Trotzdem kann man ebenfalls eine moderate Wirkung der Intervention auf standardsprechende Sprecher:innen prognostizieren, in der Gestalt, dass sie nach der Intervention sympathischer eingeschätzt werden.



**Abbildung 11.26:** Prätest-Posttest-Vergleich der allgemeinen Einschätzung der persönlichen Präferenz von Dialektsprecherinnen

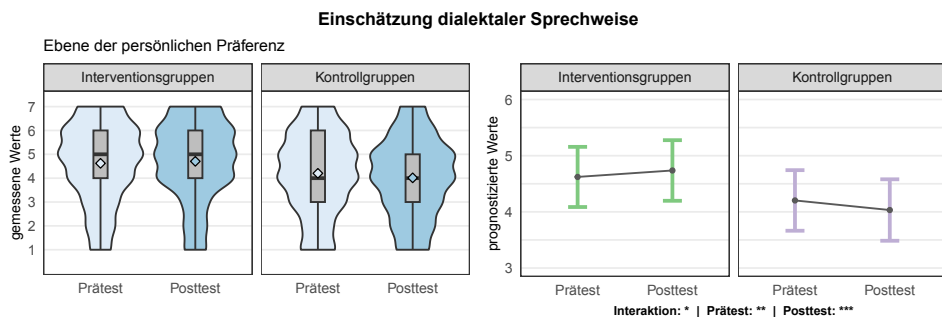
Auf Ebene der persönlichen Präferenz kommen ebenfalls sowohl in dialektsprachlicher als auch in standardsprachlicher Sprachverwendung signifikante Unterschiede zum Vorschein. Die dialektsprachlichen Sprecherinnen (s. Abb. 11.26) werden von den Interventionsgruppen in Relation zu den Kontrollgruppen nach erfolgter Intervention signifikant positiver eingeschätzt.

Auch die Testzeitpunkte der Interventionsgruppen unterscheiden sich mit Werten, die in der Nähe der Signifikanzgrenze liegen. Ebenso ist die schon öfter beobachtete Vergrößerung der Bewertungsdistanz beider Gruppen von signifikanten Unterschieden im Prätest zu höchstsignifikanten Unterschieden im Posttest zu erkennen. Alle Werte gemeinsam deuten darauf hin, dass die Intervention dafür gesorgt hat, dass dialektsprachliche Sprecherinnen positiver eingeschätzt werden.



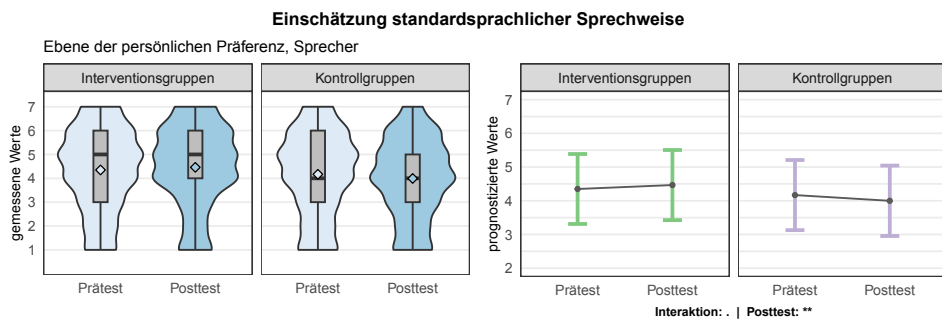
**Abbildung 11.27:** Prätest-Posttest-Vergleich der allgemeinen Einschätzung der persönlichen Präferenz von Dialektsprechern

Bei den dialektsprechenden Sprechern (s. Abb. 11.27) sieht man hingegen keine großen Veränderungen innerhalb der Gruppen. Die Prä- und Posttestwerte der Gruppen selbst unterscheiden sich höchstsignifikant voneinander, was allerdings keinen Aufschluss über die Wirkung der Intervention gibt.



**Abbildung 11.28:** Prätest-Posttest-Vergleich der allgemeinen Einschätzung der persönlichen Präferenz von dialektaler Sprechweise

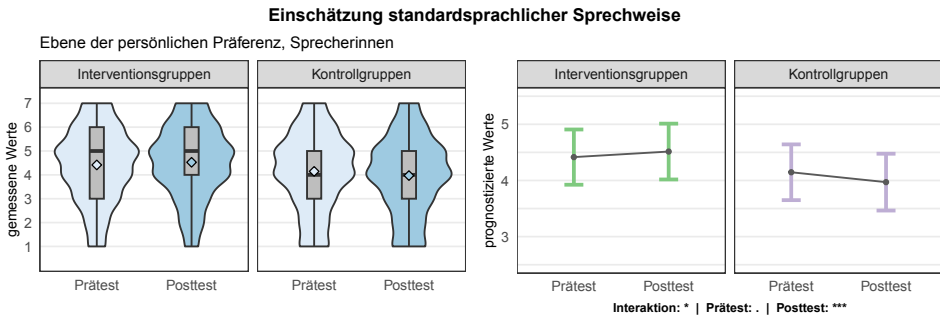
Wenn man zusätzlich auf die differenzierte Betrachtung der Geschlechter verzichtet, ähnelt das Bild demjenigen der Sprecherinnen (s. Abb. 11.28). Der Unterschied der Testzeitpunkte in Relation der Gruppen zueinander bleibt signifikant.<sup>32</sup> Damit lässt sich mit hoher Wahrscheinlichkeit annehmen, dass dialektsprechende Personen nach der Intervention allgemein positiver eingeschätzt werden.



**Abbildung 11.29:** Prätest-Posttest-Vergleich der allgemeinen Einschätzung der persönlichen Präferenz von Standardsprechern

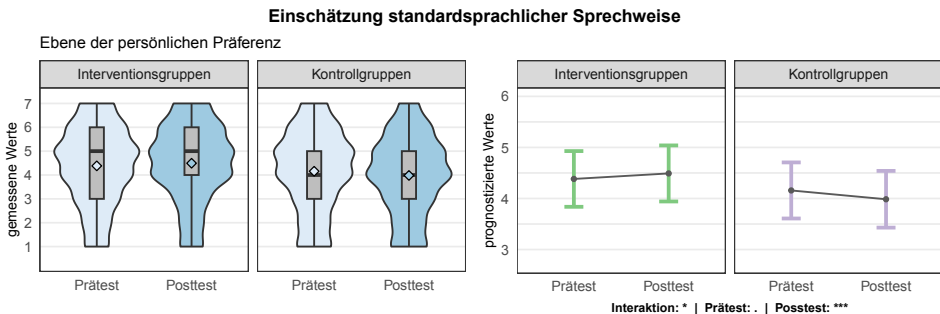
<sup>32</sup> Interessant ist der Blick zu den deskriptiven Plots. Wenn man z. B. den Posttest-Plot der Interventionsgruppen betrachtet, weisen die Werte vier, fünf und sechs eine ähnliche Häufigkeit auf. Dabei sehen die Prätestergebnisse dieser Gruppe zwar linksschief, jedoch mehr nach normalverteilten Ergebnissen aus. Die Intervention hat offensichtlich zu einer breiteren Streuung in den neutralen und positiven Einstellungsbereichen geführt (4, 5, 6) und einer Abnahme in den negativen (speziell 2 & 3). Diese Tendenzen sind auch nicht mit der gleichzeitigen Darstellung mehrerer Variablen erklärbar, da man solche Tendenzen auch in Einzelvariablen (z. B. Abbildungen 11.7 und 11.11) beobachten kann.

Auch auf standardsprachlicher Ebene findet man ähnliche verallgemeinerbare Ergebnisse, wie sie schon bei der dialektsprachlichen Ebene auftraten. Die Sprecher werden im Posttest von den Interventionsgruppen in Relation zu den Kontrollgruppen etwas positiver eingeschätzt (Abb. 11.29), wobei sich die Werte in der Nähe der Signifikanzgrenze befinden und die Konfidenzintervalle relativ breit sind. Die Posttestergebnisse beider Gruppen unterscheiden sich jedoch hochsignifikant voneinander.



**Abbildung 11.30:** Prätest-Posttest-Vergleich der allgemeinen Einschätzung der persönlichen Präferenz von Standardsprecherinnen

Bei den Sprecherinnen sieht das Ergebnis ähnlich aus – sie wurden von den Interventionsgruppen in Relation zu den Kontrollgruppen im Posttest signifikant positiver eingeschätzt (s. Abb. 11.30). Die Prätestergebnisse unterscheiden sich mit Werten nahe der Signifikanzgrenze, die Posttestergebnisse unterscheiden sich sogar höchstsignifikant voneinander.



**Abbildung 11.31:** Prätest-Posttest-Vergleich der allgemeinen Einschätzung der persönlichen Präferenz von standardsprachlicher Sprechweise

Wenn man die Geschlechter gemeinsam analysiert, findet man ein fast identisches Ergebnis vor, wie man es bei den Sprecherinnen gesehen hat – die standardsprachliche Sprechweise wird auf Ebene der persönlichen Präferenz also nach erfolgter Intervention mit großer Wahrscheinlichkeit positiver eingeschätzt (Abb. 11.31). Bei den Sprecherinnen ist die Effektstärke etwas größer.

### 11.1.3 Zusammenfassung der allgemeinen Prä-Posttest-Unterschiede

Die Betrachtung der Prä-Posttest-Unterschiede zeigt auf allen drei Ebenen, Kompetenzebene, Sympathieebene und Ebene der persönlichen Präferenz, signifikante Unterschiede. Abbildung 11.32 fasst die Ergebnisse der Kompetenzebene noch einmal zusammen. Dabei sind in den Zeilen die einzelnen Guises und in den Spalten die möglichen Analysekatégorien aufgelistet. So sind etwa in der ersten Zeile die Unterschiede zwischen den Prä- und Posttestergebnissen des dialektsprechenden Lehrers in formeller Situation gegeben. Ein Signifikanzstern in der Spalte Interaktion bedeutet, dass die Messungen einen signifikanten Unterschied der Testzeitpunkte Prä- und Posttest in Relation mit den beiden Gruppen (Interventions- und Kontrollgruppen) erbracht haben. Zwei Signifikanzsterne bei Posttest zeigen an, dass sich die Posttestergebnisse der Interventions- und der Kontrollgruppe hochsignifikant voneinander unterscheiden.

| Veränderung Kompetenzeinschätzung von Sprecher:innen, detailliert: |           |          |          |             |         |          |                |
|--|-----------|----------|----------|-------------|---------|----------|----------------|
|  |           |          |          | Interaktion | Prätest | Posttest | IV-<br>Gruppen |
| Lehrkräfte   | formell   | Dialekt  | männlich | *           |         | **       |                |
|  |           |          | weiblich | *           |         | **       | *              |
|  |           | Standard | männlich | *           |         | **       |                |
|  |           |          | weiblich |             | *       |          |                |
|  | informell | Dialekt  | männlich |             | *       | *        |                |
|  |           |          | weiblich | *           |         | ***      |                |
|  |           | Standard | männlich |             |         | *        |                |
|  |           |          | weiblich | .           |         | **       |                |
| Schüler:<br>innen  | formell   | Dialekt  | männlich |             |         | *        |                |
|  |           |          | weiblich | .           | **      | ***      |                |
|  |           | Standard | männlich |             | *       | **       |                |
|  |           |          | weiblich |             | *       | **       | *              |
|  | informell | Dialekt  | männlich |             | **      | ***      |                |
|  |           |          | weiblich |             |         | *        | *              |
|  |           | Standard | männlich |             | **      | *        |                |
|  |           |          | weiblich |             |         | *        | *              |
| Kompetenzeinschätzung, generell, geschlechterspezifisch:           |           |          |          |             |         |          |                |
| Sprecher:<br>innen   | Dialekt   | männlich | .        |             | **      | ***      |                |
|  |           | weiblich | **       |             | **      | ***      | *              |
|  | Standard  | männlich | .        |             | **      | ***      |                |
|  |           | weiblich |          |             | **      | ***      |                |
| Kompetenzeinschätzung, generell:                                   |           |          |          |             |         |          |                |
| Sprecher:<br>innen   | Dialekt   |          |          | **          | **      | ***      | *              |
|  | Standard  |          |          | *           | **      | ***      |                |

Abbildung 11.32: Übersichtstabelle Prä-Posttest-Unterschiede auf der Kompetenzebene

Man sieht, dass speziell die dialektsprechenden Lehrerinnen nach erfolgter Intervention kompetenter eingeschätzt werden. Dieses Ergebnis zeigt sich auch in der allgemeinen Betrachtung der geschlechterspezifischen Ergebnisse, die bei den dialektsprechenden

Sprecherinnen hochsignifikante Unterschiede des Relationswerts Gruppen – Testzeitpunkte zeigen. Das allgemeine Ergebnis aller dialektsprechenden Personen ist ebenfalls hochsignifikant und zeigt, dass der Materialeinsatz mit hoher Wahrscheinlichkeit zu einer höheren Kompetenzeinschätzung der dialektsprechenden Personen geführt hat. Dies spricht auch stark dafür, dass das in der Bevölkerung weitverbreitete Stereotyp, Dialektsprechende seien weniger kompetent, abgebaut wurde.

| Veränderung Sympathieeinschätzung von Sprecher:innen, detailliert: |           |          |          |    | Interaktion | Prätest | Posttest | IV-Gruppen |
|--|-----------|----------|----------|----|-------------|---------|----------|------------|
| Lehrkräfte   | formell   | Dialekt  | männlich |    |             |         | **       |            |
|  |           |          | weiblich | *  |             |         | ***      | .          |
|  |           | Standard | männlich | *  |             |         | **       |            |
|  |           |          | weiblich |    | *           |         |          |            |
|  | informell | Dialekt  | männlich |    | *           | *       | *        |            |
|  |           |          | weiblich | ** | *           | ***     |          |            |
|  |           | Standard | männlich | ** |             | **      |          |            |
|  |           |          | weiblich | .  |             | **      |          |            |
| Schüler:innen  | formell   | Dialekt  | männlich |    |             |         | *        |            |
|  |           |          | weiblich | .  | **          | **      |          |            |
|  |           | Standard | männlich |    |             |         |          |            |
|  |           |          | weiblich | *  |             | **      |          |            |
|  | informell | Dialekt  | männlich |    | .           | *       |          |            |
|  |           |          | weiblich |    |             |         |          |            |
|  |           | Standard | männlich |    |             | .       |          |            |
|  |           |          | weiblich |    | *           | *       |          |            |
| Sympathieeinschätzung, generell, geschlechterspezifisch:           |           |          |          |    |             |         |          |            |
| Sprecher:innen   | Dialekt   | männlich |          |    | ***         | ***     |          |            |
|  |           | weiblich | *        |    | **          | ***     |          |            |
|  | Standard  | männlich | *        |    |             | ***     |          |            |
|  |           | weiblich |          |    | **          | ***     |          |            |
| Sympathieeinschätzung, generell:                                   |           |          |          |    |             |         |          |            |
| Sprecher:innen   | Dialekt   |          | *        |    | ***         | ***     |          |            |
|  | Standard  |          | *        |    | *           | ***     |          |            |

**Abbildung 11.33:** Übersichtstabelle Prä-Posttest-Unterschiede auf der Sympathieebene

Ähnlich verhält es sich auf Sympathieebene (Abb. 11.33). Es zeigt sich auch hier, dass speziell bei den Lehrkräften Veränderungen hin zu einer höheren Sympathieeinschätzung erfolgt sind. In dieser Gruppe fällt allerdings auf, dass einerseits die standardsprechenden Lehrer, aber auch die dialektsprechenden Lehrerinnen sympathischer nach erfolgter Intervention eingeschätzt werden. Hier tritt folgendes Phänomen auf, das uns schon in Abschnitt 9.2 begegnet ist: Bei den Sprechern entsprechen die Stereotype der landläufigen Meinung, dass standardsprechende Personen weniger sympathisch sind. Dieses Stereotyp scheint die Intervention zu verringern.

Bei den Lehrerinnen ist es jedoch differenzierter. Sie zeigen vor der Intervention niedrige Sympathiewerte in dialektaler Sprechweise, die danach höher werden. Aber

auch bei den standardsprachlichen Guises der Lehrerin in informeller Situation und der Schülerin in formeller Situation ist ein signifikanter Aufwärtstrend in den Einstellungen bemerkbar.

Die geschlechterspezifischen Ergebnisse unterstreichen dieses Ergebnis noch: dialektsprechende Sprecherinnen und standardsprechende Sprecher werden nach der Intervention signifikant sympathischer eingeschätzt. Auch die allgemeinen Ergebnisse zur Verwendung von Dialekt und Standard zeigen einen signifikanten Zugewinn der Sympathieeinschätzung (in Relation der Gruppen und Testzeitpunkte [= Interaktion] zueinander). Die geschlechterspezifischen Ergebnisse sind in dieser Kategorie jedoch die aussagekräftigeren.

| Veränderung allgemeine Einschätzung von Sprecher:innen, detailliert: |           |          |          |             |         |          |                |
|--|-----------|----------|----------|-------------|---------|----------|----------------|
|  |           |          |          | Interaktion | Prätest | Posttest | IV-<br>Gruppen |
| Lehrkräfte   | formell   | Dialekt  | männlich |             |         | .        | *              |
|  |           |          | weiblich |             |         |          | **             |
|  |           | Standard | männlich | *           |         |          | **             |
|  |           |          | weiblich |             | *       | .        |                |
|  | informell | Dialekt  | männlich |             | *       | *        | *              |
|  |           |          | weiblich | *           | .       | ***      |                |
|  |           | Standard | männlich |             |         | *        |                |
|  |           |          | weiblich | .           | .       | *        |                |
| Schüler:<br>innen  | formell   | Dialekt  | männlich |             |         | .        | **             |
|  |           |          | weiblich | .           |         | **       |                |
|  |           | Standard | männlich |             |         |          | *              |
|  |           |          | weiblich |             |         | *        |                |
|  | informell | Dialekt  | männlich |             | **      | ***      |                |
|  |           |          | weiblich |             |         |          | .              |
|  |           | Standard | männlich |             | .       | *        |                |
|  |           |          | weiblich |             |         | *        |                |
| Allgemeine Einschätzung, generell, geschlechterspezifisch:           |           |          |          |             |         |          |                |
| Sprecher:<br>innen   | Dialekt   | männlich |          |             | ***     | ***      |                |
|  |           | weiblich | *        | *           | ***     | .        |                |
|  | Standard  | männlich | .        |             | **      |          |                |
|  |           | weiblich | *        | .           | ***     |          |                |
| Allgemeine Einschätzung, generell:                                   |           |          |          |             |         |          |                |
| Sprecher:<br>innen   | Dialekt   |          | *        | **          | ***     |          |                |
|  | Standard  |          | *        | .           | ***     |          |                |

**Abbildung 11.34:** Übersichtstabelle Prä-Posttest-Unterschiede auf der Ebene der persönlichen Präferenz

Die Ebene der persönlichen Präferenz (Abb. 11.34) bestätigt die bisher gezeigten Ergebnisse, wobei signifikante Effekte hier seltener vorkommen – bei den Lehrkräften etwas häufiger als bei den Schüler:innen. Es stechen die dialektsprechenden Sprecherinnen heraus. In formeller Situation sieht man bei der Lehrerin und bei der Schülerin Tendenzen zu einer positiveren Einschätzung von Dialektspacherinnen im Vergleich zu den Kontrollgruppen; in informeller Situation ist diese Tendenz bei der Lehrerin deut-

lich ausgeprägt. Auch der standardsprechende Lehrer in formeller Situation wird nach erfolgter Intervention positiver eingeschätzt.

Bei den geschlechterspezifischen Ergebnissen sieht man, dass sowohl die dialekt-sprechenden als auch die standardsprechenden Sprecherinnen deutlich positiver von den Interventionsgruppen im Posttest und auch in Relation mit den Kontrollgruppen eingeschätzt werden. Auch der standardsprechende Sprecher zeigt diese Veränderung, wenn auch schwächer ausgeprägt. Die geschlechterspezifischen Ergebnisse sind in dieser Kategorie die aussagekräftigeren, obwohl auch eine allgemeine Betterbewertung beider Varietäten im Posttest der Interventionsgruppen messbar ist. So wie auf Sympathieebene sind diese Ergebnisse jedoch nur in Bezug auf die Geschlechter und die Unterschiede in den Prätestergebnissen sinnvoll zu interpretieren.

## 11.2 Dialekt und Standarddeutsch

### 11.2.1 Detaillierte Ergebnisse

In Abschnitt 11.1 wurden Einstellungsveränderungen anhand der mittleren Zu- oder Abnahme der Einschätzungen der Rater:innen vorgenommen. Es existieren jedoch auch andere Möglichkeiten, dies zu untersuchen. In diesem Kapitel fokussieren wir auf die Bewertungsunterschiede zwischen den Varietäten – die Varietätendistanz. Dafür werden pro Proband von den standardsprachlichen Guises die dazugehörigen dialektsprachlichen Guises subtrahiert. Dies funktioniert in R ganz einfach, indem man einen Variablenvektor von einem anderen subtrahiert und den neu entstandenen Vektor unter neuem Namen speichert. Wichtig dabei ist nur, dass es sich um die gleiche Personenanzahl und gleiche Reihung in beiden Variablen handelt.<sup>33</sup> Dies ist bei einem Tidy Dataset aber sowieso Standard.

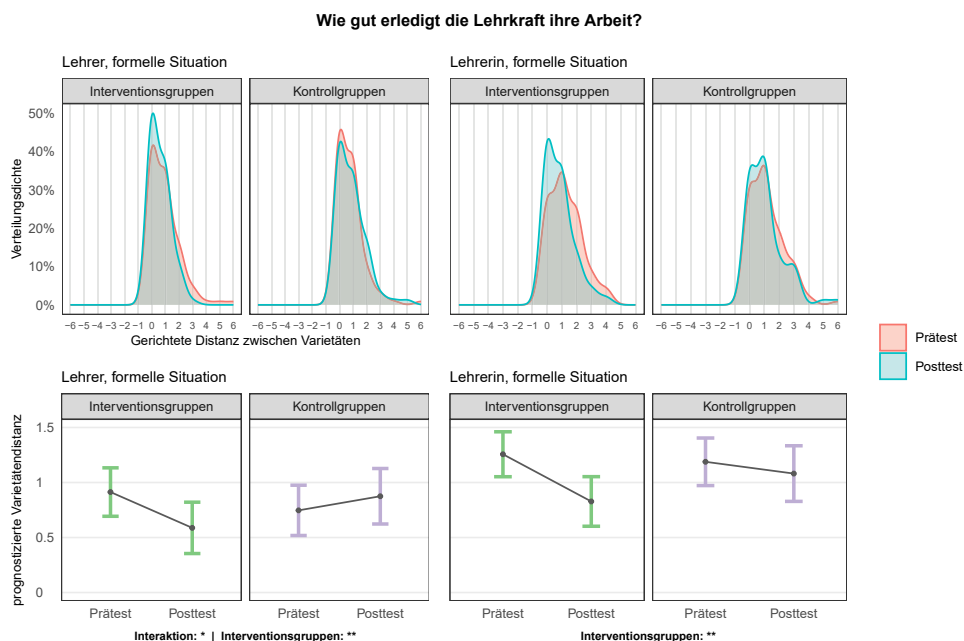
Die neue Variable zeigt damit den Abstand der Varietäten. Da es sich um eine sieben-teilige Likertskala handelt, reicht die neue Skala von  $-6$  bis  $+6$ . Wenn also ein Rater/eine Raterin die standardsprechende Lehrerin in der formellen Situation mit der maximal möglichen Sieben und die dialekt-sprechende Lehrerin mit einer minimalen Eins einschätzt, dann wird von der Sieben die Eins subtrahiert und man erhält 6. Wenn im Gegensatz bei beiden Guises die gleiche Zahl verwendet wurde, landet man bei 0. Solche „gerichteten“ Distanzen lassen sich mit Density Plots gut abbilden (wie es z. B. bei Abb. 11.35 in der oberen Hälfte realisiert ist).

Bei diesen Diagrammen ist es durch unterschiedliche Farbwahl und eine durchscheinende Deckkraft der Farben möglich, die Prä- und Posttestergebnisse übereinanderzulegen und somit intuitiv vergleichen zu können. Bei den inferenzstatistischen Modellen, die die Signifikanz der Veränderungen testen, kann man auf zwei unterschiedliche Weisen vorgehen. Wie man schon bei Abbildung 11.35 sieht, ist die Bewertungsrichtung zumeist recht einseitig. In diesem Beispiel gibt es einige Probanden, die mit einem Einser-, Zweier- oder sogar Dreierabstand den Standard besser einschätzen

---

33 Für die genauen Berechnungen s. das Dokument `sekiidistvar.R` auf meiner OSF-Seite <https://osf.io/7mhwf/>.

als den Dialekt. Im Gegensatz dazu findet sich fast keine Person, die den Dialekt besser einschätzt als den Standard, da bei  $-1$  die Linie schon fast bei 0 % auf der y-Achse angekommen ist. Man könnte die Modelle also entweder einseitig oder zweiseitig rechnen – also einmal nur die Seite, die Bewertungsunterschiede zeigt, oder beide Seiten ungeachtet der Bewertungsrichtung. In der Praxis zeigt sich, dass die Modellierung auf die unterschiedlichen Arten keinen großen Einfluss auf die Ergebnisse hat. Die geringere Zahl der Ratings einer einzelnen Richtung (Gesetz der großen Zahlen) scheint sich mit der größeren Effektstärke auszugleichen.<sup>34</sup> Bei den Effektplots, die den Dichtedialogrammen untergeordnet sind, handelt es sich um ungerichtete Modelle, die alle Daten beinhalten. Zwecks Übersichtlichkeit werden nur Diagramme gezeigt, die signifikante Unterschiede enthalten.



**Abbildung 11.35:** Prätest-Posttest-Vergleich der Distanz der Dialekt-Standard-Einschätzung bei der Frage nach der Kompetenz der Lehrkräfte in formeller Situation

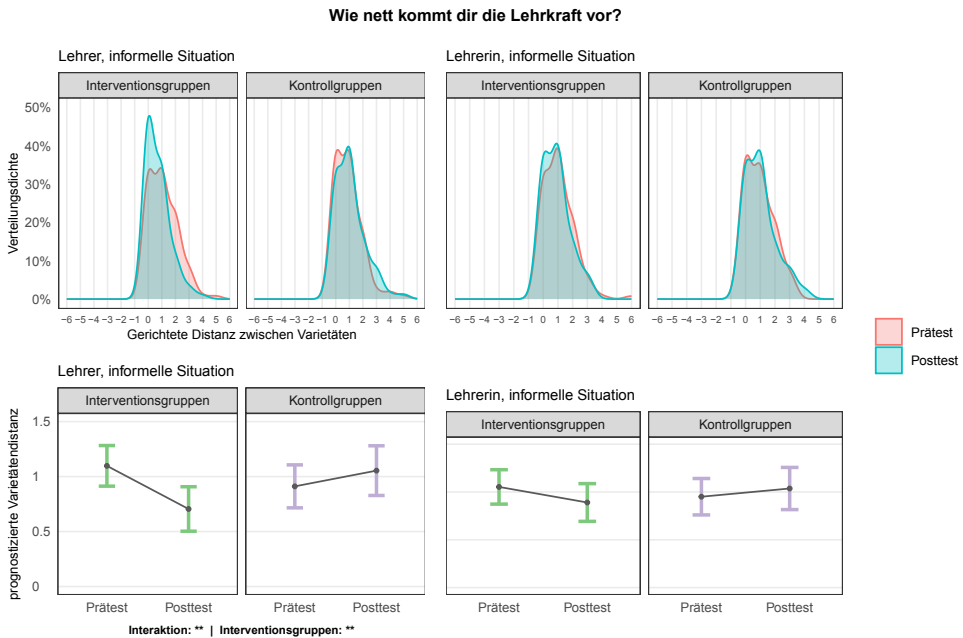
Wie schon in den vorigen Abschnitten starten wir die Analyse bei den Lehrkräften in formeller Situation auf Kompetenzebene (s. Abb. 11.35). Der Lehrer wird von den Interventionsgruppen im Vergleich zu den Kontrollgruppen zwischen den Testzeitpunkten mit hochsignifikant geringerer Varietätendistanz eingeschätzt.

Auch das Ergebnis innerhalb der Interventionsgruppen zeigt zwischen den Testzeitpunkten einen hochsignifikanten Rückgang. Man kann diese Tendenz überdies

34 Ich habe alle relevanten Modelle auf beide Weisen berechnet.

auch recht deutlich am Dichtediagramm erkennen, da die 0er-Werte im Posttest deutlich zugenommen und die Werte der Distanz zwei bis sechs abgenommen haben.

Bei der Lehrerin ist ebenfalls der Unterschied zwischen den Testzeitpunkten innerhalb der Interventionsgruppen hochsignifikant, was auf eine deutliche Einstellungsveränderung der Proband:innen in Bezug auf die Kompetenzeinschätzung von Dialekt-sprechenden hindeutet. Für beide Geschlechter kann also angenommen werden, dass die Intervention mit hoher Wahrscheinlichkeit dafür gesorgt hat, dass Varietäten in der Kompetenzeinschätzung ähnlicher wahrgenommen werden.



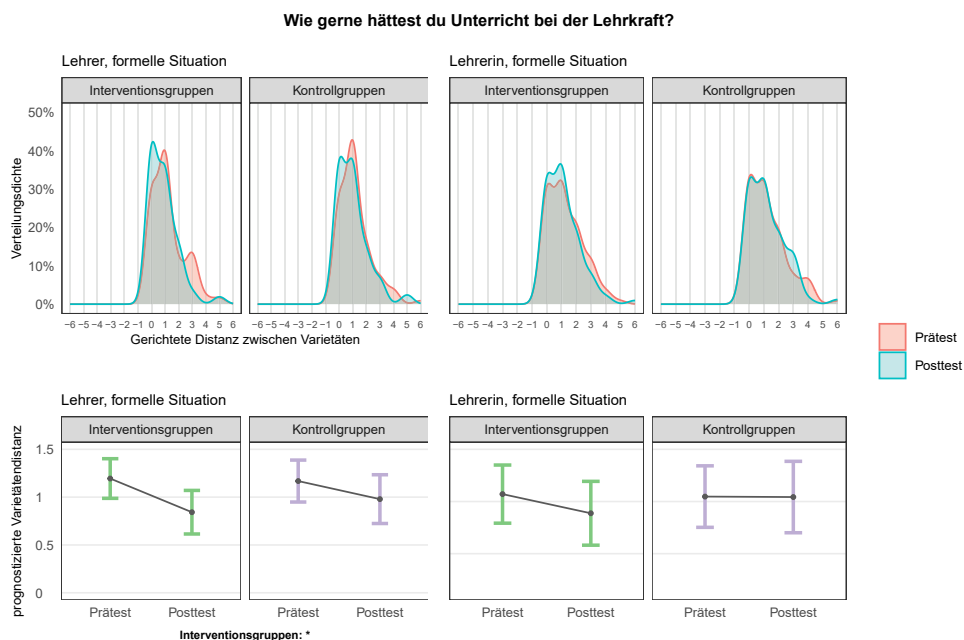
**Abbildung 11.36:** Prätest-Posttest-Vergleich der Distanz der Dialekt-Standard-Einschätzung bei der Frage nach der Sympathie der Lehrkräfte in informeller Situation

Die nächsten signifikanten Ergebnisse findet man auf Sympathieebene in informeller Situation beim Lehrer (s. Abb. 11.36). Sowohl der Bewertungsunterschied in Relation zur Kontrollgruppe als auch der Bewertungsunterschied zwischen den beiden Testergebnissen der Interventionsgruppen zeigen, dass der Lehrer hochsignifikant sympathischer nach erfolgter Intervention eingeschätzt wird, was wiederum auf eine Wirkung der Intervention hindeutet.

Auch bei der Lehrerin gehen die Tendenzen in die gleiche Richtung, sie sind jedoch nicht so ausgeprägt und zeigen auch keine signifikanten Unterschiede.

Ein ähnliches Ergebnis ist auch auf Ebene der persönlichen Präferenz in formeller Situation festzustellen (s. Abb. 11.37). Die Bewertungsdifferenz der Varietäten Dialekt und Standarddeutsch geht innerhalb der Interventionsgruppen zwischen den Test-

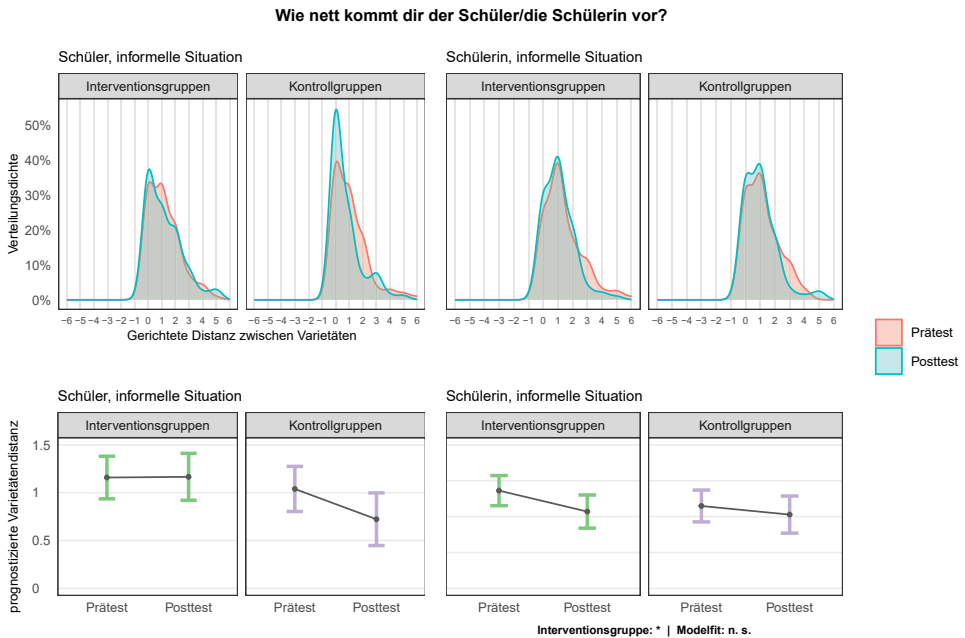
zeitpunkten signifikant zurück. Es ist zwar auch ein Rückgang in den Kontrollgruppen erkennbar, dieser ist jedoch nicht so ausgeprägt und auch nicht signifikant.



**Abbildung 11.37:** Prätest-Posttest-Vergleich der Distanz der Dialekt-Standard-Einschätzung bei der Frage nach persönlicher Präferenz für Lehrkräfte in formeller Situation

Auch bei der Lehrerin sieht man wiederum die gleiche Tendenz des Rückgangs der Bewertungsdifferenz der Varietäten, allerdings weniger ausgeprägt und nicht signifikant. Bei den Schüler:innen fallen die Unterschiede weniger stark aus als bei den Lehrkräften. Es existiert über alle Situationen und Ebenen nur ein signifikantes Ergebnis, und zwar auf Sympathieebene in informeller Situation bei der Schülerin (s. Abb. 11.38). Hier findet man einen signifikanten Rückgang der Bewertungsdifferenz der Varietäten innerhalb der Interventionsgruppen. Allerdings unterscheiden sich die Posttestergebnisse der Interventions- und Kontrollgruppen nicht voneinander und auch der Modellfit liefert keine signifikanten Werte. Daher muss man mit den Schlussfolgerungen vorsichtig sein. Von einer Wirkung der Intervention kann bei dieser Datenlage jedenfalls nicht ausgegangen werden. Auffallend ist in diesem Diagramm der Distanzrückgang in den Kontrollgruppen beim Schüler. Vergleicht man dieses Ergebnis mit den regulären Einschätzungen des dialekt- und standardsprechenden Schülers in informeller Situation (Abb. 11.8) fällt Folgendes auf: Die Kontrollgruppen schätzen im Prätest den dialekt-sprechenden Schüler deutlich sympathischer ein als den standardsprechenden, im Posttest findet man in beiden Sprachsituationen deutlich mehr 3er- und 4er-Werte.

Da der 4er-Wert auf der Siebenerskala die Mitte darstellt, ist es wahrscheinlich, dass der Distanzrückgang daher stammt, dass aufgrund geringen Interesses am Test in der Posttestsituation vermehrt die Mitte angegeben wurde.

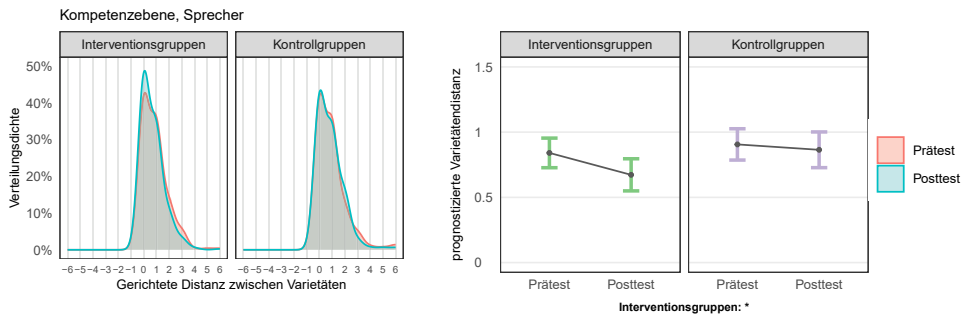


**Abbildung 11.38:** Prätest-Posttest-Vergleich der Distanz der Dialekt-Standard-Einschätzung bei der Frage nach der Sympathie der Schüler:innen in informeller Situation

### 11.2.2 Generelle Ergebnisse

Die generellen Ergebnisse können speziell in diesem Kapitel, in dem nur in wenigen Kategorien Unterschiede gefunden werden, aufschlussreiche weitere Informationen bereitstellen. In Abbildung 11.39 ist die Einschätzung der Sprecher auf Kompetenzebene abgebildet. Die deskriptiven Plots zeigen, dass es bei den Interventionsgruppen nach erfolgter Intervention die Tendenzen hin zu einer neutraleren Bewertung gibt, da deutlich mehr Nulldifferenz-Bewertungen abgegeben wurden. Auch das inferenzstatistische Ergebnis bestätigt dies, indem es hier zu einer signifikanten Distanzreduktion in der Varietätenbewertung gekommen ist.

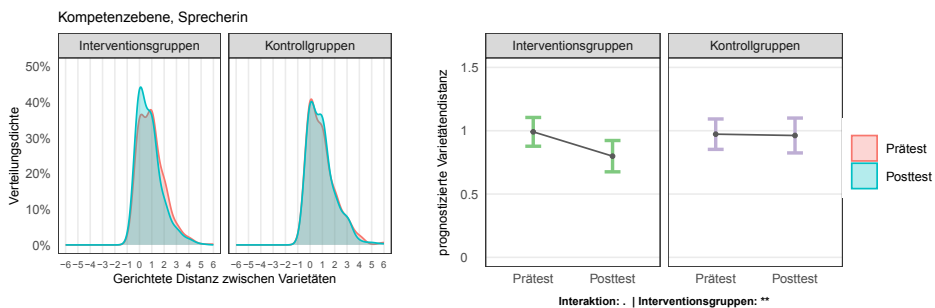
## Veränderung der Varietätendistanz



**Abbildung 11.39:** Prätest-Posttest-Vergleich der Distanz der allgemeinen Dialekt-Standard-Einschätzung bei der Frage nach der Kompetenz bei Sprechern

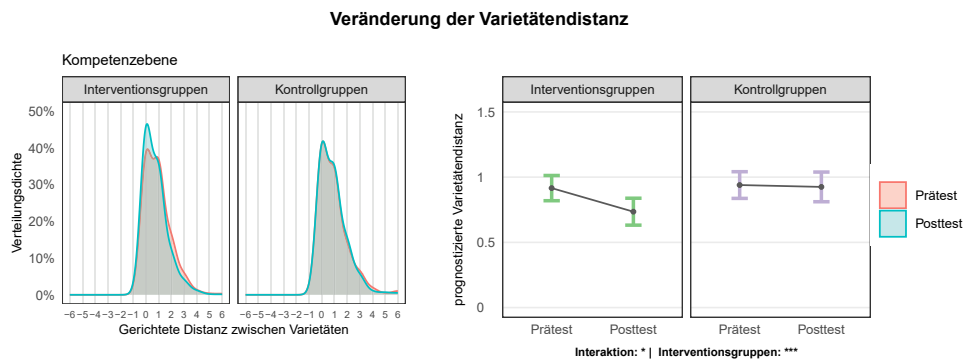
Bei den Sprecherinnen (s. Abb. 11.40) ist dieses Ergebnis noch ausgeprägter. Die Distanzreduktion innerhalb der Interventionsgruppen zeigt hochsignifikante Werte und auch die Messergebnisse zum Distanzrückgang in Relation zur Kontrollgruppe befinden sich nahe der Signifikanzgrenze.

## Veränderung der Varietätendistanz



**Abbildung 11.40:** Prätest-Posttest-Vergleich der Distanz der allgemeinen Dialekt-Standard-Einschätzung bei der Frage nach der Kompetenz bei Sprecherinnen

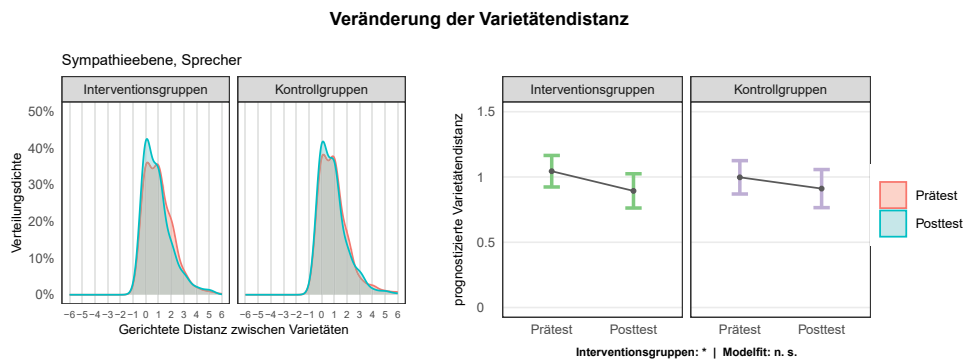
Wenn man beide Ergebnisse gemeinsam betrachtet (Abb. 11.41), dann steigt die Signifikanz in beiden relevanten Bereichen. In den Interventionsgruppen zeigt der Rückgang der Bewertungsdistanz der Varietäten nun höchstsignifikante Werte, aber auch in Interaktion mit der Kontrollgruppe wird nun die Signifikanzschwelle übersprungen. Die Nullhypothese  $H_{10n}$ : „Schüler:innen schätzen nach erfolgter Intervention dialektsprechende Personen auf Kompetenzebene gleich ein wie zuvor“ kann somit auch bei Betrachtung der Varietätendistanz verworfen werden.



**Abbildung 11.41:** Prättest-Posttest-Vergleich der Distanz der allgemeinen Dialekt-Standard-Einschätzung bei der Frage nach der Kompetenz

Es ist damit sehr wahrscheinlich, dass Rater:innen nach der Intervention Personen, die Dialekt oder Standarddeutsch sprechen, auf Kompetenzebene ähnlicher einschätzen.

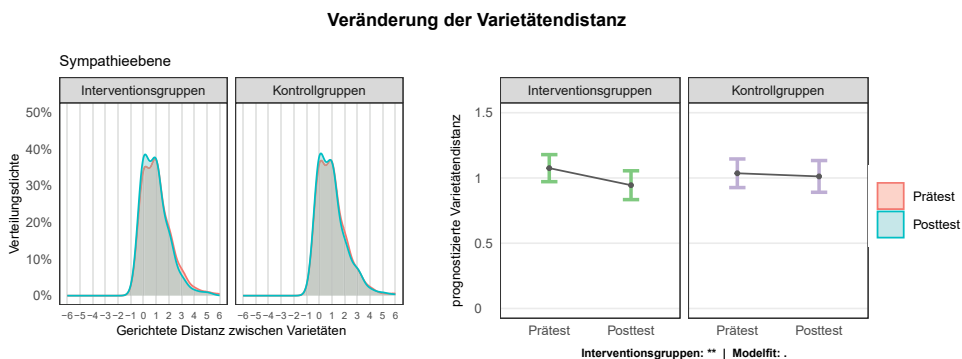
Das Ergebnis ist auch aus statistischer Sicht ideal für den Versuchsaufbau. Die Mittelwerte und Konfidenzintervalle der Kontrollgruppe und des Prättestergebnisses der Interventionsgruppen sind nämlich praktisch identisch. Dass nur das Posttestergebnis der Interventionsgruppen ausschert, spricht daher sehr dafür, dass dies einen Effekt der Intervention darstellt.



**Abbildung 11.42:** Prättest-Posttest-Vergleich der Distanz der allgemeinen Dialekt-Standard-Einschätzung bei der Frage nach der Sympathie bei Sprechern

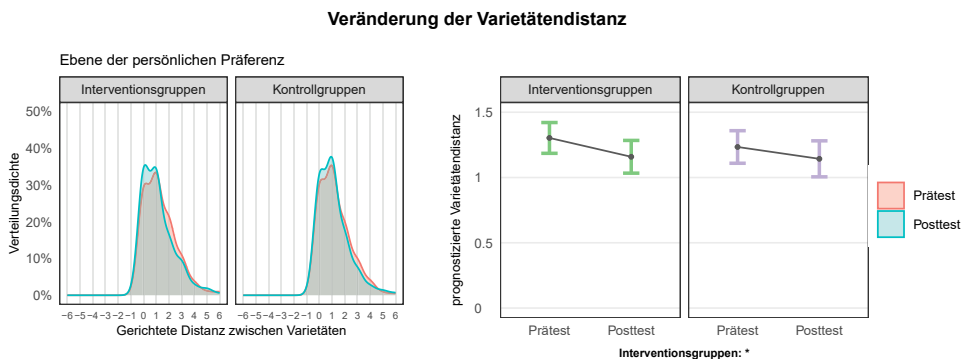
Auf Sympathieebene findet man bei geschlechterspezifischer Betrachtung nur bei den Sprechern ein signifikantes Ergebnis (s. Abb. 11.42). Die Bewertungsdistanz der Varietäten innerhalb der Interventionsgruppen geht signifikant zurück. Zwar geht auch diejenige der Kontrollgruppen zurück – dieser Wert ist jedoch nicht signifikant. Problematisch bei diesem Ergebnis ist, dass die Posttestergebnisse zwischen Interventions- und Kontrollgruppe nicht voneinander abweichen und auch, dass der Modellfit keine signifikanten Ergebnisse liefert – man kann eine Wirkung der Intervention also nicht prognostizieren.

Wenn man nun beide Geschlechter gemeinsam analysiert (s. Abb. 11.43), sieht das Ergebnis etwas besser aus. Auch wenn die Messwerte der Sprecherinnen keine signifikanten Unterschiede beinhalten, zeigen sie die erhofften Tendenzen, die bei Vereinigung mit den Sprecherdaten hochsignifikante Veränderungen innerhalb der Interventionsgruppen zeigen.



**Abbildung 11.43:** Prätest-Posttest-Vergleich der Distanz der allgemeinen Dialekt-Standard-Einschätzung bei der Frage nach der Sympathie

Die Kontrollgruppen bleiben dagegen über die Testzeitpunkte hinweg bei ihren Einschätzungen. Der Modellfit befindet sich jedoch nur in der Nähe der Signifikanzgrenze. Damit kann auf Sympathieebene die Nullhypothese  $H_{1In}$  (Schüler:innen schätzen nach erfolgter Intervention standardsprechende Personen auf Sympathieebene gleich ein wie zuvor.) vorsichtig abgelehnt werden. Der Effekt ist aber deutlich schwächer ausgeprägt als auf Kompetenzebene.



**Abbildung 11.44:** Prätest-Posttest-Vergleich der Distanz der allgemeinen Dialekt-Standard-Einschätzung bei der Frage nach persönlicher Präferenz

Auf Ebene der persönlichen Präferenz findet man weder bei den Sprechern noch bei den Sprecherinnen signifikante Unterschiede. Nur wenn man die Ergebnisse zusammen analysiert (s. Abb. 11.44), zeigt sich ein signifikanter Rückgang der Bewertungsdis-

tanz der Varietäten Dialekt und Standarddeutsch in den Interventionsgruppen. Der Rückgang ist auch ausgeprägter als bei den Kontrollgruppen, allerdings sind die Post-testergebnisse der Gruppen praktisch identisch, von einer Veränderung durch die Intervention kann also auf dieser Ebene nicht per se ausgegangen werden.

### 11.2.3 Zusammenfassung Dialekt und Standarddeutsch

Die Ergebnisse der drei Einstellungsebenen zeigen auch bei der Betrachtung des Rückgangs der Bewertungsdistanz ähnliche Ergebnisse, wie sie schon in Abschnitt 11.1 aufgetreten sind. Zum einen treten Veränderungen vermehrt bei den Lehrkräfte-Guises auf. Bei den Schüler:innen zeigt sich nur bei der Schülerin auf Sympathieebene in informeller Situation ein messbarer Rückgang der Varietätendistanz. Eine generelle Tendenz lässt sich auch darin finden, dass signifikante Ergebnisse in diesem Untersuchungssetting vermehrt auftreten, wenn man die Interventionsgruppen alleine betrachtet. In den Interaktionen der Testzeitpunkte mit den Gruppen sind Veränderungen seltener feststellbar.

| Varietätendistanz auf Kompetenzebene, detailliert:                      |           |          |             |            |
|---|-----------|----------|-------------|------------|
| Lehrkräfte  | formell   | männlich | Interaktion | IV-Gruppen |
|   |           | weiblich | *           | **         |
|   | informell | männlich |             | **         |
|   |           | weiblich |             |            |
| Schüler:innen   | formell   | männlich |             |            |
|   |           | weiblich |             |            |
|   | informell | männlich |             |            |
|   |           | weiblich |             |            |
| Varietätendistanz auf Kompetenzebene, generell, geschlechterspezifisch: |           |          |             |            |
| Sprecher:innen  |           | männlich |             | *          |
|   |           | weiblich | .           | **         |
| Varietätendistanz auf Kompetenzebene, generell:                         |           |          |             |            |
| Sprecher:innen  |           |          | *           | ***        |

**Abbildung 11.45:** Die Varietätendistanz auf Kompetenzebene

Auf Kompetenzebene (Abb. 11.45) liegt bei den Lehrkräften in formeller Situation ein hochsignifikanter Rückgang der Bewertungsdifferenz der Varietäten in den Interventionsgruppen vor – diese Tendenzen, wenn auch schwächer ausgeprägt und nicht signifikant, findet man jedoch in allen Gruppen und Situationen. Dies spiegelt sich auch in den generellen Ergebnissen wider. So ist für beide Geschlechter ein Distanzrückgang messbar, für die Sprecherinnen ist er hochsignifikant, für die Sprecher signifikant. Das generelle Ergebnis, welches sogar höchstsignifikanten Rückgang der Bewertungsdistanz zwischen den Varietäten anzeigt, und bei dem auch die Relation der Gruppen mit den Testzeitpunkten signifikante Unterschiede zeigt, stellt ein deutliches Indiz dafür dar, dass sich die Kompetenzeinschätzung von dialektsprechenden und standardsprechenden Personen nach erfolgter Intervention aneinander annähert.

| Varietätendistanz auf Sympathieebene, detailliert:                      |           |          |             |            |
|---|-----------|----------|-------------|------------|
|   |           |          | Interaktion | IV-Gruppen |
| Lehrkräfte  | formell   | männlich |             |            |
|   |           | weiblich |             |            |
|   | informell | männlich | **          | **         |
|   |           | weiblich |             |            |
| Schüler:innen   | formell   | männlich |             |            |
|   |           | weiblich |             |            |
|   | informell | männlich |             |            |
|   |           | weiblich |             | *          |
| Varietätendistanz auf Sympathieebene, generell, geschlechterspezifisch: |           |          |             |            |
| Sprecher:innen  |           | männlich |             | *          |
|   |           | weiblich |             |            |
| Varietätendistanz auf Sympathieebene, generell:                         |           |          |             |            |
| Sprecher:innen  |           |          |             | **         |

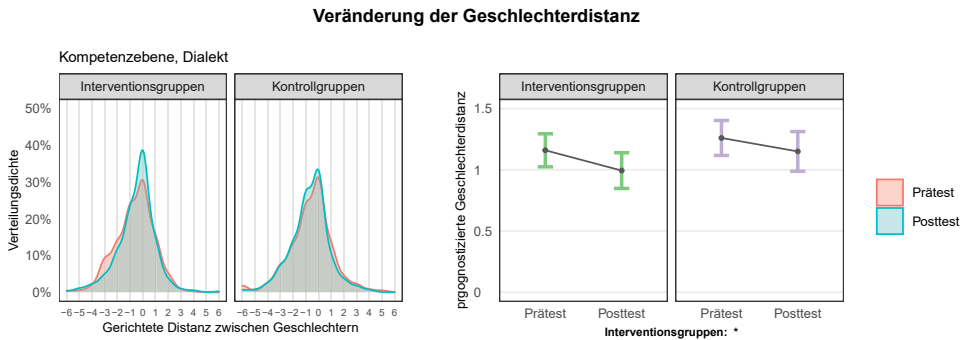
**Abbildung 11.46:** Die Varietätendistanz auf Sympathieebene

Auf Sympathieebene findet man in informeller Situation beim Lehrer und der Schülerin einen signifikanten bzw. hochsignifikanten Rückgang der Bewertungsdifferenz, wobei auch auf dieser Ebene die Tendenzen im Großteil der Interventionsgruppen einen Rückgang der Bewertungsdifferenzen zeigen, wenn auch weniger ausgeprägt. Die geschlechterspezifischen Ergebnisse zeigen dann auch nur bei den Sprechern signifikante Veränderungen innerhalb der Interventionsgruppen, die mit Sicherheit auf den deutlichen Distanzrückgang beim Lehrer in informeller Situation zurückzuführen sind. Das generelle Ergebnis auf Sympathieebene zeigt hochsignifikante Werte, was auf die generellen Tendenzen eines Rückgangs der Bewertungsdifferenz in den Interventionsgruppen zurückzuführen ist. Es kann somit angenommen werden, dass die Intervention auch Auswirkungen auf die Sympathieebene zeigt, insofern, dass Sprecher:innen der Varietäten Dialekt und Standarddeutsch danach ähnlicher auf Sympathieebene eingeschätzt werden.

Auf Ebene der persönlichen Präferenz nimmt die Menge der signifikanten Ergebnisse nochmals ab. Hier ist nur beim Lehrer in formeller Situation ein signifikanter Rückgang der Varietätendistanz innerhalb der Interventionsgruppen messbar und dieser auch als solcher für die Gesamtpopulation prognostizierbar. Das generelle Ergebnis zeigt zwar ebenfalls signifikante Werte, jedoch sind diese mit Vorsicht zu genießen. Innerhalb der Interventionsgruppen liegt zwar ein signifikanter Rückgang der Bewertungsdistanz vor, doch das Ergebnis der Kontrollgruppen unterscheidet sich davon nur minimal (vgl. Abb. 11.44).



Schon einmal vorweg – die Ergebnisse sind nur wenig ergiebig. In den detaillierten Betrachtungen findet man nur einen Distanzrückgang, der sich in der Nähe der Signifikanzschwelle befindet (s. Abb. 11.48) – und zwar bei den dialektsprechenden Lehrkräften in formeller Situation auf Kompetenzebene. In dieser Kategorie liegen sowohl die Unterschiede zwischen den Testzeitpunkten der Interventionsgruppen als auch die Prä-Posttest-Ergebnisse in Relation der Gruppen zueinander nahe an der Signifikanzgrenze, was sowohl in der deskriptiven Darstellung als auch im Effektplot gut zu erkennen ist. Ansonsten findet man bei Betrachtung der einzelnen Guises keine Ergebnisse, die auf eine Veränderung durch die Intervention hindeuten.

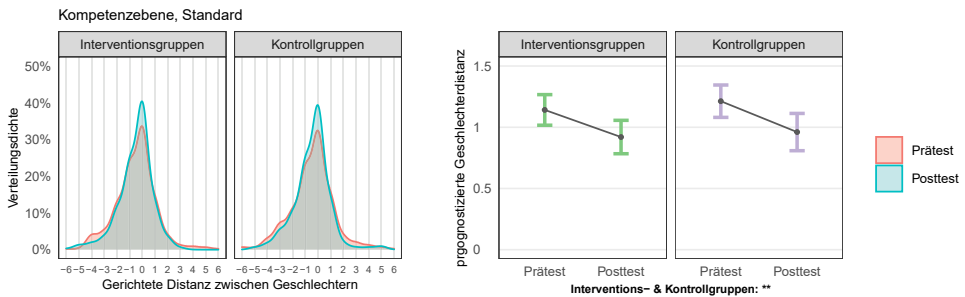


**Abbildung 11.49:** Prätest-Posttest-Vergleich der Distanz der allgemeinen Geschlechtereinschätzung bei der Frage nach der Kompetenz und dialektaler Sprachverwendung

Etwas anders stellt sich die Situation dar, wenn man die Rollen und die Formalitätsgrade nicht mehr differenziert betrachtet. Abbildung 11.49 zeigt die Distanzveränderung zwischen den Geschlechtern bei Kompetenzbeurteilung dialektaler Sprechweise. Die Interventionsgruppen weisen dabei einen signifikanten Rückgang der Distanz auf, allerdings sieht das Ergebnis der Kontrollgruppen recht ähnlich aus. Man kann zwar erkennen, dass die Posttest-Einschätzung der Interventionsgruppen deutlich unter derjenigen der Kontrollgruppen liegt und auch der Rückgang in ersteren Gruppen ausgeprägter ist, die Tatsache, dass in den Kontrollgruppen ebenfalls ein Rückgang der Bewertungsdistanz vorliegt, ist jedoch für die Interpretation problematisch.

Noch problematischer wird es, wenn man das allgemeine Ergebnis für die Bewertungsdistanz bei standardsprachlicher Sprachverwendung betrachtet (Abb. 11.50). Dort ist zwar der Rückgang in den Interventionsgruppen sogar hochsignifikant, allerdings ist der Rückgang in den Kontrollgruppen sogar noch leicht ausgeprägter und ebenfalls hochsignifikant. Die Interventionsgruppen zeigen zwar in den Posttestergebnissen niedrigere Distanzwerte, allerdings zeigen sie das auch in den Prätestwerten. Eine Wirkung der Intervention unter diesen Voraussetzungen zu prognostizieren, ist nicht möglich. Das generelle Ergebnis über beide Varietäten hinweg zeigt für die Interventionsgruppen zwar sogar höchstsignifikante Werte, die Kontrollgruppen unterscheiden sich jedoch nicht stark genug – die Werte hier sind hochsignifikant unterschiedlich. Eine Distanzverringierung zwischen den Geschlechtern ist zwar messbar, allerdings auch in

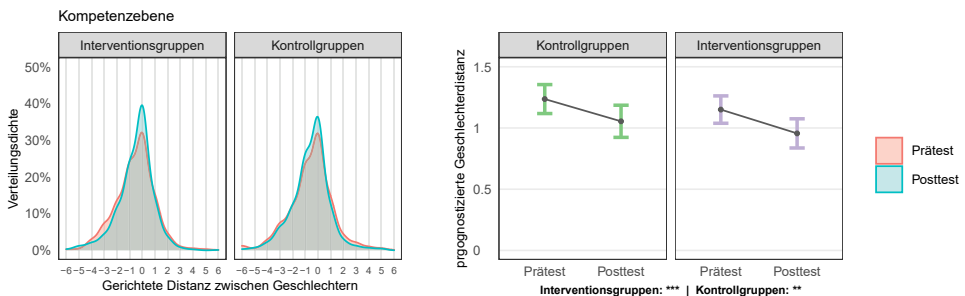
## Veränderung der Geschlechterdistanz



**Abbildung 11.50:** Prätest-Posttest-Vergleich der Distanz der allgemeinen Geschlechtereinschätzung bei der Frage nach der Kompetenz und standardsprachlicher Sprachverwendung

den Kontrollgruppen, wenn auch etwas schwächer ausgeprägt. Nur in isolierten Bereichen wie in Abbildung 11.48 kann die Interpretation gewagt werden, dass hier Wirkungen der Intervention sichtbar werden. Da sich diese Ergebnisse jedoch nur in der Nähe der Signifikanzschwelle befinden, muss dies mit Vorsicht geschehen, allerdings fügt sich das Ergebnis nahtlos in die Ergebnisse der beiden vorhergehenden Kapitel.

## Veränderung der Geschlechterdistanz



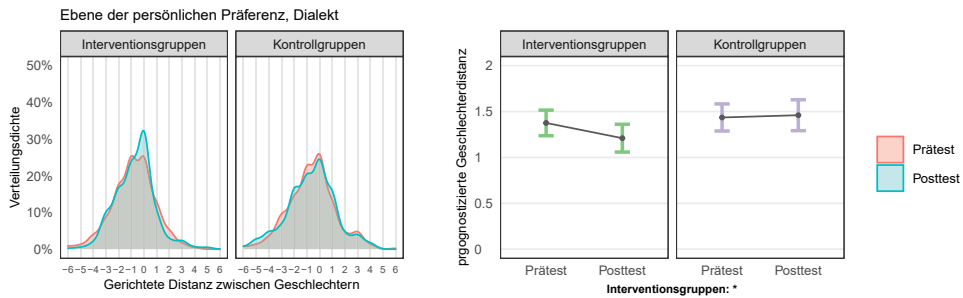
**Abbildung 11.51:** Prätest-Posttest-Vergleich der Distanz der allgemeinen Geschlechtereinschätzung bei der Frage nach der Kompetenz

Auf Sympathieebene sind weder in den detaillierten noch verallgemeinerten Betrachtungen signifikante Veränderungen messbar. Es finden sich zwar Tendenzen für den Distanzrückgang bei dialektaler Sprechweise in formeller Situation, allerdings befinden sich diese nicht in der Nähe der Signifikanzschwelle.

Auf Ebene der persönlichen Präferenz findet man ebenfalls keine signifikanten Unterschiede bei detaillierter Betrachtung der Ergebnisse, allerdings auf genereller Ebene bei dialektaler Sprechweise (s. Abb. 11.52). Hier liegt ein signifikanter Distanzrückgang innerhalb der Interventionsgruppen vor. Das Ergebnis in Relation zu den Kontrollgruppen ist zwar nicht signifikant, allerdings sieht man am Effektplot, dass

hier kein Distanzrückgang vorliegt. Es kann also angenommen werden, dass die Intervention einen Rückgang der Bewertungsdistanz zwischen dialektsprechenden Sprechern und Sprecherinnen auf Ebene der persönlichen Präferenz fördert.

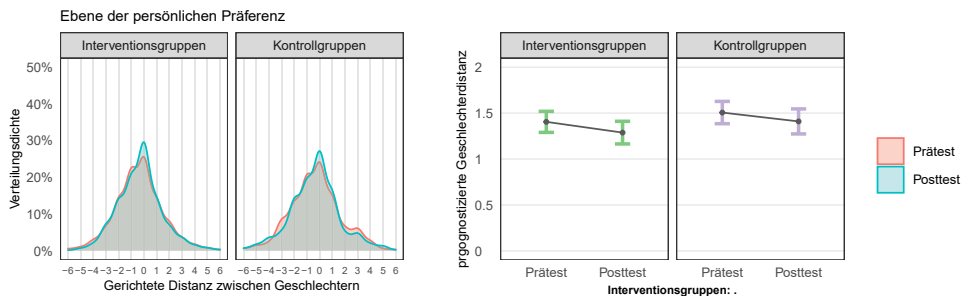
#### Veränderung der Geschlechterdistanz



**Abbildung 11.52:** Prätest-Posttest-Vergleich der Distanz der allgemeinen Geschlechtereinschätzung bei der Frage nach der persönlichen Präferenz und dialektaler Sprachverwendung

Nimmt man beide Varietäten zusammen (Abb. 11.53), verändert sich das Bild wiederum. Die Bewertungsdistanz innerhalb der Interventionsgruppen geht zwar ebenfalls zurück, allerdings befinden sich die Werte nur in der Nähe der Signifikanzgrenze. Gleichzeitig findet man auch in den Kontrollgruppen einen Distanzrückgang, der jedoch nicht signifikant ist.

#### Veränderung der Geschlechterdistanz



**Abbildung 11.53:** Prätest-Posttest-Vergleich der Distanz der allgemeinen Geschlechtereinschätzung bei der Frage nach der persönlichen Präferenz

Zusammenfassend ist festzustellen, dass sich die vorliegenden Bewertungsunterschiede der Geschlechter nicht deutlich verändern. Die Tendenzen sind zumeist nicht signifikant und bei Vorliegen von Signifikanzwerten häufig nicht von den Ergebnissen der Kontrollgruppe unterscheidbar. Einzige Ausnahme stellt die Kompetenzeinschätzung der dialektsprechenden Lehrkräfte in formeller Situation dar. Auch wenn sich die Ergebnisse nur in der Nähe der Signifikanzgrenze befinden, passen sie sehr gut zu

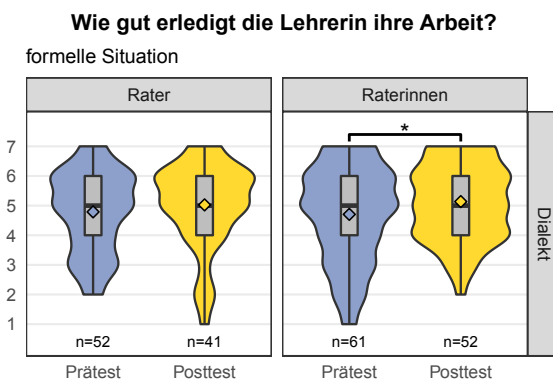
denjenigen aus Abschnitt 11.1 und 11.2. Dazu passt auch das Ergebnis der persönlichen Präferenz zu dialektaler Sprechweise, in der ein signifikanter Rückgang der Bewertungsdifferenz zwischen den Geschlechtern zu verzeichnen ist.<sup>35</sup>

## 11.4 Raterinnen und Rater

In Abschnitt 9.3 sind vor allem zwei Ergebnisse ins Auge gestochen: Die Raterinnen schätzen generell etwas positiver ein als die Rater, tendenziell schätzen die Geschlechter aber recht ähnlich ein, unabhängig davon, ob Sprecherinnen oder Sprecher bewertet werden. Dies muss aber nicht bedeuten, dass die Intervention bei den Rater:innen die gleichen Effekte zeigt. Ob es hier Auffälligkeiten gibt, sehen wir uns in diesem Kapitel an.

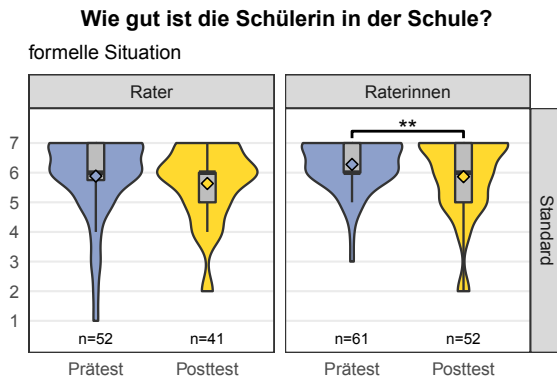
### 11.4.1 Detaillierte Ergebnisse

Ein erstes signifikantes Ergebnis findet man bei Betrachtung der dialektspredenden Lehrerin in formeller Situation (s. Abb. 11.54). Die Raterinnen schätzen sie nach der Intervention signifikant kompetenter ein. Auch bei den Ratern ist eine Aufwärtsbewegung erkennbar, diese ist jedoch weniger ausgeprägt und auch nicht signifikant.



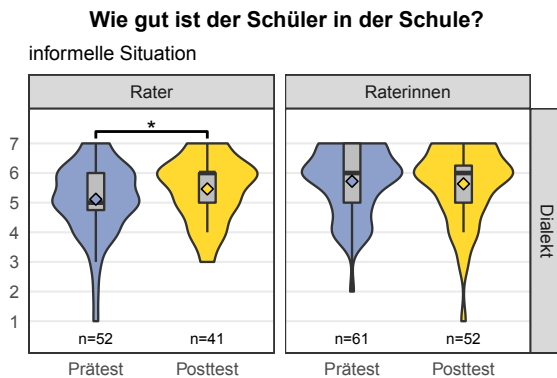
**Abbildung 11.54:** Prätest-Posttest-Vergleich der Rater:innen bei der Frage nach der Kompetenz der dialekt-sprechenden Lehrerin in formeller Situation

<sup>35</sup> Für den Rückgang der Bewertungsdistanz in den Kontrollgruppen könnte auch die Wahl der Likert-Skala eine Erklärung bieten. Da es auf der siebenteiligen Skala einen Wert gibt, der exakt in der Mitte liegt, könnte es dazu geführt haben, dass Rater:innen in der Posttestung aus Langeweile oder fehlendem Interesse, die Umfrage auszuführen, diesen Wert vermehrt angaben. Da die Kontrollgruppen nicht den Mehrwert der Unterrichtsmaterialien hatten, liegt es nahe, dass dies in diesen Gruppen häufiger vorkam.



**Abbildung 11.55:** Prätest-Posttest-Vergleich der Rater:innen bei der Frage nach der Kompetenz der standard-sprechenden Schülerin in formeller Situation

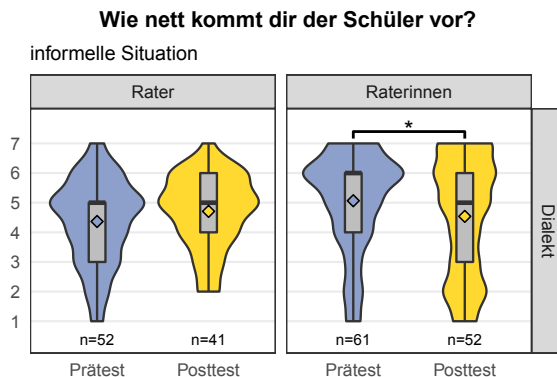
Die standardsprechende Schülerin in formeller Situation (Abb. 11.55) wird im Posttest hingegen von den Raterinnen hochsignifikant weniger kompetent eingeschätzt. Allerdings lagen die Prätestwerte in einem sehr hohen Bereich – einer der höchsten aller gemessenen Werte. Diese Einschätzung geht zurück und nähert sich somit der Einschätzung des dialektsprachlichen Guises an. Auch bei den Ratern ist diese Bewegung erkennbar. Die Prätestwerte liegen allerdings nicht ganz so hoch, daher ist der Rückgang auch weniger ausgeprägt.



**Abbildung 11.56:** Prätest-Posttest-Vergleich der Rater:innen bei der Frage nach der Kompetenz des dialekt-sprechenden Schülers in informeller Situation

Beim dialekt-sprechenden Schüler in informeller Situation ist zu beobachten, dass er von den Ratern nach der Intervention signifikant kompetenter eingeschätzt wird (Abb. 11.56). Die Werte der Rater nähern sich damit denjenigen der Raterinnen an und bleiben nur in geringem Maße darunter. Auch die Prätestwerte dieser Abbildung sind interessant, zeigen sie doch, dass die Raterinnen den dialekt-sprechenden Schüler kom-

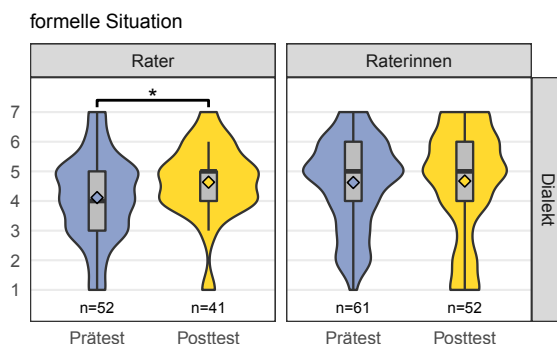
petenter einschätzen als die Rater – das männliche Geschlecht wird vom weiblichen also kompetenter eingeschätzt als vom männlichen selbst, ein auffälliger Umstand, der schon in Abschnitt 9.3 besprochen wurde.



**Abbildung 11.57:** Prätest-Posttest-Vergleich der Rater:innen bei der Frage nach der Sympathie des dialekt-sprechenden Schülers in informeller Situation

Der dialekt-sprechende Schüler in informeller Situation wird darüber hinaus aber auch von den Raterinnen nach der Intervention signifikant weniger sympathisch eingeschätzt (Abb. 11.57). Auch in diesem Ergebnis ist zu erkennen, dass sich die doch recht unterschiedlichen Werte zwischen den Geschlechtern der Prätestung in der Posttestung aneinander annähern – bei den Ratern nehmen die Sympathiewerte sogar zu.

### Wie gerne hättest du den Schüler als Tischnachbar?

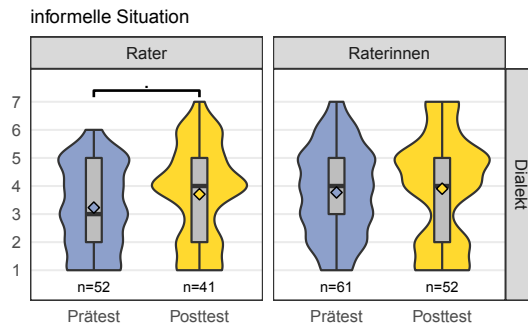


**Abbildung 11.58:** Prätest-Posttest-Vergleich der Rater:innen bei der Frage nach persönlicher Präferenz des dialekt-sprechenden Schülers in formeller Situation

Auf Ebene der persönlichen Präferenz findet man sowohl beim Schüler als auch bei der Schülerin einen Prä-Posttest-Unterschied. Der dialekt-sprechende Schüler (s. Abb. 11.58) wird von den Ratern nach der Intervention signifikant positiver eingeschätzt. Wiederum ist zu bemerken, dass sich die Posttestergebnisse der Raterinnen und Rater aneinander

annähern. Ähnlich verhält es sich bei der dialektsprechenden Schülerin in informeller Situation (s. Abb. 11.59). Sie wird von den Ratern ebenfalls nach der Intervention positiver eingeschätzt, allerdings befindet sich dieses Ergebnis nur in der Nähe der Signifikanzgrenze. Dennoch ist wiederum zu beobachten, dass sich die Posttestergebnisse beider ratenden Geschlechter aneinander annähern.

#### Wie gerne hättest du die Schülerin als Tischnachbarin?

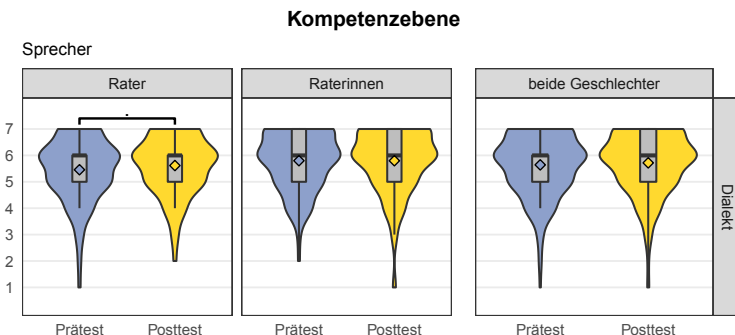


**Abbildung 11.59:** Prätest-Posttest-Vergleich der Rater:innen bei der Frage nach persönlicher Präferenz der dialektsprechenden Schülerin in informeller Situation

### 11.4.2 Generelle Ergebnisse

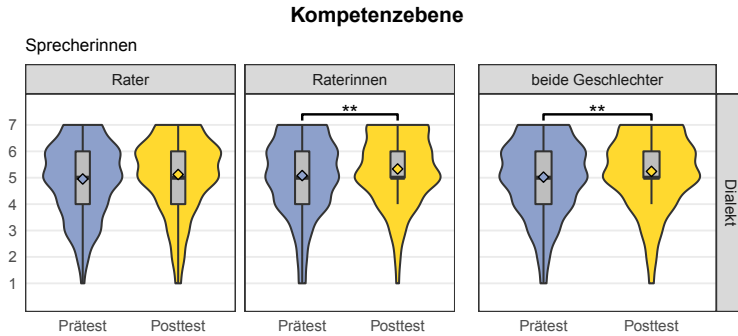
Wie verhält es sich nun, wenn man die generellen Ergebnisse betrachtet? Speziell auf Kompetenzebene sind hier durchaus interessante Erkenntnisse zu gewinnen.

Abbildung 11.60 zeigt die Ergebnisse für die dialektsprechenden Sprecher auf Kompetenzebene. In diesem Plot sind keine großen Veränderungen bemerkbar, einzig die Rater schätzen die Sprecher im Posttest kompetenter ein – dieses Ergebnis liegt jedoch nur in der Nähe der Signifikanzgrenze. Auch in diesem Plot ist gut zu erkennen, dass die Raterinnen dialektsprechende Sprecher höher einschätzen als die Rater. Der allgemeine Plot, der beide Geschlechter beinhaltet, zeigt keine Unterschiede auf.



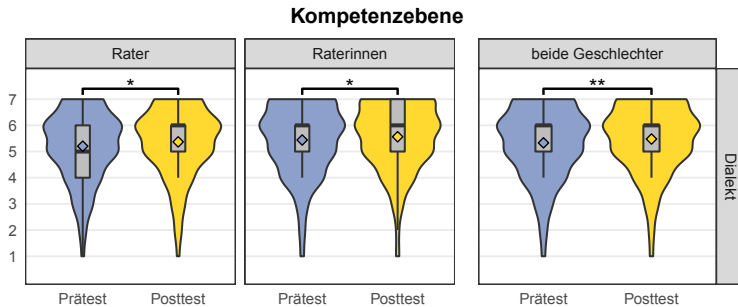
**Abbildung 11.60:** Allgemeiner Prätest-Posttest-Vergleich der Rater:innen bei der Frage nach der Kompetenz der Dialekt Sprecher

Bei den dialektsprechenden Sprecherinnen (s. Abb. 11.61) erkennt man, dass sie von den Raterinnen nach der Intervention hochsignifikant kompetenter eingeschätzt werden. Auch bei den Ratern ist eine leichte Bewegung hin zur kompetenteren Einschätzung bemerkbar, diese ist allerdings nicht signifikant.



**Abbildung 11.61:** Allgemeiner Prätest-Posttest-Vergleich der Raterinnen bei der Frage nach der Kompetenz der Dialektsprecherinnen

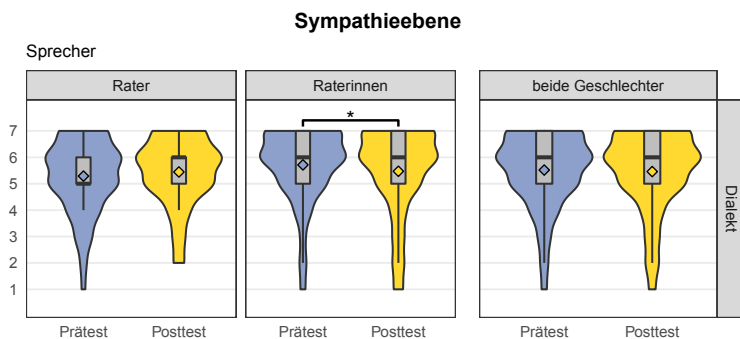
Wenn man beide Geschlechter gemeinsam betrachtet, überträgt sich das hochsignifikante Ergebnis von den Raterinnen auf die Gesamtbetrachtung. Man kann also prognostizieren, dass dialektsprechende Sprecherinnen mit hoher Wahrscheinlichkeit nach erfolgter Intervention hochsignifikant kompetenter eingeschätzt werden – und zwar vor allem von den Raterinnen.



**Abbildung 11.62:** Allgemeiner Prätest-Posttest-Vergleich der Raterinnen bei der Frage nach der Kompetenz der Dialektsprecher:innen

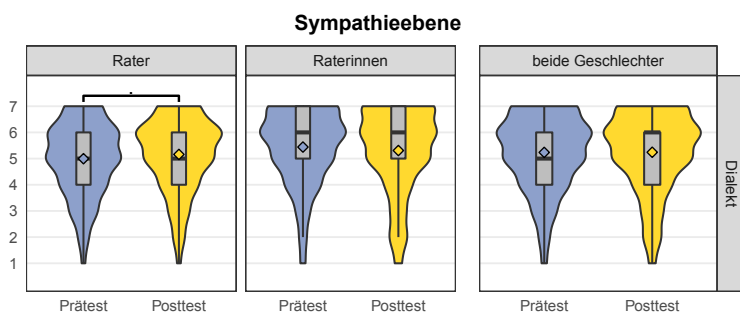
Eine weitere Analysemöglichkeit in diesem Kapitel liegt darin, dass man auf die Unterscheidung von Sprecherinnen und Sprechern verzichtet und sozusagen die Gesamteinschätzung der Raterinnen und Rater auf den einzelnen Ebenen und in den beiden Varietäten betrachtet. In Abbildung 11.62 wurde dies auf Kompetenzebene realisiert. Man erkennt, dass sowohl die Rater als auch die Raterinnen nach der Intervention dialektsprechende Personen signifikant kompetenter einschätzen. Wenn man beide Ge-

schlechter zusammenfasst, ergeben sich sogar hochsignifikante Unterschiede, was wahrscheinlich auf das Gesetz der großen Zahlen zurückzuführen ist. Nichtsdestotrotz lässt sich annehmen, dass sowohl Rater als auch Raterinnen durch die Intervention dialektsprechende Personen signifikant kompetenter einschätzen als zuvor.



**Abbildung 11.63:** Allgemeiner Prätest-Posttest-Vergleich der Raterinnen bei der Frage nach der Sympathie der Dialektsprecher

Auf Sympathieebene (s. Abb. 11.63) werden die dialektsprechenden Sprecher von den Raterinnen im Posttest signifikant weniger sympathisch eingeschätzt. Dies hängt wahrscheinlich damit zusammen, dass die Prätestergebnisse der Raterinnen deutlich höher ausfallen als die der Rater und sich die Posttestergebnisse beider Geschlechter aneinander annähern. Bei den Ratern sieht man nämlich sogar einen leichten Zuwachs der Sympathiewerte. Da zwischen den Geschlechtern die Werte in gegenteilige Richtungen gehen, ist es auch nicht weiter verwunderlich, dass die zusammengefassten Werte beider Geschlechter sehr ähnlich sind und keine signifikanten Veränderungen zeigen.

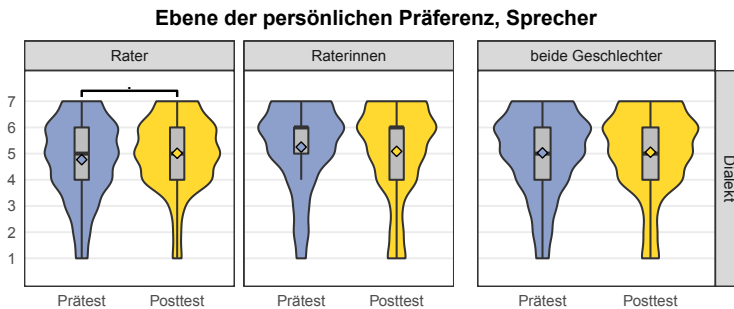


**Abbildung 11.64:** Allgemeiner Prätest-Posttest-Vergleich der Raterinnen bei der Frage nach der Sympathie zu dialektaler Sprachverwendung

Bei den Sprecherinnen sind auf dieser Ebene keine signifikanten Unterschiede vorhanden und auch wenn man Sprecherinnen und Sprecher gemeinsam analysiert (s. Abb. 11.64), sieht man zwar ähnliche Tendenzen wie bei den Sprechern, allerdings kein

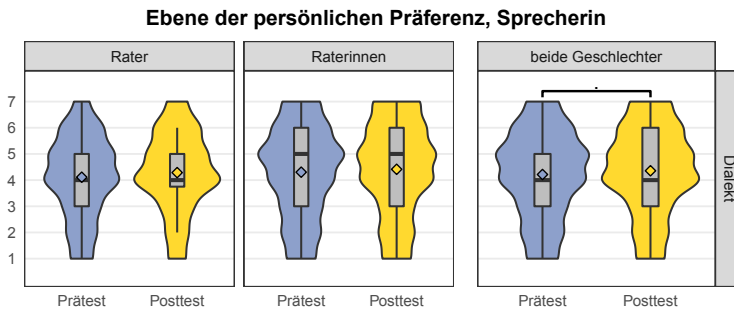
signifikantes Ergebnis. Einzig bei den Ratern findet sich eine Zunahme der Sympathieeinschätzung dialektaler Sprechweise, die sich nahe an der Signifikanzschwelle befindet.

Auf Ebene der persönlichen Präferenz werden die dialektsprechenden Sprecher von den Ratern etwas positiver eingeschätzt (Abb. 11.65). Dieses Ergebnis befindet sich allerdings nur in der Nähe der Signifikanzschwelle. Ansonsten findet man bei den Raterinnen eine leichte Tendenz zur negativeren Bewertung, was die Posttestergebnisse zwischen den Geschlechtern näher zusammenrücken lässt.



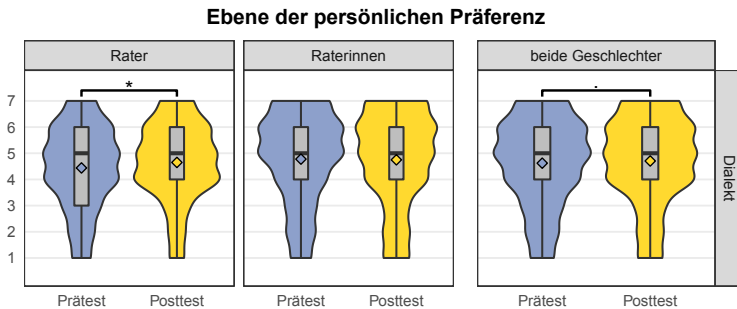
**Abbildung 11.65:** Allgemeiner Prätest-Posttest-Vergleich der Raterinnen bei der Frage der persönlichen Präferenz bezüglich Dialektsprechern

Bei den dialektsprechenden Sprecherinnen (s. Abb. 11.66) geht die Tendenz sowohl bei den Raterinnen als auch bei den Ratern zu einer etwas positiveren Einschätzung auf dieser Ebene. Diese Tendenz nähert sich dann bei einer gemeinsamen Betrachtung der Geschlechter der Signifikanzschwelle, überschreitet diese jedoch nicht.



**Abbildung 11.66:** Allgemeiner Prätest-Posttest-Vergleich der Raterinnen bei der Frage der persönlichen Präferenz bezüglich Dialektsprecherinnen

Etwas anders sieht es aus, wenn man dialektsprechende Sprecherinnen und Sprecher gemeinsam analysiert (Abb. 11.67). Die Rater schätzen dialektale Sprechweise nach der Intervention allgemein positiver ein. Bei den Raterinnen sieht man dagegen keine Veränderung. Allerdings sieht es auch hier so aus, dass sich die negativere Einschätzung der Rater im Prätest derjenigen der Raterinnen im Posttest annähert. Bei Betrachtung beider Geschlechter sieht man eine leichte Tendenz zur positiveren Bewertung dialektaler Sprechweise, die sich in der Nähe der Signifikanzgrenze befindet.



**Abbildung 11.67:** Allgemeiner Prätest-Posttest-Vergleich der Raterinnen bei der Frage der persönlichen Präferenz bezüglich dialektaler Sprachverwendung

### 11.4.3 Zusammenfassung Sprecherinnen und Sprecher

Bei der differenzierten Betrachtung von Raterinnen und Ratern fallen durchaus einige Besonderheiten auf. Standardsprachliche Sprechweise zeigt nur bei der Kompetenzbewertung der Schülerin in formeller Situation eine niedrigere Einschätzung der Raterinnen im Posttest. Da das Prätestergebnis dieser Kategorie jedoch sehr hoch ausfällt, handelt es sich dabei wahrscheinlich um eine Annäherung der Bewertungen von Dialekt und Standard – die standardsprechende Schülerin wird nicht mehr per se als hochkompetent eingeschätzt. Alle anderen signifikanten Ergebnisse betreffen die dialektale Sprechweise. Eine generelle Tendenz ist darin zu erkennen, dass sich die Einschätzungen der Raterinnen und Rater, die im Prätest teils deutlich auseinanderliegen, im Posttest aneinander annähern. Dies könnte darauf hindeuten, dass ein eventueller Genderbias in den Spracheinstellungen abgebaut wird.

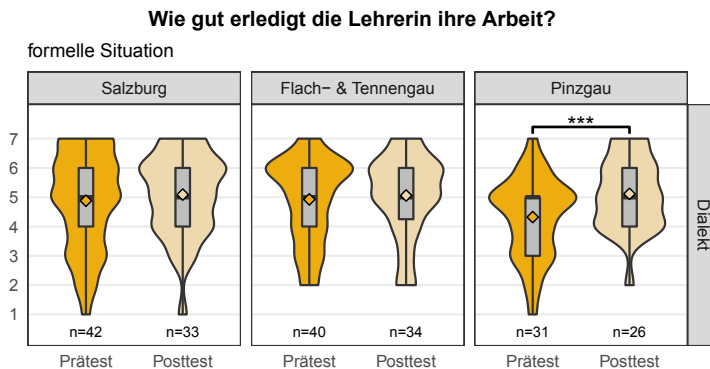
Auf Kompetenzebene sieht man bei beiden Geschlechtern eine Zunahme der Kompetenzeinschätzung des eigenen Geschlechts – bei den Raterinnen noch stärker als bei den Ratern. Dieses Ergebnis ist bemerkenswert, weist es doch darauf hin, dass die Reflexionsprozesse, die durch die Materialien ausgelöst werden sollten, vor allem die Kompetenzeinschätzung des eigenen Geschlechts beeinflussen. Speziell bei den Raterinnen, die das männliche Geschlecht ja kompetenter einschätzen als ihr eigenes, sorgt die hochsignifikante Betterbewertung der Sprecherinnen dafür, dass sich der Abstand der Kompetenzeinschätzung zwischen Sprecherinnen und Sprechern deutlich verringert (vgl. Abb. 11.61).

Dazu passt auch die Einschätzung der Raterinnen auf Sympathieebene. Die Sprecher werden von ihnen in der Posttestung nämlich signifikant weniger sympathisch

eingeschätzt, was ebenfalls darauf hindeutet, dass sich bei den ratenden Schülerinnen ein potenzieller Confirmation Bias durch die Intervention verringert hat. Ansonsten kommen die Effekte sowohl auf Sympathie- als auch auf Ebene der persönlichen Präferenz denjenigen der Kompetenzebene nicht nahe.

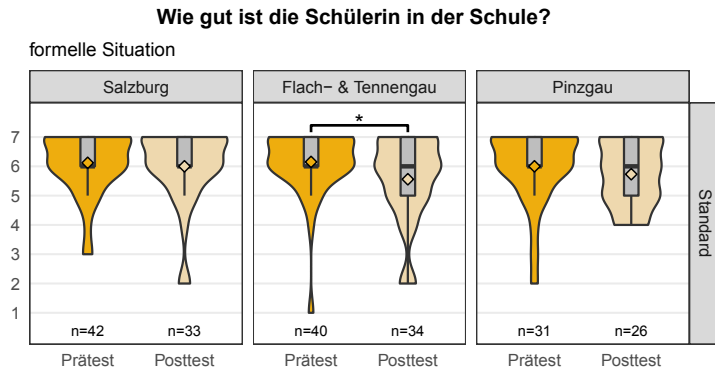
## 11.5 Städtische und ländliche Regionen

In diesem Kapitel werden etwaige Unterschiede im Einschätzungsverhalten und etwaigen Einschätzungsveränderungen in Bezug auf unterschiedliche Regionen beleuchtet. Wie schon in Abschnitt 8.3 beschrieben wurde, sollte die Auswahl der Schulen drei Regionen innerhalb des Landes Salzburg repräsentieren. Zum einen wäre dies der Zentralraum der Stadt Salzburg, dann befinden sich zwei Schulen in Pendeldistanz zum Zentralraum – eine im Flachgau und eine im Tennengau – und zu guter Letzt wurden zwei Schulen außerhalb der Pendeldistanz im Pinzgau gefunden. Durch dieses regionale Setting kann nun überprüft werden, ob die Nähe zum Zentralraum die Wirksamkeit der Intervention beeinflusst.



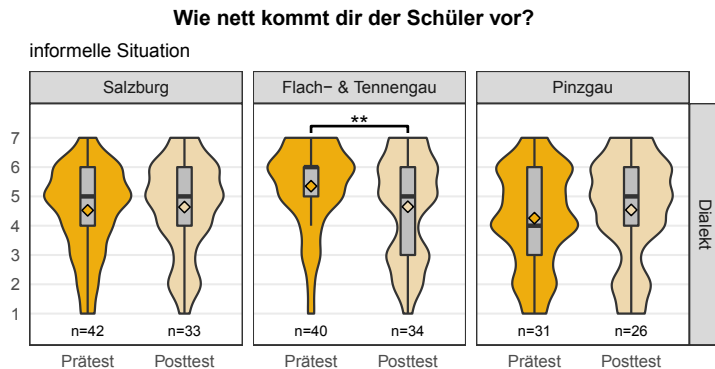
**Abbildung 11.68:** Prätest-Posttest-Vergleich des Stadt-Land-Unterschieds bei der Frage nach der Kompetenz der dialektprechenden Lehrerin in formeller Situation

Auf Kompetenzebene findet man zwei signifikante Ergebnisse, eines bei den Lehrkräften und eines bei den Schüler:innen. Abbildung 11.68 zeigt die Ergebnisse der dialektprechenden Lehrerin in formeller Situation. Man kann bei allen drei Gruppen eine Tendenz zur positiveren Einschätzung nach erfolgter Intervention erkennen. Allerdings weist nur das Ergebnis der Rater:innen des Pinzgaus höchstsignifikante Differenzwerte auf. Die positivere Einschätzung ist in dieser Region auch deskriptiv am ausgeprägtesten. Interessant ist nun jedoch, dass sich der Pinzgau speziell in den Prätestergebnissen im Vergleich zu den beiden anderen Gruppen unterscheidet. Diese liegen deutlich unter den beiden anderen Regionen. Die Posttestergebnisse der drei Gruppen werden sich ähnlicher.



**Abbildung 11.69:** Prätest-Posttest-Vergleich des Stadt-Land-Unterschieds bei der Frage nach der Kompetenz der standardsprechenden Schülerin in formeller Situation

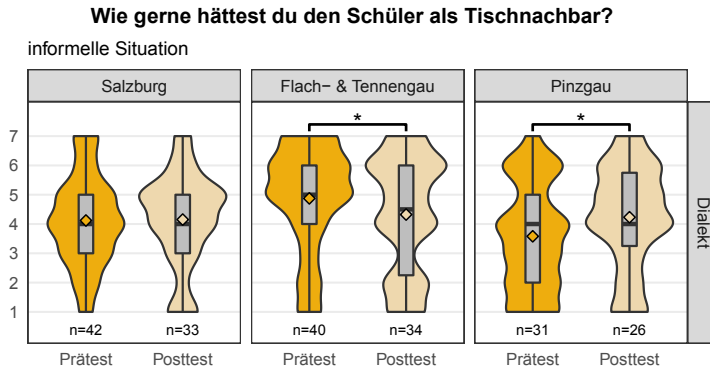
Bei der standardsprechenden Schülerin in formeller Situation (s. Abb. 11.69) sehen die Ergebnisse anders aus. Hier sind in Salzburg keine Unterschiede erkennbar. Die Werte in dieser Region liegen sehr hoch, was die hohe Kompetenzzuschreibung standard-sprachlicher Varietätenverwendung in den Stadtschulen abbildet. Diese bleibt auch nach erfolgter Intervention praktisch unverändert. In den beiden anderen Regionen ist hingegen ein Rückgang der Einschätzungen im Posttest von ebenfalls sehr hohen Prätestwerten erkennbar – im Flach- & Tennengau ist dieser Rückgang signifikant. Dies deutet darauf hin, dass die Intervention in beiden letzteren Regionen dafür gesorgt hat, die standardsprechende Schülerin nicht pauschal hochkompetent einzuschätzen, sondern diese Einschätzung reflektierter anzugehen. Damit nähern sich die Werte dieser Gruppen auch denjenigen der Dialektguises an.



**Abbildung 11.70:** Prätest-Posttest-Vergleich des Stadt-Land-Unterschieds bei der Frage nach der Sympathie des dialektssprechenden Schülers in informeller Situation

Auf Sympathieebene wird der dialektssprechende Schüler in informeller Situation von den Rater:innen der Region Flach- & Tennengau nach erfolgter Intervention hochsig-

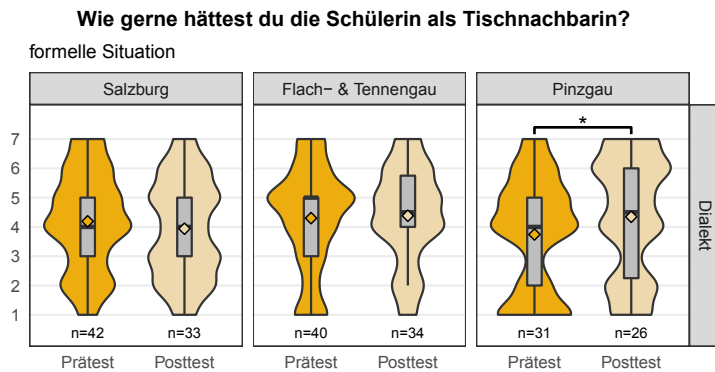
nifikant weniger sympathisch eingeschätzt (s. Abb. 11.70). Wenn man dieses Ergebnis allerdings mit den anderen Gruppen vergleicht, fällt auf, dass die Prätestergebnisse in den Schulen des Flach- & Tennengaus besonders hoch ausgefallen sind. Durch den Rückgang der Einschätzungen in dieser Region nähern sich die Posttestergebnisse aller Regionen damit stark aneinander an.



**Abbildung 11.71:** Prätest-Posttest-Vergleich des Stadt-Land-Unterschieds bei der Frage nach persönlicher Präferenz zum dialektsprechenden Schüler in informeller Situation

Auf Ebene der persönlichen Präferenz lassen sich in dialektaler Sprachverwendung sowohl bei der Schülerin als auch beim Schüler Unterschiede finden. Beim Schüler (s. Abb. 11.71) sieht man in informeller Situation sowohl in den Schulen des Flach- & Tennengaus als auch in denen des Pinzgaus eine signifikante Veränderung – allerdings in unterschiedliche Richtungen. Im Flach- & Tennengau fällt das Prätestergebnis – wie schon im Diagramm davor – sehr hoch aus; im Pinzgau dagegen stellt es das niedrigste aller drei Gruppen dar. Wiederum scheint es so zu sein, dass sich die sehr unterschiedlichen Einschätzungen der Regionen im Prätest im Posttest deutlich aneinander annähern.

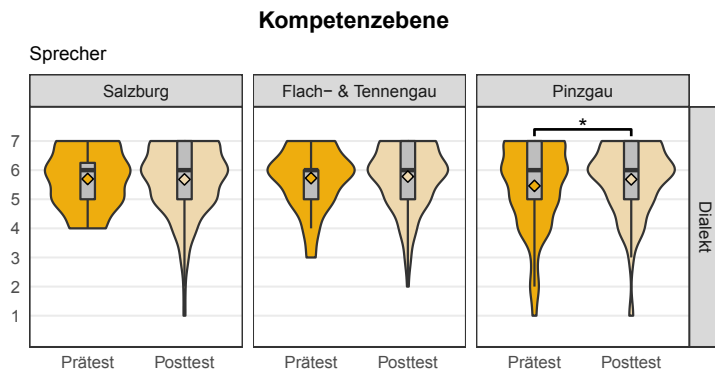
Die dialektsprechende Schülerin in formeller Situation (s. Abb. 11.72) wird nach erfolgter Intervention im Pinzgau signifikant positiver eingeschätzt. Im Vergleich mit den anderen Regionen fällt auf, dass das Prätestergebnis des Pinzgaus deutlich unter denjenigen der beiden anderen Gruppen liegt. Das Posttestergebnis ähnelt dann demjenigen der Rater:innen des Flach- & Tennengaus. Das Posttestergebnis in der Stadt Salzburg liegt unter den beiden anderen Regionen und wird im Vergleich zum Prätestergebnis in dieser Region sogar leicht negativer. Die persönliche Präferenzeinschätzung zur dialektsprechenden Schülerin zeigt also regionale Effekte. Wird sie im Pinzgau nach der Intervention signifikant positiver eingeschätzt und liegen die Einschätzungswerte auch im Flach- & Tennengau in diesem Bereich, kann man in der Stadt keine verbesserte Einschätzung wahrnehmen.



**Abbildung 11.72:** Prätest-Posttest-Vergleich des Stadt-Land-Unterschieds bei der Frage nach persönlicher Präferenz zur dialektprechenden Schülerin in formeller Situation

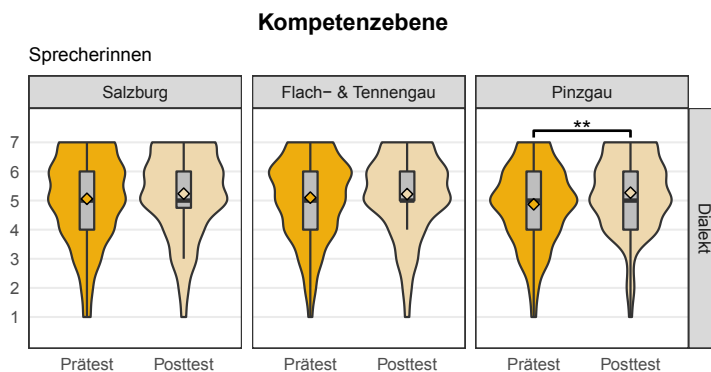
### 11.5.1 Generelle Ergebnisse

In einem nächsten Schritt werden nach schon bekanntem Muster die Situationen und Rollen nicht mehr unterschieden, um zu generellen Ergebnissen zu gelangen. Auf Kompetenzebene findet man Ergebnisse sowohl bei den Dialektsprechern als auch bei den Dialektsprecherinnen.



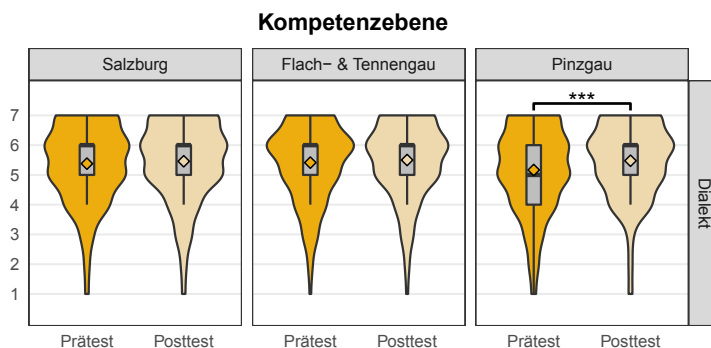
**Abbildung 11.73:** Allgemeiner Prätest-Posttest-Vergleich des Stadt-Land-Unterschieds bei der Frage nach der Kompetenz bezüglich Dialektsprechern

Die Dialektsprecher (Abb. 11.73) werden vergleichsweise kompetent eingeschätzt – die Rater:innen der Stadt Salzburg und des Flach- & Tennengaus zeigen dabei zwischen den Testzeitpunkten keine großen Veränderungen in ihrem Abstimmungsverhalten. Im Pinzgau liegen die Prätestergebnisse unter den beiden anderen Regionen. Sie steigen zum Posttest hin jedoch signifikant an und ähneln dann den Werten der beiden anderen Regionen.



**Abbildung 11.74:** Allgemeiner Prätest-Posttest-Vergleich des Stadt-Land-Unterschieds bei der Frage nach der Kompetenz bezüglich Dialektsprecherinnen

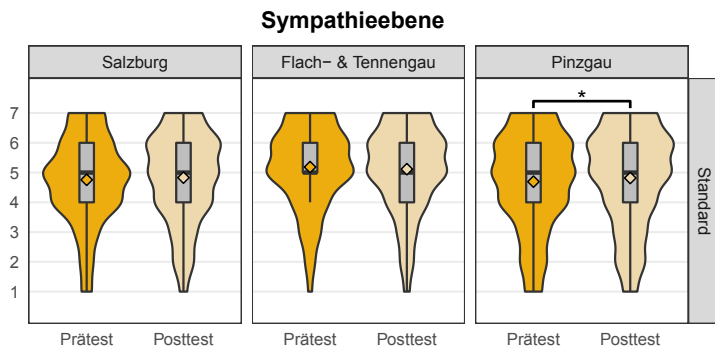
Bei den Dialektsprecherinnen (Abb. 11.74) liegen die Prätestwerte aller Regionen unter denen der Sprecher. Ebenfalls vergeben die Rater:innen des Pinzgaus die niedrigsten Werte. Zur Posttestung hin nehmen die Werte in allen drei Gruppen zu. Bei den Rater:innen des Pinzgaus ist die Zunahme allerdings am ausgeprägtesten und zeigt hochsignifikante Unterschiede. Dass die Intervention dafür sorgt, dass dialektsprechende Sprecherinnen in ländlichen Regionen deutlich kompetenter eingeschätzt werden und diese Einschätzungsveränderung dort auch stärker ausfällt als im Ballungsraum, ist also sehr wahrscheinlich.



**Abbildung 11.75:** Allgemeiner Prätest-Posttest-Vergleich des Stadt-Land-Unterschieds bei der Frage nach der Kompetenz bezüglich dialektaler Sprachverwendung

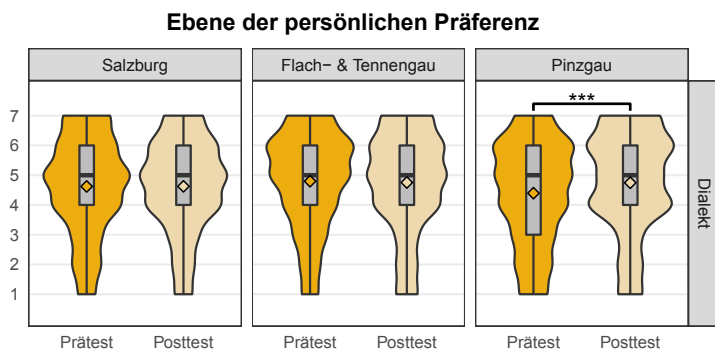
Wenn man nun ebenfalls die Geschlechterdifferenzierung auflöst (s. Abb. 11.75), sieht man, dass dialektsprechende Sprecher nach der Intervention in den Schulen des Pinzgaus höchstsignifikant kompetenter eingeschätzt werden. Wie schon bei den Sprecherinnen liegen die Prätestwerte dieser Region allerdings auch unter denjenigen der beiden anderen Regionen. Eine Tendenz zur kompetenteren Einschätzung nach der Intervention findet sich aber auch bei den beiden anderen Gruppen. Es ist also wieder-

rum wahrscheinlich, dass die Intervention das Potenzial hegt, dialektsprechende Personen kompetenter einzuschätzen. Am Land scheint diese Wirkung am stärksten auszufallen.



**Abbildung 11.76:** Allgemeiner Prätest-Posttest-Vergleich des Stadt-Land-Unterschieds bei der Frage nach der Sympathie bezüglich standardsprachlicher Sprachverwendung

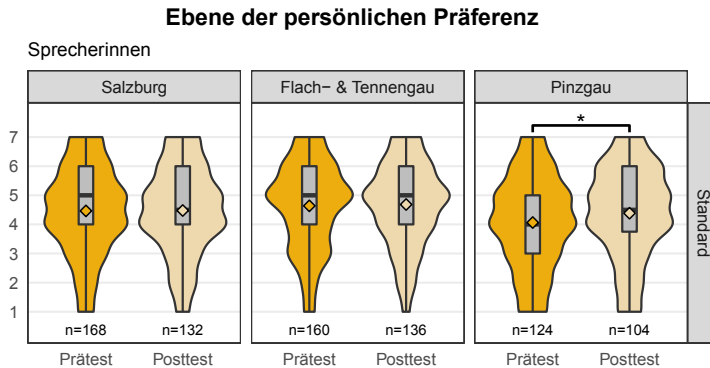
Die Sympathieebene offenbart nur im geschlechterindifferenten Diagramm eine signifikante Besserbewertung der standardsprechenden Sprecher:innen der Rater:innen des Pinzgaus (s. Abb. 11.76). Auffällig auf dieser Ebene ist, dass in den beiden Regionen Salzburg und Pinzgau deutlich negativere Sympathieeinschätzungen abgegeben wurden als dies im Flach- & Tennengau erfolgt ist. Ebenfalls ist in diesen beiden Gruppen ein Anstieg der Bewertungen zum Posttest hin zu erkennen, im Flach- & Tennengau sieht man dagegen jedoch einen leichten Rückgang. Die Posttestwerte aller Gruppen gleichen sich somit aneinander an.



**Abbildung 11.77:** Allgemeiner Prätest-Posttest-Vergleich des Stadt-Land-Unterschieds bei der Frage nach persönlicher Präferenz bezüglich dialektaler Sprachverwendung

Auf Ebene der persönlichen Präferenz zeigt sich in der geschlechterindifferenten Abbildung eine höchstsignifikant positivere Einschätzung der dialektalen Sprechweise im Pinzgau (Abb. 11.77). Besonders auffallend in dieser Region ist der Umstand, dass die

negativen Bewertungen 1–3 stark zurückgehen. Bei den anderen beiden Gruppen kann man keine Veränderung in den arithmetischen Mittelwerten erkennen. Die Werte, die die Jugendlichen der Stadtschulen abgegeben haben, liegen etwas unter denen der beiden anderen Regionen.



**Abbildung 11.78:** Allgemeiner Prätest-Posttest-Vergleich des Stadt-Land-Unterschieds bei der Frage nach persönlicher Präferenz bezüglich Standardsprecherinnen

Man findet auf Ebene der persönlichen Präferenz bei den standardsprechenden Sprecherinnen noch ein weiteres signifikantes Ergebnis. Hier schätzen die Rater:innen des Pinzgaus im Posttest dialektale Sprachverwendung signifikant positiver ein. Wie schon auf Sympathieebene zeigt die Gruppe Flach- & Tennengau deutlich höhere Einschätzungswerte als die beiden anderen Gruppen.

### 11.5.2 Zusammenfassung städtische und ländliche Regionen

Sowohl in den detaillierten als auch in den generellen Ergebnissen fallen Tendenzen besonders bei zwei Gruppen ins Auge: dem Flach- & Tennengau und dem Pinzgau. Im Flach- & Tennengau findet man in den Kategorien, die signifikante Ergebnisse zeigen, durchwegs Prätestwerte vor, die deutlich höher liegen als diejenigen der beiden anderen Regionen. Die Ergebnisse werden dann zur Posttestung hin niedriger und gleichen sich somit den Werten der beiden anderen Gruppen an. Dabei ist dieses Verhalten nicht etwa auf eine bestimmte Sprechweise, Situation oder Ebene beschränkt. Sowohl auf Kompetenzebene in formeller Situation bei der standardsprechenden Schülerin als auch auf Sympathieebene und Ebene der persönlichen Präferenz in informeller Situation beim dialektalsprechenden Schüler kann dieses Verhalten beobachtet werden.

Im Pinzgau verhält es sich genau andersherum. Hier liegen die Prätestergebnisse oft deutlich unter denjenigen der anderen beiden Gruppen. Zum Posttest hin kommt es dann zu einer positiveren Einschätzung, die durch die niedrigen Prätestwerte signifikante bis höchstsignifikante Veränderungen zeigt. Dabei werden die anderen Gruppen in ihren Werten jedoch nie deutlich überschritten, es findet eher eine Angleichung an deren Werte statt.

Ein weiterer auffälliger Unterschied liegt in den deutlich positiveren Einschätzungen der Rater:innen aus dem Flach- & Tennengau gegenüber der standardsprachlichen Sprechweise auf Sympathie- und Ebene der persönlichen Präferenz. Dies ist so erklärbar, dass durch die regionale Nähe zur Stadt Salzburg auch die dort häufigere Verwendung der Standardsprache bekannt sein dürfte und daher die Bedeutung, diese Varietät zu erwerben, ausgeprägter sein müsste. Das Messergebnis, dass auf Sympathie- und Ebene der persönlichen Präferenz die Standardvarietät in dieser Gruppe deutlich positiver eingeschätzt wird, würde dies sozusagen indirekt anzeigen.

Die Messergebnisse sind jedoch im Kontrast zur zugrunde liegenden Hypothese H<sub>5</sub> (Schüler:innen am Land schätzen dialektsprechende Personen kompetenter ein als Schüler:innen in der Stadt. Schüler:innen in Pendeldistanz liegen mit ihrer Einschätzung zwischen den beiden Regionen). Wie auch schon in der Betrachtung der Prätestergebnisse in Abschnitt 9.5 zeigt sich folgendes Ergebnis: Im Pinzgau, der Region außerhalb der Pendeldistanz zum Zentralraum, wird dialektale Sprechweise durchgängig am negativsten eingeschätzt, in den Bezirken Tennengau & Flachgau, den Regionen in Pendeldistanz zum Zentralraum, am positivsten. Die Rater:innen, die in der Stadt Salzburg ins Gymnasium gehen, liegen mit ihren Werten dazwischen.<sup>36</sup> Bei der standardsprachlichen Sprechweise verhalten sich die Einschätzungen der Sprecher:innen eher hypothesenkonform zu H<sub>6</sub> (Schüler:innen in der Stadt schätzen standardsprechende Personen sympathischer ein als Schüler:innen am Land. Schüler:innen in Pendeldistanz liegen mit ihrer Einschätzung zwischen den beiden anderen Regionen). So werden standardsprechende Personen von den Rater:innen des Flach- & Tennengaus am positivsten eingeschätzt, dann folgen die Ergebnisse aus der Stadt Salzburg und am Schluss die aus dem Pinzgau.

Generell scheint es jedoch so zu sein, dass die Einstellungen der Rater:innen des Pinzgaus am stärksten von der Intervention verändert wurden, da hier die Werte am stärksten gestiegen sind. Auch der Rückgang der Tendenz im Flach- & Tennengau, die Standardsprache viel positiver einzuschätzen als die anderen Gruppen, könnte mit der Intervention zusammenhängen, da dieser Rückgang als reflektiertere Einschätzung dieser Varietät interpretiert werden kann.

## 11.6 Erstsprache

Dieser Abschnitt widmet sich potenziell unterschiedlichen Wirkungen der Intervention auf Personen mit Deutsch als Erstsprache, Deutsch und einer weiteren Erstsprache sowie einer anderen Erstsprache als Deutsch. Von den 113 Proband:innen dieser Gruppen, die beim Prätest anwesend waren, gab der größte Teil, 93 Personen, an, nur Deutsch als Erstsprache zu besitzen. Die beiden anderen Gruppen kommen insgesamt nur auf 20 Personen in den Prätests und nur auf 16 in den Posttests. Für Signifikanz-

---

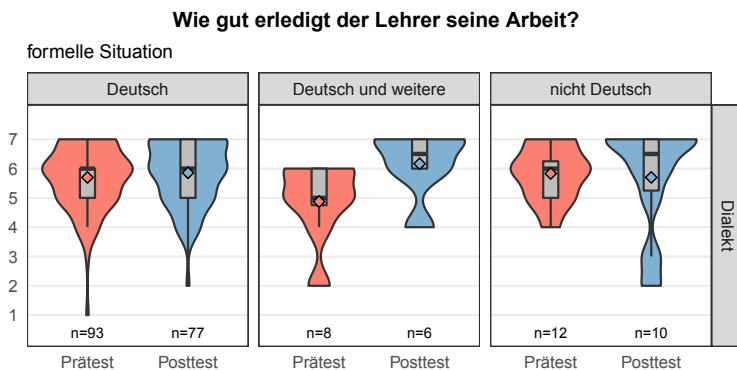
<sup>36</sup> Das Abstimmungsverhalten deckt sich in etwa mit der Nähe der dialektalen Varietät in den Aufnahmen zu den Rater:innen. So entspricht die Sprechweise am ehesten derjenigen des Flach- & Tennengaus, etwas weniger derjenigen der Stadt Salzburg und am wenigsten derjenigen des Pinzgaus.

tests ist die Anzahl der ratenden Personen zu niedrig, daher wurde in diesen beiden Gruppen darauf vollständig verzichtet. In der Gruppe der Deutsch-Erstsprachler:innen wurden signifikante Gruppenunterschiede, falls vorhanden, in den Plots integriert.

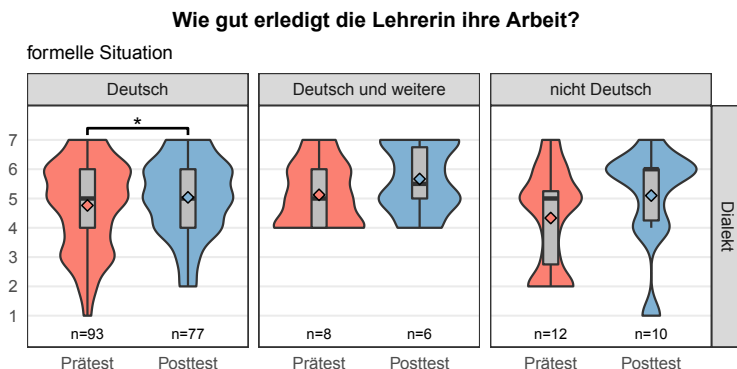
### 11.6.1 Detaillierte Ergebnisse

Auf Kompetenzebene in formeller Situation gibt es mehrere Ergebnisse zu berichten. Der dialektsprechende Lehrer (s. Abb. 11.79) wird vor allem von der Gruppe mit Deutsch und einer weiteren Erstsprache deutlich kompetenter nach erfolgter Intervention eingeschätzt. Diese Gruppe zeigt jedoch auch die niedrigsten Prätestwerte. Bei den beiden anderen Gruppen sind keine großen Veränderungen bemerkbar.

Die dialektsprechende Lehrerin wird nach der Intervention von allen drei Gruppen kompetenter eingeschätzt – bei der Gruppe mit Deutsch als Erstsprache sind die Unterschiede auch signifikant (s. Abb. 11.80). Die positiveren Einschätzungen bei den beiden anderen Gruppen sind, wie man an den deskriptiven Plots sehen kann, jedoch noch

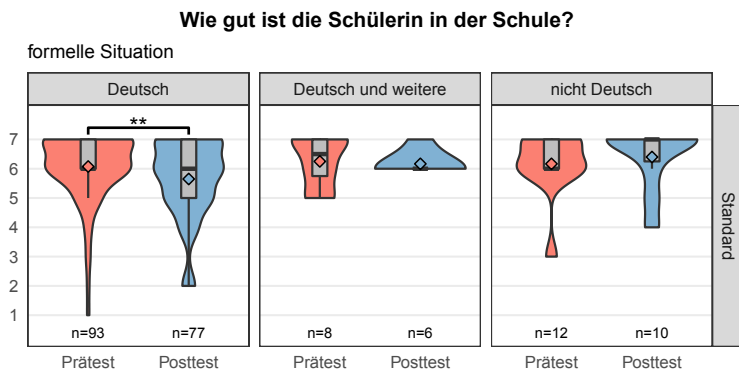


**Abbildung 11.79:** Prätest-Posttest-Vergleich der Erstsprachenunterschiede bei der Frage nach der Kompetenz des dialektsprechenden Lehrers in formeller Situation



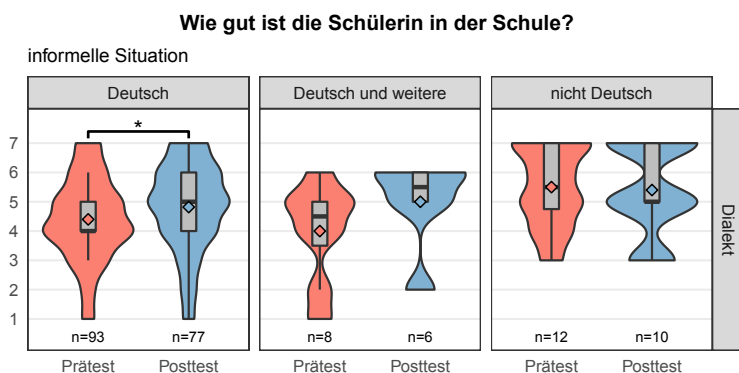
**Abbildung 11.80:** Prätest-Posttest-Vergleich der Erstsprachenunterschiede bei der Frage nach der Kompetenz der dialektsprechenden Lehrerin in formeller Situation

ausgeprägter, wobei die Gruppe mit einer anderen Erstsprache als Deutsch in den Posttestwerten sich den Deutsch-Erstsprachler:innen annähert und die Gruppe mit Deutsch und einer weiteren Erstsprache noch deutlich darüber liegt.



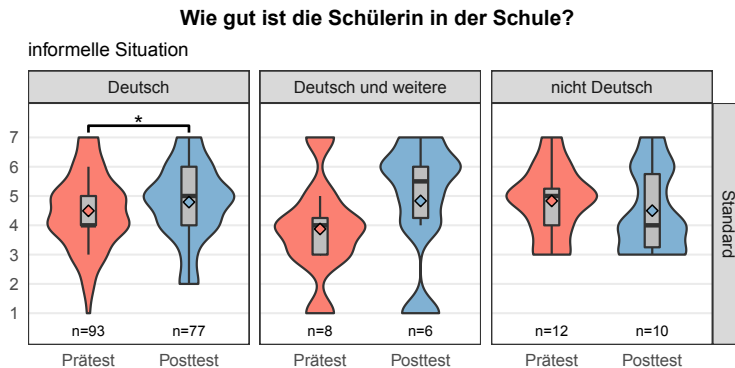
**Abbildung 11.81:** Prätest-Posttest-Vergleich der Erstsprachenunterschiede bei der Frage nach der Kompetenz der dialektspredenden Schülerin in formeller Situation

Eines der wenigen signifikanten Ergebnisse bei standardsprachlicher Sprechweise findet man bei der Schülerin in formeller Situation (s. Abb. 11.81). Sie wird im Prätest über alle Gruppen hinweg sehr kompetent eingeschätzt. Im Posttest wird jedoch in der Gruppe mit Deutsch als Erstsprache die Kompetenzeinschätzung hochsignifikant negativer. Sie bleibt jedoch weiterhin auf hohem Niveau, gleicht sich aber den Einschätzungen der dialektspredenden Schülerin an, deren arithmetisches Mittel ebenfalls zwischen fünf und sechs liegt. Bei den beiden anderen Gruppen sieht man dagegen keine großen Veränderungen.



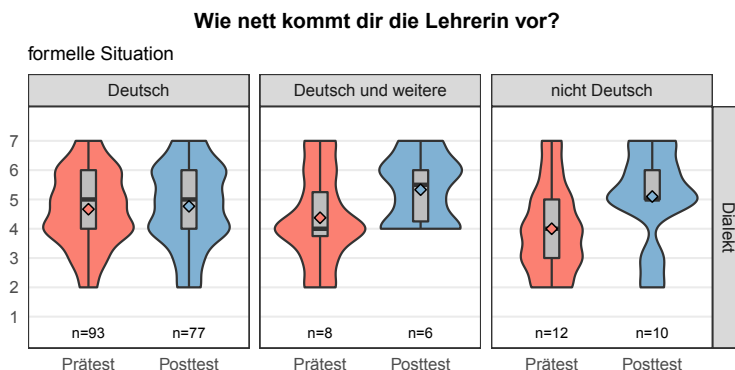
**Abbildung 11.82:** Prätest-Posttest-Vergleich der Erstsprachenunterschiede bei der Frage nach der Kompetenz der dialektspredenden Schülerin in informeller Situation

Auf Kompetenzebene in informeller Situation findet man ebenfalls bei der Schülerin Unterschiede, und zwar sowohl in dialektaler (Abb. 11.82) als auch standardsprachlicher Sprechweise (Abb. 11.83).



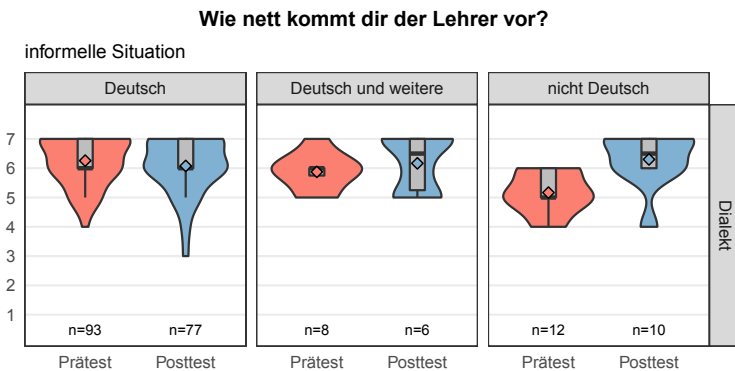
**Abbildung 11.83:** Prätest-Posttest-Vergleich der Erstsprachenunterschiede bei der Frage nach der Kompetenz der standardsprechenden Schülerin in informeller Situation

In diesen beiden Abbildungen fällt zum einen auf, dass die Einschätzungen generell viel niedriger ausfallen als in formeller Situation, was mit dem informellen Gesprächsinhalt der Guises zusammenhängen kann. Zum anderen wird die Sprecherin sowohl bei dialektaler als auch standardsprachlicher Sprechweise im Posttest von der Gruppe mit Deutsch als Erstsprache signifikant kompetenter eingeschätzt. Die Einschätzungen der beiden Varietäten selbst fallen sehr ähnlich aus. Bei der Gruppe mit Deutsch und einer weiteren Sprache sieht man ebenfalls bei beiden Varietätenverwendungen eine deutlich positivere Einschätzung im Posttest. In der Gruppe mit einer anderen Erstsprache als Deutsch sieht man einen leichten Abwärtstrend – etwas ausgeprägter bei standardsprachlicher Sprechweise.



**Abbildung 11.84:** Prätest-Posttest-Vergleich der Erstsprachenunterschiede bei der Frage nach der Sympathie der dialektprechenden Lehrerin in informeller Situation

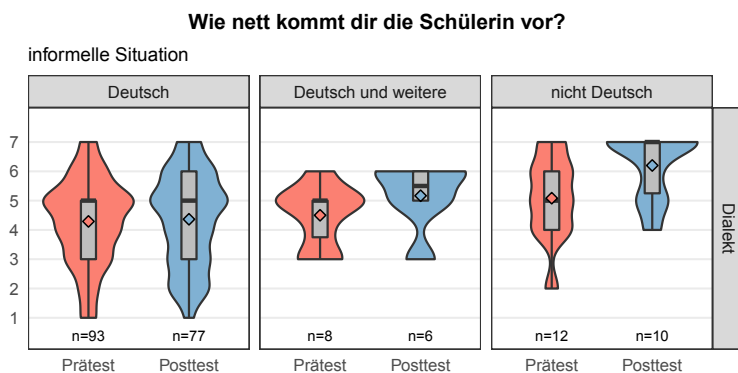
Auf Sympathieebene findet man Unterschiede nur in den dialektsprachlichen Guises vor. In formeller Situation (s. Abb. 11.84) wird die Lehrerin von den Gruppen mit einer anderen Erstsprache als Deutsch bzw. Deutsch und einer weiteren Erstsprache im Posttest deutlich sympathischer eingeschätzt. Dagegen sieht man keine großen Bewegungen bei den Rater:innen mit Deutsch als Erstsprache. Der dialektsprechende Lehrer in informeller Situation wird mit ähnlichem Muster eingeschätzt, seine Werte liegen jedoch deutlich über denen der Lehrerin (s. Abb. 11.85). Die Gruppe mit einer anderen Erstsprache als Deutsch kommt im Posttest zu deutlich positiveren Einschätzungen und mit ihren Werten über den beiden anderen Gruppen zu liegen. Auffällig ist auch, dass die gleiche Gruppe im Prätest die niedrigsten Werte zeigt, daher ist die Differenz zwischen den Testzeitpunkten vergleichsweise groß. Die beiden anderen Gruppen zeigen zwischen beiden Testzeitpunkten sehr hohe Werte, bei der Gruppe mit Deutsch und einer weiteren Erstsprache nehmen sie zum Posttest hin sogar noch etwas zu.



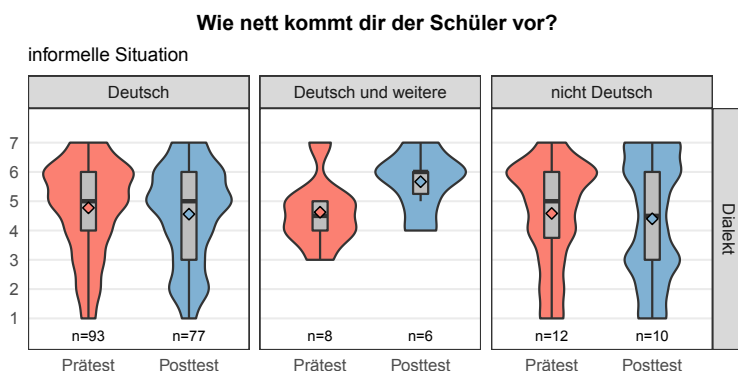
**Abbildung 11.85:** Prätest-Posttest-Vergleich der Erstsprachenunterschiede bei der Frage nach der Sympathie des dialektsprechenden Lehrers in informeller Situation

Bei der dialektsprechenden Schülerin in informeller Situation (s. Abb. 11.86) sieht man in den Gruppen „Deutsch und weitere“ und „nicht Deutsch“ eine deutliche Tendenz zur positiveren Einschätzung nach der Intervention. Bei den Personen mit Deutsch als Erstsprache ist dagegen keine größere Veränderung auszumachen. Interessant ist auch, dass letztere Gruppe in beiden Testzeitpunkten die niedrigsten Werte angibt. Die dialektsprechende Schülerin in informeller Situation wird also von den Deutsch-Erstsprachler:innen am wenigsten sympathisch eingeschätzt.

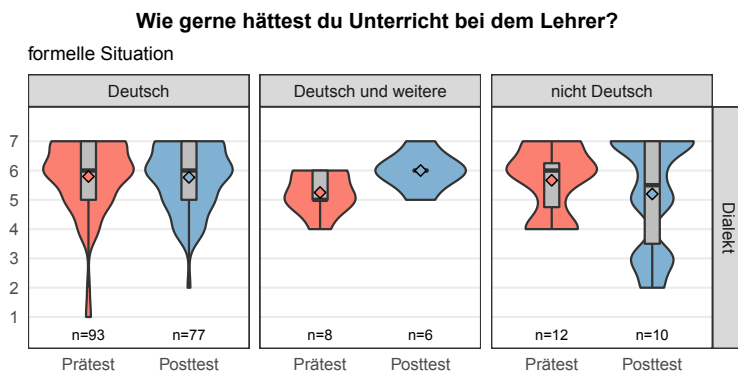
Beim dialektsprechenden Schüler in informeller Situation (Abb. 11.87) stellen sich die Ergebnisse etwas differenzierter dar. Die Gruppen „Deutsch“ und „nicht Deutsch“ schätzen nach der Intervention den Schüler etwas weniger sympathisch ein, die Gruppe mit Deutsch und einer weiteren Erstsprache geht hier den gegenteiligen Weg – der dialektsprechende Schüler wird deutlich sympathischer eingeschätzt und diese Einschätzung liegt auch deutlich über denjenigen der beiden anderen Gruppen.



**Abbildung 11.86:** Prätest-Posttest-Vergleich der Erstsprachenunterschiede bei der Frage nach der Sympathie der dialektprechenden Schülerin in informeller Situation



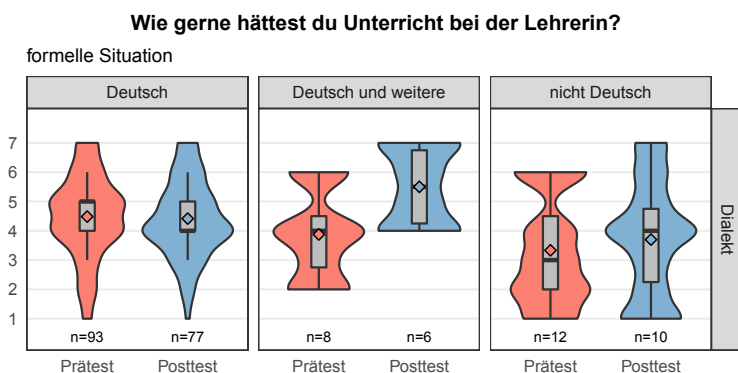
**Abbildung 11.87:** Prätest-Posttest-Vergleich der Erstsprachenunterschiede bei der Frage nach der Sympathie des dialektprechenden Schülers in informeller Situation



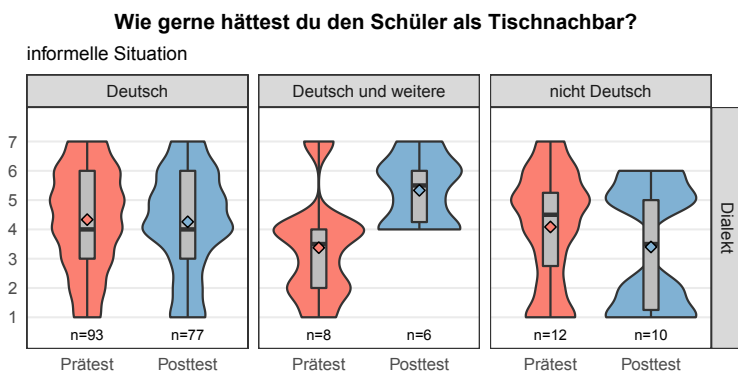
**Abbildung 11.88:** Prätest-Posttest-Vergleich der Erstsprachenunterschiede bei der Frage nach persönlicher Präferenz zum dialektprechenden Lehrer in formeller Situation

Auf Ebene der persönlichen Präferenz zeigen sich ebenfalls nur in den dialektsprachlichen Guises signifikante Unterschiede. In formeller Situation wird der dialektsprechende Lehrer von der Gruppe „Deutsch und weitere“ im Posttest deutlich positiver eingeschätzt (s. Abb. 11.88). In der Gruppe mit einer anderen Erstsprache als Deutsch gehen die Einschätzungen hingegen zurück – jedoch von hohem Ausgangsniveau. In der Gruppe mit Deutsch als Erstsprache sind dagegen keine Veränderungen bemerkbar. Beide Einschätzungen zeigen hohe Werte.

Bei der dialektsprechenden Lehrerin in formeller Situation (Abb. 11.89) zeigt sich, dass die Gruppen „Deutsch und weitere“ sehr deutlich und „nicht Deutsch“ etwas weniger deutlich die Lehrerin nach der Intervention positiver einschätzen. Die Posttestwerte aller drei Gruppen sind auffällig unterschiedlich. In der Gruppe „Deutsch“ liegt das arithmetische Mittel bei ca. 4,5; in der Gruppe „Deutsch und weitere“ liegt es bei ca. 5,5 und in der Gruppe „nicht Deutsch“ bei unter vier.



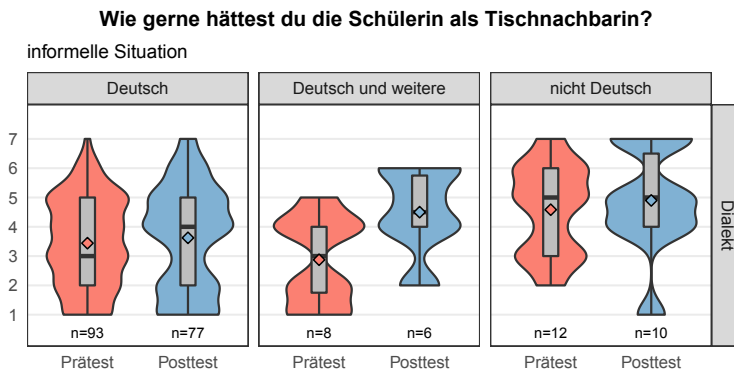
**Abbildung 11.89:** Prätest-Posttest-Vergleich der Erstsprachenunterschiede bei der Frage nach persönlicher Präferenz zum dialektsprechenden Lehrer in formeller Situation



**Abbildung 11.90:** Prätest-Posttest-Vergleich der Erstsprachenunterschiede bei der Frage nach persönlicher Präferenz zum dialektsprechenden Schüler in informeller Situation

In informeller Situation stechen die dialektsprachlichen Schüler:innen-Guises heraus. Der Schüler (s. Abb. 11.90) wird von der Gruppe mit Deutsch und einer weiteren Erstsprache in der Posttestung deutlich positiver eingeschätzt, von der Gruppe mit einer anderen Erstsprache als Deutsch aber negativer. Die Gruppe mit Deutsch als Erstsprache zeigt hingegen keine Veränderung auf.

Die Schülerin (s. Abb. 11.91) wird ebenfalls von der Gruppe „Deutsch und weitere“ deutlich positiver eingeschätzt. Auch in den beiden anderen Gruppen sieht man Zuwächse in den Werten, wenn auch nur in geringem Maße. Vergleicht man nun die Posttest-Werte zwischen den Gruppen, sieht man, dass die dialektsprechende Schülerin von den Deutsch-Erstsprachler:innen am negativsten eingeschätzt wird. Beim Schüler ist dies anders. Hier ist es die Gruppe mit einer anderen Erstsprache als Deutsch, die ihn am negativsten einschätzt – ein auffälliger Unterschied, der so auch schon auf Sympathieebene vorlag.

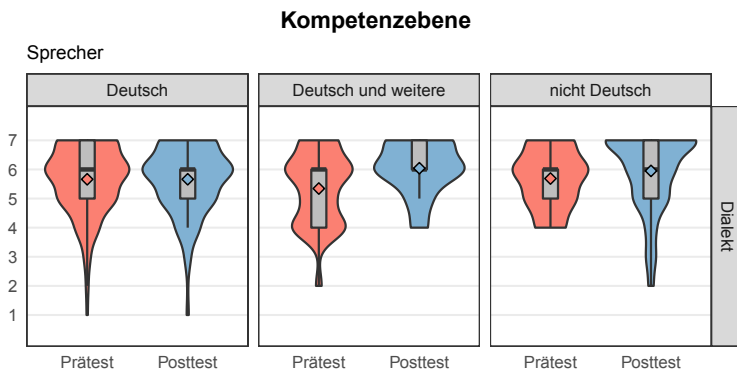


**Abbildung 11.91:** Prätest-Posttest-Vergleich der Erstsprachenunterschiede bei der Frage nach persönlicher Präferenz zum dialektsprechenden Schüler in informeller Situation

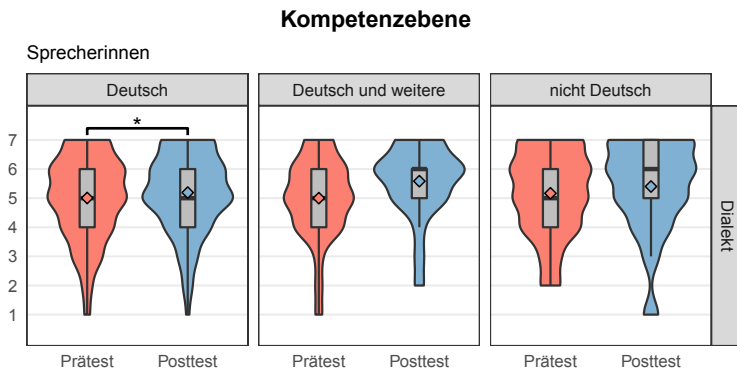
### 11.6.2 Generelle Ergebnisse

Wenn man die formellen und informellen Situationen sowie die Rollen Schüler:innen und Lehrkräfte in der Analyse zusammennimmt, gibt es keine signifikanten und sichtbaren Unterschiede bei standardsprachlicher Varietätenverwendung. Bei dialektaler Sprechweise sind vor allem die geschlechterspezifischen Betrachtungen relevant, da sie sich teils deutlich voneinander unterscheiden.

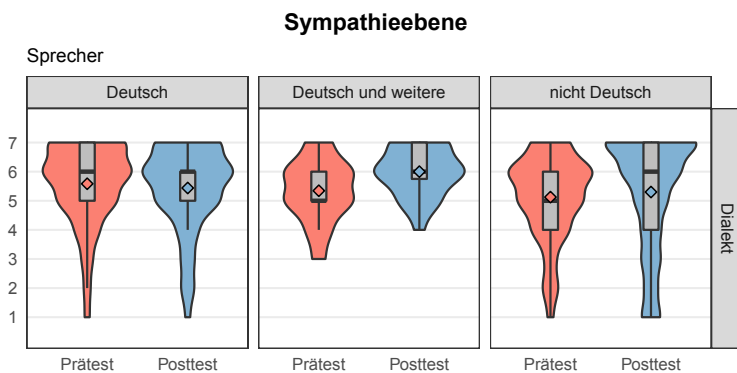
Auf Kompetenzebene werden den Sprechern verhältnismäßig hohe Werte von allen drei Gruppen zugesprochen (s. Abb. 11.92). Dabei sieht man so gut wie keine Prä-Posttest-Veränderung bei der Gruppe der Deutsch-Erstsprachler:innen. Die Gruppe „Deutsch und weitere“ zeigt im Prätest im Vergleich etwas niedrigere Werte, die zum Posttest hin jedoch deutlich ansteigen. Auch bei der Gruppe „nicht Deutsch“ zeigt sich es eine Tendenz zur höheren Kompetenzeinschätzung der Sprecher – bei beiden letzteren Gruppen scheint es im Gegensatz zu den Deutsch-Erstsprachler:innen Interventionseffekte zu geben.



**Abbildung 11.92:** Allgemeiner Prätest-Posttest-Vergleich der Erstsprachenunterschiede bei der Frage nach Kompetenz bezüglich Dialektsprechern



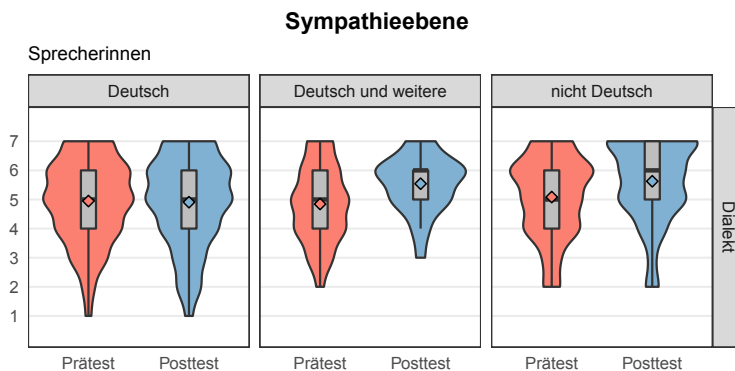
**Abbildung 11.93:** Allgemeiner Prätest-Posttest-Vergleich der Erstsprachenunterschiede bei der Frage nach der Kompetenz bezüglich Dialektsprecherinnen



**Abbildung 11.94:** Allgemeiner Prätest-Posttest-Vergleich der Erstsprachenunterschiede bei der Frage nach der Sympathie bezüglich Dialektsprechern

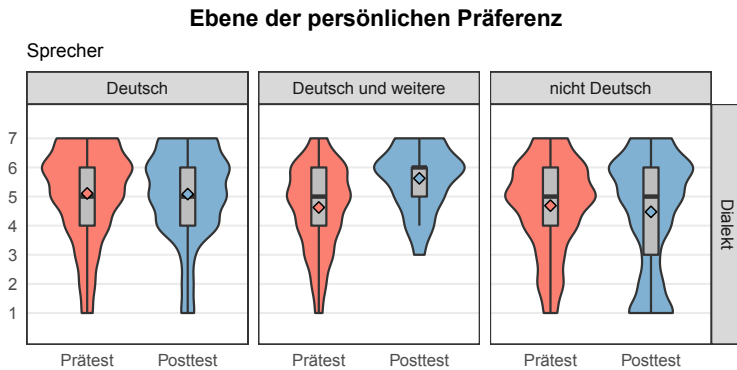
Bei den Sprecherinnen (s. Abb. 11.93) liegen die Ratingergebnisse deutlich unter denjenigen der Sprecher. Allerdings zeigt sich im Gegensatz zu diesen, dass in allen drei Gruppen Tendenzen zur positiveren Einschätzung vorliegen – bei der Gruppe mit Deutsch als Erstsprache sind diese auch signifikant. Bei den beiden anderen Gruppen treten ähnliche Muster zum Vorschein wie schon bei den Sprechern. Eine deutliche Zunahme der Kompetenzeinschätzung ist bei „Deutsch und weitere“ zu verzeichnen, aber auch bei der Gruppe „nicht Deutsch“ ist der Anstieg gut auszumachen.

Auf Sympathieebene und in dialektaler Sprechweise sind die Ergebnisse beim Sprecher ähnlich wie auf Kompetenzebene (s. Abb. 11.94). Bei der Gruppe der Deutsch-Erstsprachler:innen sieht man nur wenig Veränderung zwischen den Testsituationen, es ist sogar eine leichte Abwärtsbewegung zum Posttest hin zu bemerken. Im Gegensatz dazu stehen die beiden anderen Gruppen. Die Gruppe „Deutsch und weitere“ zeigt eine deutliche Tendenz zu größerer Sympathieeinschätzung der Sprecher, wobei die Werte der Prätestung etwas niedriger ausfallen als bei der Gruppe mit Deutsch als Erstsprache. Die Gruppe „nicht Deutsch“ liegt mit ihren Prätestergebnissen sogar am niedrigsten, es ist aber auch hier eine Zunahme feststellbar.



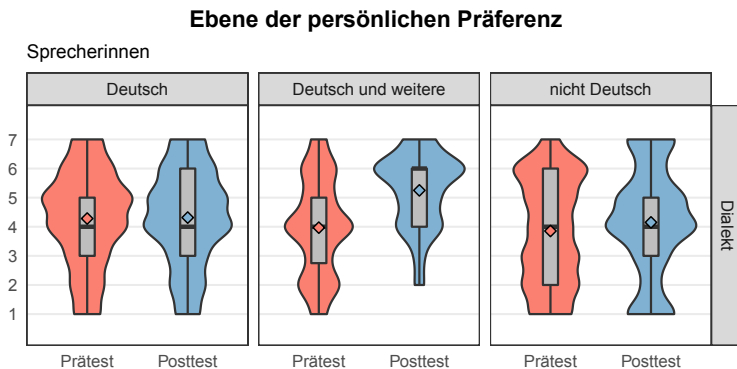
**Abbildung 11.95:** Allgemeiner Prätest-Posttest-Vergleich der Erstsprachenunterschiede bei der Frage nach der Sympathie bezüglich Dialektsprecherinnen

Die dialektsprechenden Sprecherinnen (Abb. 11.95) werden auch auf Sympathieebene im Vergleich zu den Sprechern niedriger eingeschätzt. Bei der Gruppe der Deutsch-Erstsprachler:innen ändert sich an dieser Einschätzung praktisch nichts. Bei den beiden andern Gruppen sieht man jedoch einen deutlichen Anstieg, der etwa im selben Ausmaß vorliegt, wobei die Prätestergebnisse sehr denjenigen der Deutsch-Erstsprachler:innen ähneln.



**Abbildung 11.96:** Allgemeiner Prätest-Posttest-Vergleich der Erstsprachenunterschiede bei der Frage nach persönlicher Präferenz bezüglich Dialektsprechern

Auf Ebene der persönlichen Präferenz findet man bei den dialektsprechenden Sprechern nun einen auffälligen Unterschied zu den Ergebnissen der Sympathieebene (s. Abb. 11.96). Verhalten sich die beiden Gruppen „Deutsch“ und „Deutsch und weitere“ noch sehr ähnlich in ihrem Abstimmungsverhalten in Prä- und Posttest (wenn auch die Abstimmungswerte deutlich unter denjenigen der Sympathieebene liegen) zeigen die Rater:innen der Gruppe „nicht Deutsch“ eine völlig andere Tendenz – im Posttest schätzen sie die dialektsprechenden Sprecher weniger positiv ein als im Prätest. Damit divergiert die Einschätzung dieser Gruppe deutlich von derjenigen der Gruppe mit Deutsch und einer weiteren Erstsprache.



**Abbildung 11.97:** Allgemeiner Prätest-Posttest-Vergleich der Erstsprachenunterschiede bei der Frage nach persönlicher Präferenz bezüglich Dialektsprecherinnen

Interessant ist dies auch, wenn man dieses Ergebnis mit demjenigen der Sprecherinnen vergleicht (s. Abb. 11.97). Zum einen werden die Sprecherinnen über alle Gruppen hinweg im Prätest deutlich negativer eingeschätzt als die Sprecher. In der Gruppe „Deutsch“ sieht man keine Veränderung nach der Intervention. Die Gruppe „Deutsch

und weitere“ zeigt deutlich positivere Einschätzungen und erreicht im Posttest fast das Niveau der Sprecher. Die Gruppe „nicht Deutsch“ zeigt aber ebenfalls positivere Einschätzungen, wenn auch nicht so ausgeprägt.

### 11.6.3 Zusammenfassung Erstsprache

Auch wenn die Ergebnisse der Betrachtung unterschiedlicher Erstsprachen wegen der geringen Personenanzahl der Gruppen mit Deutsch und einer weiteren Erstsprache sowie einer anderen Erstsprache als Deutsch keine inferenzstatistischen Berechnungen zulassen, zeigen sich dennoch bemerkenswerte Muster. Zum einen sind Veränderungen vor allem bei den dialektsprachlichen Guises bemerkbar. In standardsprachlicher Sprechweise zeigen sich nur zwei signifikante Ergebnisse auf Kompetenzebene bei der Schülerin. In formeller Situation wird sie von den Deutsch-Erstsprachler:innen von sehr hohem Niveau ausgehend etwas weniger kompetent eingeschätzt, in informeller Situation etwas kompetenter. Im Vergleich sieht man hier am ehesten, dass es zu einer Nivellierung der beiden Situationen gekommen ist, die sich ja inhaltlich voneinander unterscheiden. Die hochsignifikante Abnahme der Kompetenzeinschätzung auf formeller Ebene könnte man dahingehend deuten, dass von den Schüler:innen mit Deutsch als Erstsprache nach der Intervention diese Abstimmung etwas reflektierter vonstatten gegangen ist – eine Tendenz, die in den beiden anderen Gruppen nicht bemerkbar ist.

Ansonsten zeigen vor allem die dialektalen Guises auffallende Veränderungen. Signifikante Unterschiede in der Gruppe der Deutsch-Erstsprachler:innen findet man ausschließlich auf Kompetenzebene bei den Lehrerinnen und Schülerinnen – diese wurden nach der Intervention signifikant kompetenter eingeschätzt. Auch die Gruppe mit Deutsch und einer weiteren Erstsprache zeigt deutlich positivere Kompetenzeinschätzungen der Sprecherinnen, bei der Gruppe mit einer anderen Erstsprache als Deutsch sind die Ergebnisse nicht so deutlich: Bei der dialektsprechenden Lehrerin und Schülerin in formeller Situation nimmt die Kompetenzeinschätzung zu, bei der dialektsprechenden Schülerin in informeller Situation nimmt sie ab.

Auf Sympathieebene sieht man generelle Tendenzen, die dialektsprechenden Sprecher:innen sympathischer einzuschätzen. Bei der Gruppe mit Deutsch als Erstsprache sind keine Prä-Posttest-Veränderungen erkennbar, ganz im Gegensatz zu den beiden anderen Gruppen. Diese schätzen sowohl in formeller als auch informeller Situation die Schüler:innen wie auch die Lehrkräfte sympathischer im Posttest ein. Die einzige Ausnahme bildet hier das Abstimmungsverhalten beim dialektsprechenden Schüler in informeller Situation. Dieser wird von der Gruppe mit einer anderen Erstsprache als Deutsch im Posttest weniger sympathisch eingeschätzt.

Auf Ebene der persönlichen Präferenz sieht man wiederum bei den Deutsch-Erstsprachler:innen die geringsten Veränderungen. Einzig hervorzuheben ist das positivere Abstimmungsergebnis bei der dialektsprechenden Schülerin in informeller Situation, die jedoch ein sehr niedriges Prätestniveau zeigt. Die Gruppe „Deutsch und weitere“ zeigt wiederum eine Tendenz zur positiveren Einschätzung nach der Intervention. Die

Gruppe „nicht Deutsch“ zeigt eine Tendenz zur positiveren Einschätzung bei den Schülerinnen und zur Schlechterbewertung bei den Schülern.

Die generellen Ergebnisse unterstreichen die Detailliergebenisse. Die dialektsprechenden Sprecherinnen werden von allen Gruppen nach erfolgter Intervention kompetenter eingeschätzt, die Sprecher nur von den Gruppen „Deutsch und weitere“ und „nicht Deutsch“. Diese beiden Gruppen schätzen auch beide Geschlechter sympathischer ein. Auf Ebene der persönlichen Präferenz werden beide Geschlechter von der Gruppe „Deutsch und weitere“ positiver eingeschätzt, von der Gruppe „nicht Deutsch“ allerdings nur die Sprecherinnen. Die Sprecher streichen hier niedrigere Werte ein.

Auch wenn es die Proband:innenzahlen nicht zulassen, liegt es meiner Einschätzung nach nahe, dass die Intervention etwas unterschiedliche Auswirkungen auf die drei Gruppen zeigt. Auf die Rater:innen mit Deutsch als Erstsprache hat sie vor allem die Kompetenzeinschätzung dialektsprechender Personen – und hier vor allem die der Sprecherinnen – steigen lassen. Für die beiden anderen Gruppen scheint es so zu sein, dass sich durch die Intervention neben der höheren Kompetenzeinschätzung von Personen, die dialektal sprechen, auch die Sympathieeinschätzung und die der persönlichen Präferenz verändert haben. So werden alle Personen, die dialektal sprechen, auf diesen beiden Ebenen positiver eingeschätzt, mit Ausnahme der Sprecher, die auf Ebene der persönlichen Präferenz von der Gruppe mit einer anderen Erstsprache als Deutsch etwas negativer eingeschätzt wurden.

## 11.7 Bildungshintergrund

Auch bei der Untersuchung des Bildungshintergrunds trifft man auf Probleme der Stichprobengröße. Da nur die Interventionsgruppen analysiert werden, reduziert sich die Anzahl der Rater:innen um ca. die Hälfte. Zusätzlich führt die Teilung in vier Gruppen dazu, dass in zwei Gruppen die wünschenswerte Anzahl von zumindest 25 Personen unterschritten wird. Dies betrifft vor allem die Gruppe „0–25“, in der nur 17 Personen in der Prätestung und nur 14 in der Posttestung abgestimmt haben.<sup>37</sup> Auch in der Gruppe „100–200“ waren in der Posttestung nur 21 Rater:innen anwesend. Es wurden dennoch Signifikanztestungen in Form von gemischten linearen Modellen mit den Teilnehmer:innen als *Random Intercept* gerechnet und signifikante Ergebnisse in den deskriptiven Plots eingefügt.

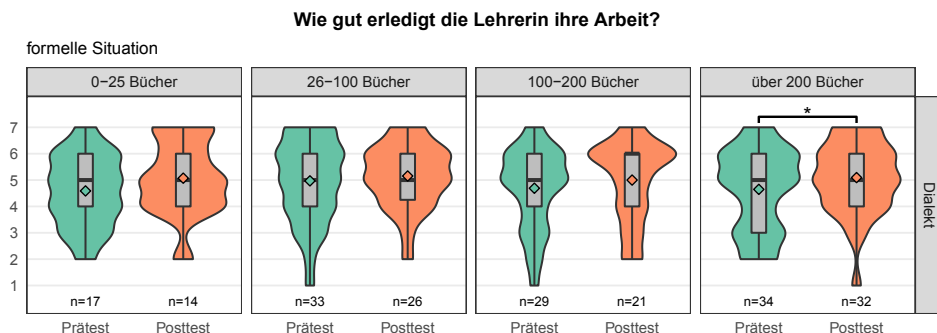
### 11.7.1 Detaillierte Ergebnisse

Bei der explorativen Betrachtung über alle Kategorien fällt auf, dass nur in wenigen Teilbereichen signifikante Unterschiede ausgemacht werden können – alle betreffen die Kompetenzebene und die Sprecherinnen, die Ebene also, die bisher schon die

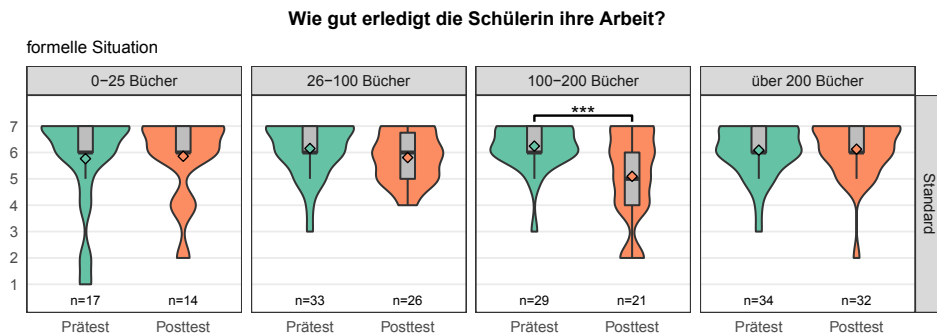
37 Zur Erinnerung: Im Fragebogen gab es eine Frage, mit der der Bildungshintergrund indirekt über die Anzahl der Bücher zu Hause festgestellt wurde. Dabei gab es fünf Antwortmöglichkeiten. Die beiden ersten „keine oder nur sehr wenige (0–10 Bücher)“ und „genug, um ein Regalbrett zu füllen (11–25 Bücher)“ gaben insgesamt sehr wenige Personen an. Daher wurden sie in der Analyse zu einer Kategorie „0–25 Bücher“ zusammengeführt.

meisten Unterschiede aufwies. Auf Sympathie- wie Ebene der persönlichen Präferenz und bei den Sprechern scheinen die Gruppen mit unterschiedlichem Bildungshintergrund ähnliche Einschätzungen zu vertreten.

In Abbildung 11.98 sind die Ergebnisse für die dialekt sprechende Lehrerin in formeller Situation, aufgeteilt in die Bildungskategorien, abgebildet. Man erkennt, dass das arithmetische Mittel in allen Gruppen zunimmt und sich die Posttestwerte aller Gruppen auf recht ähnlichen Werten befinden. Die Zunahme in der Kompetenzeinschätzung ist jedoch nur bei der Gruppe „über 200“ signifikant – man sieht hier an der Varianz deutlich, dass zum Posttest hin viel weniger Einschätzungen im negativen Bereich vorliegen.



**Abbildung 11.98:** Prätest-Posttest-Vergleich des Bildungshintergrunds bei der Frage nach der Kompetenz der dialekt sprechenden Lehrerin in formeller Situation

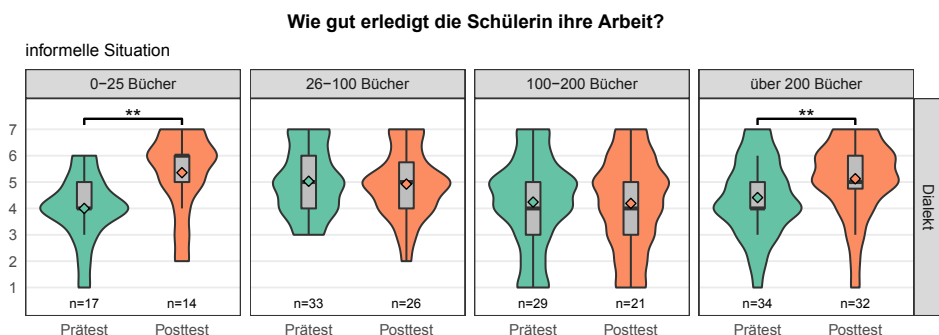


**Abbildung 11.99:** Prätest-Posttest-Vergleich des Bildungshintergrunds bei der Frage nach der Kompetenz der standardsprechenden Schülerin in formeller Situation

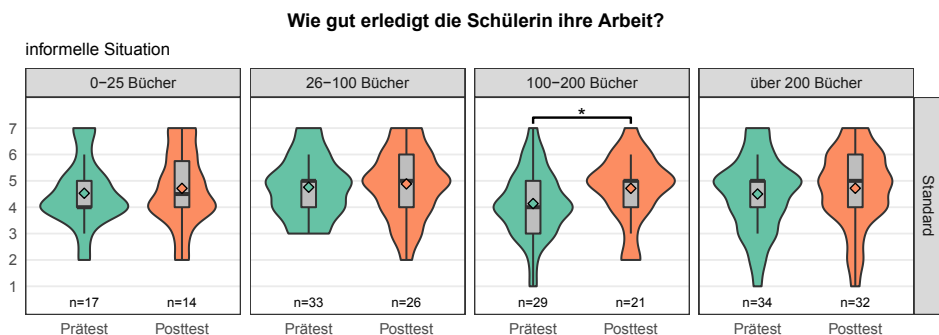
In formeller Situation findet man bei der standardsprechenden Schülerin folgendes Ergebnis, welches in Abbildung 11.99 abgebildet ist: Die Einschätzungen dieses Guises im Prätest fallen für alle Gruppen recht hoch aus. Interessant ist nun, dass in den Gruppen „26–100“ und „100–200“ ein Rückgang der Einschätzungen im Posttest zu bemer-

ken ist. Bei letzterer Gruppe ist dieser Rückgang sehr ausgeprägt und daher sind die inferenzstatistischen Messergebnisse auch höchstsignifikant.

In informeller Situation zeigen sich deutliche Veränderungen bei der dialekt-sprechenden Schülerin (s. Abb. 11.100). In diesem Fall ist zu erkennen, dass es bei den Gruppen mit dem niedrigsten und dem höchsten Bildungshintergrund zu einer deutlich positiveren Kompetenzeinschätzung im Posttest kommt. In beiden Fällen zeigt der Unterschied hochsignifikante Messwerte, wobei die Ergebnisse der Gruppe „0–25“ wegen der zu geringen Zahl der Proband:innen mit Vorsicht zu genießen sind. Zwei Gründe sprechen jedoch dafür, dass die Signifikanzwerte bei etwaiger Replikation der Untersuchung standhalten würden:



**Abbildung 11.100:** Prätest-Posttest-Vergleich des Bildungshintergrunds bei der Frage nach der Kompetenz der dialekt-sprechenden Schülerin in informeller Situation



**Abbildung 11.101:** Prätest-Posttest-Vergleich des Bildungshintergrunds bei der Frage nach der Kompetenz der standardsprechenden Schülerin in informeller Situation

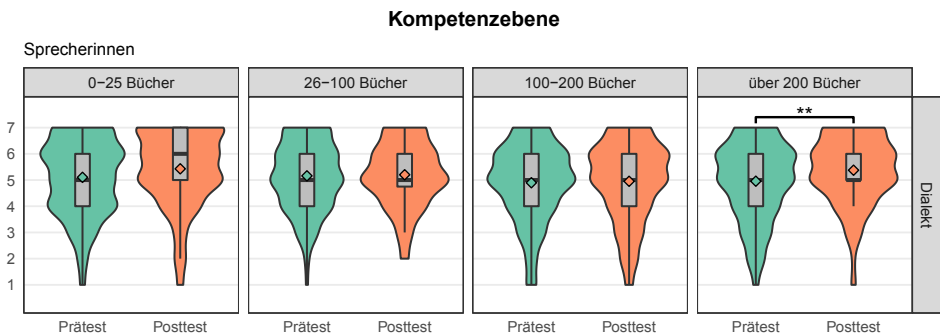
Zum einen ist die Zunahme der Kompetenzeinschätzung, wie man deskriptiv beobachten kann, sehr ausgeprägt – noch stärker als bei der Gruppe mit dem höchsten Bildungshintergrund. Zum anderen sehen speziell die Prätestergebnisse dieser Gruppe normalverteilt aus, auch die Posttestergebnisse kommen einer Normalverteilung nahe, sind jedoch etwas linksverschoben, was eindeutig mit der positiveren Einschätzung in Zu-

sammenhang steht. Bei Normalverteilung der Werte ist die Gruppengröße nicht mehr ganz so entscheidend. Meine Einschätzung ist daher, dass diese Werte Aussagekraft besitzen.

In informeller Situation sind darüber hinaus Veränderungen bei der standardsprechenden Schülerin bemerkbar (Abb. 11.101). Man erkennt in allen Gruppen eine Tendenz zur höheren Kompetenzeinschätzung im Posttest – vor allem jedoch bei der Gruppe „100–200“, bei der die Veränderung signifikant ist. Dies kommt jedoch daher, dass die standardsprechende Schülerin von dieser Gruppe im Prätest deutlich weniger kompetent eingeschätzt wurde, sich die Werte im Posttest jedoch denjenigen der anderen Gruppen annähern. Speziell die Werte des arithmetischen Mittels liegen bei den Gruppen eng beieinander, nur die Gruppe mit dem niedrigsten Bildungshintergrund liegt etwas darunter.

### 11.7.2 Generelle Ergebnisse

Die generellen Ergebnisse bleiben auf die Kompetenzebene und die dialektssprechenden Sprecherinnen beschränkt. Wie man in Abbildung 11.102 erkennen kann, schätzen vor allem die beiden Gruppen mit dem niedrigsten und höchsten Bildungshintergrund die Sprecherinnen in der Posttestung kompetenter ein, wobei nur letztere Gruppe hochsignifikante Unterschiede aufweist.



**Abbildung 11.102:** Allgemeiner Prätest-Posttest-Vergleich des Bildungshintergrunds bei der Frage nach der Kompetenz bezüglich Dialektsprecherinnen

Bei den beiden anderen Gruppen sind nur minimale Bewegungen in Richtung einer positiveren Einschätzung vorhanden. Ein weiteres interessantes Ergebnis ist, dass die Gruppen mit dem niedrigsten und höchsten Bildungshintergrund im Posttest auch deutlich höhere Werte zeigen als die beiden anderen Gruppen. Speziell der Vergleich der Gruppen „100–200“ und „über 200“ zeigt, dass beide in den Prätestungen sehr ähnliche Einschätzungen aufweisen, aber nur letztere Gruppe einen Anstieg zum Posttest hin zu verzeichnen hat.

### 11.7.3 Zusammenfassung Bildungshintergrund

Unterschiedlicher Bildungshintergrund führt speziell auf Kompetenzebene und bei den Sprecherinnen zu unterschiedlichen Tendenzen in der Einschätzung zwischen der Prä- und der Posttestung. Den dialektalen Sprachgebrauch der Sprecherinnen schätzen vor allem die Gruppen mit dem niedrigsten und höchsten Bildungshintergrund nach der Intervention kompetenter ein als zuvor. Ihre Ergebnisse liegen auch zumeist über den beiden anderen Gruppen.

In standardsprachlicher Sprechweise ist vor allem das Ergebnis der Gruppe „100–200“ in formeller Situation bei der Schülerin auffällig. Diese Gruppe schätzt von einem sehr hohen Prätestergebnis kommend die Schülerin in der Posttestung höchstsignifikant weniger kompetent ein. Der Grund hierfür könnte in einer reflektierteren Einschätzung liegen, die man so auch schon in anderen Abschnitten sehen konnte. Wenn man das Posttestergebnis mit demjenigen der informellen Ebene der standardsprechenden Schülerin vergleicht, sieht man, dass sich die Einschätzungen aneinander annähern. Ebenfalls nähert sich das Ergebnis demjenigen der dialektsprachlichen Guises der Schülerinnen an. Dies könnte bedeuten, dass sich die Schüler:innen dieser Gruppe weniger von den stereotypen Einschätzungen und den Satzinhalten in ihren Einschätzungen leiten lassen. Zusätzlich kann man davon ausgehen, dass dialektsprechende Personen nach der Intervention vor allem von Rater:innen mit hohem Bildungshintergrund kompetenter eingeschätzt werden.

## 12 Diskussion der Spracheinstellungsveränderung

Wie schon Siegel (2006), Bucholtz et al. (2014) oder Johnson et al. (2017) zeigten, scheinen Ansätze, die einer Critical Language Awareness folgen, Spracheinstellungen im Unterricht der Erstsprache verändern zu können. Die vorliegende Studie belegt dies und zeigt, dass dies auch im deutschsprachigen Raum der Fall ist. Da jedoch empirische Forschung zur Veränderung von Spracheinstellungen im Unterricht der Erstsprache nicht existiert, ist es auch nicht möglich, die Ergebnisse mit anderen Studien zu vergleichen. Durch die Analyse der Wirksamkeit der Intervention in bestimmten Teilbereichen ist es jedoch möglich, einen näheren Einblick in die Wirkung der Critical Language Awareness zu geben sowie diese Ansätze empirisch untermauern zu können.

### Kompetenzebene

Auf Kompetenzebene findet man die größten Veränderungen vor. So profitieren speziell die dialektsprechenden Sprecherinnen von der Intervention, die Gruppe also, die auch am negativsten im Prätest eingeschätzt wurde. Aber auch die dialektsprechenden Sprecher sowie teils die standardsprechenden Personen werden positiver eingeschätzt, jedoch in geringerem Ausmaß.

Diese Veränderungen werden zusätzlich durch den Rückgang der Distanz zwischen Dialekt und Standardsprache bestätigt. Hier sieht man, dass ebenfalls bei den Sprecherinnen ein deutlicher Rückgang dieser Distanz im Vergleich zu den Sprechern zu verzeichnen ist, wobei er bei Letzteren ebenfalls nachzuweisen ist. Zusätzlich sieht man einen leichten Rückgang der Bewertungsdistanz zwischen Sprecherinnen und Sprechern, allerdings sind die Ergebnisse des Geschlechtervergleichs weniger deutlich als die der zuvor genannten Bereiche.

Interessant ist auch, dass die Intervention speziell die Einstellungen zum eigenen Geschlecht verändert zu haben scheint. So haben wir in Abschnitt 9.3 gesehen, dass dialektsprechende Sprecher von den Raterinnen positiver eingeschätzt werden, als die Rater ihr eigenes Geschlecht einschätzen, ein Umstand, der so umgekehrt nicht feststellbar war. Ausgangspunkt hierfür könnte im *Confirmation Bias* (Klayman 1995) begründet sein, der dafür sorgen könnte, dass Sprecherinnen ihr eigenes Geschlecht negativer einschätzen.<sup>38</sup> Nach der Intervention schätzen die Raterinnen nun die Sprecherinnen aber hochsignifikant kompetenter ein und nähern sich damit ihren Einschätzungen der Sprecher an. Bei den Raterinnen sieht man also einerseits, dass sich die stereotypen Spracheinstellungen abgebaut haben, andererseits scheint auch ein eventu-

---

<sup>38</sup> Zur Erinnerung: Beim Confirmation Bias handelt es sich um den Effekt, dass Informationen so verarbeitet und gespeichert werden, dass sie den eigenen Erwartungen entsprechen. Implizit könnte das Ergebnis also darauf hindeuten, dass Schülerinnen heutzutage immer noch davon ausgehen, dass sie weniger kompetent sind als die Schüler.

eller *Confirmation Bias* zurückgegangen zu sein. Dies zeigt, dass die Intervention vor allem die Spracheinstellungen zu Sprecherinnen und von Raterinnen positiv beeinflusst hat und zusätzlich deren kognitive Repräsentation von weiblicher Kompetenz.

Bei den Ratern verhält es sich ähnlich. Sie schätzen die Sprecher ebenfalls etwas kompetenter ein, mit Werten nahe der Signifikanzgrenze, allerdings findet man hier keine Besserbewertung des anderen Geschlechts wie bei den Raterinnen. Trotzdem zeigt dies, dass auch bei den Ratern eine Veränderung der Spracheinstellungen stattgefunden hat und auch die stereotype Meinung, dass dialektsprechende Personen weniger kompetent seien, abgebaut werden konnte.

Zusätzlich findet sich ein Unterschied in standardsprachlicher Sprachverwendung. Die Schülerin wurde im Prätest auf sehr hohem Niveau eingeschätzt, eine Einschätzung, die nach der Intervention bei beiden Geschlechtern deutlich zurückgeht (bei den Raterinnen jedoch noch deutlicher). Dies könnte darauf hindeuten, dass in diesem Fall ebenfalls die stereotype Meinung, dass standardsprechende Personen kompetenter seien, abgebaut werden konnte. Dass diese Veränderung bei den Raterinnen noch stärker ausfiel, könnte ebenfalls auf einen Abbau des *Confirmation Bias* hindeuten. Es scheint direkt so zu sein, dass sich die Raterinnen weniger stark dazu gezwungen fühlen, nach den Einstellungsnormen abzustimmen, ein Ergebnis, das mit der *Social Identity Theory* darauf hindeuten würde, dass die Intervention die soziale Identität zur eigenen Gruppe gestärkt hat. Dies würde wiederum laut Ethnolinguistic Identity Theory dazu führen, dass für diese Gruppe weniger Bedarf besteht, zur prestigereichen Varietät zu greifen. Einen ähnlichen Abbau eines *Confirmation Bias* könnten auch die Ergebnisse des Stadt-Land-Vergleichs zeigen. So ist in dieser Kategorie zu beobachten, dass in den ländlichen Bereichen dialektale Sprachverwendung am wenigsten kompetent eingeschätzt wird, diese Einstellung sich jedoch nach der Intervention stark verändert. Die Auswirkungen für die sozial benachteiligten Gruppen (hier wären dies die Personen des Landes sowie diejenigen in Pendeldistanz und nicht die Sprecherinnen) wären die gleichen wie zuvor – es würde eine Stärkung der sozialen Identität in dieser Region zeigen. Da sich die dialektale Sprechweise in den Gaisers im Pinzgau jedoch nicht mit der im Alltag verwendeten Varietät im Pinzgau deckt, könnten die Ergebnisse auch schlicht auf den Abbau der Abneigung gegenüber einer anderen Dialektgruppe hindeuten.

Bei Betrachtung der Einstellungsveränderung in Bezug auf unterschiedliche Erstsprachen findet man, dass die Rater:innen mit Deutsch als Erstsprache nur in einer einzigen Kategorie eine Veränderung zeigen: Sie schätzen dialektsprechende Sprecherinnen nach der Intervention kompetenter ein. Rater:innen mit Deutsch und einer weiteren Erstsprache und diejenigen mit einer anderen Erstsprache als Deutsch schätzen dagegen sowohl dialektsprechende Sprecherinnen als auch Sprecher nach der Intervention kompetenter ein. Auch wenn die Personenanzahl der beiden letzteren Gruppen zu gering für Signifikanztests ist, sieht man an den deskriptiven Darstellungen, dass beide Gruppen die größten Veränderungen in den Einschätzungen zeigen. Dies würde darauf hindeuten, dass die Intervention auf Personen mit zwei Erstsprachen oder einer anderen Erstsprache als Deutsch einen starken Effekt zeigt. Eventuell ist die Wirksamkeit

daher stark ausgeprägt, da diese Gruppen sprachdiskriminierende Einstellungen auch besonders gut kennen und daher eine Intervention, die diese abbaut, speziell hier auf fruchtbaren Boden fällt. Gleichzeitig ist Nicht-Erstsprachler:innen die positive Bedeutung des Dialekts für die Identität sicherlich weniger bewusst. So konnten sie durch die Intervention besonders viel lernen – auch über die sozio-indexikalische Bedeutung des Dialekts für ihre Mitschüler:innen mit Deutsch als Erstsprache.

Zuletzt noch zur Wirkung der Intervention auf Personen mit unterschiedlichem Bildungshintergrund. Interessanterweise werden dialektsprechende Sprecher:innen vor allem von der Gruppe mit dem niedrigsten und der mit dem höchsten Bildungshintergrund nach der Intervention kompetenter eingeschätzt – aber nur Letztere zeigen hochsignifikante Messwerte. Dies zeigt, dass die stereotype Meinung, dialektsprechende Personen seien weniger kompetent, vor allem in der Gruppe zurückgeht, in der eine solche Einstellung auch zu erwarten war. Das Ergebnis bei den Rater:innen mit dem niedrigsten Bildungshintergrund lässt sich mit dem starken Anstieg der Kompetenzeinschätzung von Personen mit zwei Erstsprachen oder Deutsch als L2 erklären.

### Sympathieebene

Die Sympathieeinschätzungen zu Sprachvarietäten zeigen durch die Intervention weniger starke Veränderungen als diejenigen der Kompetenzeinschätzungen. Einen moderaten Anstieg verzeichnen hier die dialektsprechenden Sprecher:innen und die standardsprechenden Sprecher. Beide wiesen vergleichsweise niedrige Sympathiewerte in den Prätestergebnissen auf. Dies bestätigen zum Teil die Ergebnisse der Varietäten-distanz. So findet man beim Lehrer in informeller Situation einen deutlichen und bei der Schülerin in informeller Situation einen moderaten Rückgang der Einstellungsunterschiede.

Bei Betrachtung der unterschiedlichen Einschätzungsmuster der Raterinnen und Rater fällt auf, dass die Raterinnen die Sprecher im Posttest weniger sympathisch einschätzen als im Prätest. Damit nähern sich ihre Einschätzungen denjenigen der Rater an.

Der Vergleich der Regionen zeigt, dass standardsprachliche Sprechweise am Land im Posttest etwas sympathischer eingeschätzt wird. Ebenfalls wird der dialektsprechende Schüler in informeller Situation von den Rater:innen in Pendeldistanz weniger sympathisch im Posttest eingeschätzt, allerdings kommt es in dieser Gruppe nur zu einer Annäherung an die Werte der beiden anderen Gruppen.

Bei Betrachtung der unterschiedlichen Erstsprachen zeigt sich, dass die beiden Gruppen mit Deutsch und einer weiteren Erstsprache sowie einer anderen Erstsprache als Deutsch dialektsprechende Personen nach der Intervention generell sympathischer einschätzen. Bei den Rater:innen mit Deutsch als L1 sieht man hingegen keine Veränderungen. Die Kategorie „Bildungshintergrund“ zeigt keine Unterschiede in der Sympathieeinschätzung.

### **Ebene der persönlichen Präferenz**

Die Ebene der persönlichen Präferenz zeigt deutlich positivere Einschätzungen von Dialektsprecherinnen im Posttest, was die bisher berichteten Ergebnisse untermauert. Bei den Dialektsprechern findet man hingegen keine Veränderungen. Bei standard-sprachlicher Sprechweise zeigen sich bei beiden Geschlechtern eine moderat positivere Einschätzung im Posttest. Die Analysen der Varietäten- sowie der Geschlechterdistanz erbrachten keine Veränderungen. Ähnlich verhält es sich bei der differenzierten Betrachtung der Geschlechterratings. So sieht man nur leichte Tendenzen der Rater, dialektsprechende Personen nach der Intervention positiver einzuschätzen.

Im Stadt-Land-Vergleich zeigt sich, dass die Rater:innen der ländlichen Regionen dialektsprechende Personen im Posttest deutlich positiver einschätzen. Der Erstsprachenvergleich zeigt, dass die Gruppe mit Deutsch und einer weiteren Erstsprache dialektsprechende Personen nach der Intervention positiver einschätzt, die Personen mit Deutsch als L2 hier jedoch differenzieren. Dialektsprechende Sprecher werden negativer eingeschätzt, dialektsprechende Sprecherinnen positiver. In der Kategorie „Bildungshintergrund“ sind wiederum keine Unterschiede erkennbar.

## 13 Beantwortung der Hypothesen und Ausblick

Im empirischen Teil dieser Arbeit wurden sowohl hypothesengeleitete als auch explorative Überprüfungen der erhobenen Daten mit deskriptiven und inferenzstatistischen Methoden vorgenommen. Dieses doppelte Vorgehen hat es einerseits ermöglicht, die im Vorfeld aufgeworfenen Hypothesen zu überprüfen (s. Abschnitt 8.1) und andererseits, einen detaillierten Blick auf Gruppenunterschiede und Auffälligkeiten zu werfen, die mit einer Beschränkung auf Hypothesentestungen verborgen geblieben wären. Dieses Kapitel widmet sich nun der Zusammenführung der hypothesengeleiteten und explorativen Ergebnisse, um die Vorzüge dieses doppelten Vorgehens aufzuzeigen, sowie der Rückbindung an die Theoriebildung und den Forschungsstand. Ich werde dies anhand der elf im Vorfeld aufgeworfenen Fragestellungen und Hypothesen abhandeln.

- F1: Werden standardsprechende Personen kompetenter eingeschätzt als dialekt-sprechende Personen?
- H<sub>1</sub>: Standardsprechende Personen werden kompetenter eingeschätzt als dialekt-sprechende Personen.
- H<sub>1n</sub>: Es gibt keinen Unterschied in der Kompetenzeinschätzung von standard- und dialekt-sprechenden Personen.

In Abschnitt 4.1 wurde verdeutlicht, dass im österreichischen Kontext eine Reihe von Forschungsergebnissen zu Spracheinstellungen vorliegen. So zeigen etwa Soukup (2009) und Bellamy (2012), dass Standardsprecher:innen kompetenter eingeschätzt werden als Dialekt sprecher:innen. Auch Fuchs/Elspaß (2019) berichten diese Spracheinstellungen von Schüler:innen der 10. Schulstufe, die berufsbildende höhere Schulen besuchen. Daher lag es nahe, diese Spracheinstellungen auch für Schüler:innen der 10. Schulstufe in allgemeinbildenden höheren Schulen anzunehmen, was in H<sub>1</sub> zum Ausdruck kam. Die Testung von H<sub>1</sub> bestätigt den Einstellungsunterschied auf Kompetenzebene zwischen den Varietäten auch mit höchstsignifikanten Werten, daher kann H<sub>1n</sub> eindeutig verworfen werden.

Die explorative Betrachtung der Detailergebnisse offenbart nun jedoch deutliche Unterschiede, etwa in der Betrachtung der unterschiedlichen Rollen Lehrkräfte vs. Schüler:innen. So werden standardsprechende Schüler:innen zumeist kompetenter eingeschätzt als dialekt sprechende Schüler:innen, bei den Lehrkräften trifft dies jedoch nur bei der Lehrerin in formeller Situation zu. Die beiden Lehrer-Guises sowie derjenige der Lehrerin in informeller Situation zeigen keine Unterschiede in der Kompetenzeinschätzung der Sprechweise. Dieser Aspekt legt nahe, dass etwa die gesellschaftliche Rolle einen Faktor bei der Kompetenzeinschätzung darstellt, ein Umstand, der bisher noch keine große Berücksichtigung in der Spracheinstellungsforschung gefunden hat und dem weiter nachgegangen werden sollte. Den Zusammenhang zwischen Spracheinstel-

lungen (als Responsevariable) und gesellschaftlichen Rollen, Geschlecht sowie Alter (als Prädiktoren) weiter zu untersuchen, stellt aus meiner Sicht daher ein zentrales Desiderat dar.

F<sub>2</sub>: Werden dialektsprechende Personen sympathischer eingeschätzt als standardsprechende Personen?

H<sub>2</sub>: Dialektsprechende Personen werden sympathischer eingeschätzt als standardsprechende Personen.

H<sub>2n</sub>: Es gibt keinen Unterschied in der Sympathieeinschätzung von dialekt- und standardsprechenden Personen.

Ähnlich sieht es auf Sympathieebene aus. In den oben zitierten Studien wurde demonstriert, dass Dialektsprecher:innen sympathischer als Standardsprecher:innen eingeschätzt werden, wobei bei Fuchs/Elspaß (2019) ersichtlich wurde, dass der Unterschied zwischen den Varietäten auf dieser Ebene weit weniger groß ist als auf Kompetenzebene. Auch Bellamy (2012) stellte heraus, dass die einzelnen Eigenschaften, die er der Solidaritätsdimension zuschrieb, in Österreich nicht so deutlich unterscheidbar waren wie in der englischsprachigen Vergleichsgruppe. Ähnliche Ergebnisse ließen sich auch in dieser Studie finden. So weist die generelle Testung auf Sympathieebene höchstsignifikante Unterschiede zugunsten einer positiveren Einschätzung dialektsprechender Personen auf. H<sub>2n</sub> kann also ebenfalls verworfen werden.

Die detaillierte Betrachtung liefert jedoch beträchtliche Unterschiede, speziell die Geschlechter betreffend. So trifft man nur bei den Sprechern auf diese Unterschiede, die Ergebnisse der Sprecherinnen unterscheiden sich dagegen nicht voneinander. Das Bild wird noch eindeutiger bei Betrachtung der Ebene der persönlichen Präferenz. Dort werden Sprecherinnen sogar als positiver eingeschätzt, wenn sie sich der Standardsprache bedienen. Die Vermutung liegt also nahe, dass die Ergebnisse deswegen bei Bellamy (2012) und Fuchs/Elspaß (2019) nicht so eindeutig ausfielen, da der Faktor Geschlecht offenbar in der Sympathieeinschätzung eine wesentliche Rolle spielt. Dem sollte in Zukunft jedenfalls genauer nachgegangen werden.

F<sub>3</sub>: Weisen Schülerinnen unabhängig von der Sprachvarietät allgemein positivere Spracheinstellungen in Bezug auf Kompetenz- als auch Sympathieebene auf als Schüler?

H<sub>3</sub>: Schülerinnen weisen unabhängig von der Sprachvarietät allgemein positivere Spracheinstellungen in Bezug auf Kompetenz- als auch Sympathieebene auf als Schüler.

H<sub>3n</sub>: Es gibt keinen Unterschied in der Einschätzung von Spracheinstellungen auf Kompetenz- und Sympathieebene zwischen Schülerinnen und Schülern.

Sowohl Soukup (2009) als auch Bellamy (2012) untersuchten geschlechterspezifisches Ratingverhalten im österreichischen Kontext und fanden, dass Frauen generell positivere Einschätzungen treffen als Männer (s. auch Abschnitt 4.2). Dieser Frage wurde ebenfalls nachgegangen und die Unterschiede im Ratingverhalten ließen sich wieder-

rum belegen. Damit kann  $H_{3n}$  verworfen werden. Trotz dieser generellen Tendenzen kristallisierten sich in der explorativen Betrachtung der Details wiederum deutliche Unterschiede heraus. Ein zentrales Ergebnis ist, dass die Raterinnen Sprecher, egal in welcher Varietätenverwendung, kompetenter einschätzen als die Rater ihr eigenes Geschlecht. Diesen Umstand habe ich auf einen potenziellen *Confirmation Bias* zurückgeführt (s. Abschnitt 9.3). Es wäre jedoch wünschenswert, dem Grund für diesen Geschlechterunterschied in einem experimentellen (psycholinguistischen) Setting auf den Grund zu gehen.

F<sub>4</sub>: Schätzen Schülerinnen standardsprechende Personen sympathischer ein als Schüler?

H<sub>4</sub>: Schülerinnen schätzen standardsprechende Personen sympathischer ein als Schüler.

H<sub>4n</sub>: Es gibt keinen Unterschied in der Sympathieeinschätzung standard-sprechender Personen zwischen Schülerinnen und Schülern.

H<sub>4</sub> lehnt sich einerseits an die frühen geschlechterspezifischen soziolinguistischen Ergebnisse an (z. B. Trudgill 1972), dass Frauen standardnäher sprechen als Männer, sowie an die Ergebnisse von Brouwer (1989), die fand, dass Frauen auch positivere Einstellungen zur Standardsprache mitbringen. Letzteres kann durch diese Studie nicht bestätigt werden. So schätzen die Raterinnen standardsprechende Sprecherinnen ähnlich sympathisch ein wie die Rater, H<sub>4n</sub> kann daher nicht verworfen werden.

Die Analyse von Geschlechterunterschieden hat sich in dieser Arbeit jedoch als fruchtbares Forschungsfeld erwiesen. So konnten abseits der Hypothesen viele Unterschiede ausgemacht werden. Es wurde etwa ein massiver Einstellungsunterschied hinsichtlich der Lehrkräfte gefunden. Der Lehrer wurde auf jeder Ebene und in jeder Situation höchstsignifikant positiver eingeschätzt als die Lehrerin, mit teils großen Distanzen der Werte. Was die Gründe für diesen Genderbias sind, konnte aber nicht durch die Studie eruiert werden. Es sind diesbezüglich jedoch mehrere Aspekte auffällig. Zum Ersten ist der Einstellungsunterschied bei den Schüler:innen nicht in dieser Weise vorhanden. Die Schüler werden zwar ebenfalls häufiger als kompetenter eingeschätzt, aber eben nicht immer. Auf Sympathieebene und auf Ebene der persönlichen Präferenz ist es sogar so, dass die standardsprechende Schülerin in informeller Situation positiver eingeschätzt wird als die gleiche Schülerin in dialektaler Sprachverwendung. Zum Zweiten scheint es so zu sein, dass Deutschlehrerinnen standardnäher sprechen als Deutschlehrer. Zum Dritten scheint es überdies hinaus der Fall zu sein, dass es viel mehr Deutschlehrerinnen gibt als Deutschlehrer. Die positive Einschätzung vieler Rater:innen dem Deutschlehrer gegenüber könnte auch auf ein imaginiertes Idealbild eines Lehrers zurückzuführen sein.

Hier scheinen sich sowohl rollenspezifische als auch geschlechterspezifische Unterschiede zu zeigen, die durch weiterführende Forschung differenziert werden sollten. Eine Möglichkeit wäre, ein Verbal-Guise-Design zu kreieren, in dem der gleiche Satzinhalt von weiblichen und männlichen Personen eingesprochen wird, dann könnten zumindest Einstellungsunterschiede auf inhaltlicher Ebene ausgeschlossen wer-

den. In weiterführenden Studien könnte auch der Faktor Lehrkraft kontrolliert werden, indem man für eine diesbezügliche Geschlechterbalance im Studiendesign sorgt. Die getätigten Analysen der Geschlechterunterschiede geben einen Eindruck, welchen zentralen Stellenwert das Geschlecht in der Einschätzung der Sprachverwendung hat. Die Einbeziehung der Geschlechterkomponente sollte daher zum Standard weiterer Spracheinstellungsforschung werden.

F<sub>5</sub>: Schätzen Schüler:innen am Land dialektsprechende Personen kompetenter ein als Schüler:innen in der Stadt?

H<sub>5</sub>: Schüler:innen am Land schätzen dialektsprechende Personen kompetenter ein als Schüler:innen in der Stadt. Schüler:innen in Pendeldistanz liegen mit ihrer Einschätzung zwischen den beiden anderen Regionen.

H<sub>5n</sub>: Es existiert kein Unterschied zwischen den Kompetenzeinschätzungen von Schüler:innen am Land, in Pendeldistanz zur Stadt und der Stadt in Bezug auf dialektsprechende Personen.

Da vor dieser Studie noch keine Untersuchungen zu regionalen Unterschieden bei Spracheinstellungen durchgeführt wurden, griffen die beiden Hypothesen H<sub>5</sub> und H<sub>6</sub> auf die variablenanalytischen Erkenntnisse der Soziolinguistik zurück (s. Abschnitt 4.4). Da am Land noch vermehrt Dialekte gesprochen werden, lag es auf der Hand anzunehmen, dass sich dies auch in den Spracheinstellungen abbildet, etwa in einer kompetenteren Einschätzung dialektsprechender Personen am Land. Dieses Ergebnis stellte sich jedoch nicht ein. Auf Kompetenzebene konnten in den Regionen keine unterschiedlichen Einstellungen gefunden werden, daher kann auch H<sub>5n</sub> nicht abgelehnt werden. Interessant ist auch, dass die detaillierte Betrachtung der Guises im Unterschied zu den vorherigen Forschungsfragen ebenfalls keine Unterschiede auf Kompetenzebene offenbarte. Die Kompetenzeinschätzungen scheinen damit sehr homogen in den Regionen verteilt zu sein.

Da in allen drei Regionen die gleichen Guises verwendet wurden und die dialektalen am ehesten der Alltagskommunikation des Flach- & Tennengaus sowie (etwas weniger) der Stadt Salzburg entsprechen (s. Abschnitt 9.4), wäre es durchaus wünschenswert, die Erhebung im Pinzgau mit Sprecher:innen aus der Region erneut durchzuführen und die Ergebnisse zu vergleichen. Somit könnte festgestellt werden, ob die Kompetenzeinschätzung wirklich regionsübergreifend stabil sind oder die eigene Varietät im Pinzgau doch eine positivere Kompetenzeinschätzung erfährt.

F<sub>6</sub>: Schätzen Schüler:innen in der Stadt standardsprechende Personen sympathischer ein als Schüler:innen am Land?

H<sub>6</sub>: Schüler:innen in der Stadt schätzen standardsprechende Personen sympathischer ein als Schüler:innen am Land. Schüler:innen in Pendeldis-

tanz liegen mit ihrer Einschätzung zwischen den beiden anderen Regionen.

- H<sub>6n</sub>: Es existiert kein Unterschied zwischen den Sympathieeinschätzungen von Schüler:innen am Land, in Pendeldistanz zur Stadt und der Stadt in Bezug auf standardsprechende Personen.

Auf Sympathieebene treten eher die Ergebnisse auf, die in H<sub>6</sub> prognostiziert wurden. So schätzten die Rater:innen im Pinzgau standardsprachliche Sprechweise signifikant weniger sympathisch ein. Interessanterweise lagen die Sympathieeinschätzungen jedoch nicht in der Stadt Salzburg am höchsten, sondern in Pendeldistanz. Da signifikante Unterschiede gefunden wurden, kann H<sub>6n</sub> dennoch verworfen werden.

Für die Guises auf Sympathieebene in standardsprachlicher Sprechweise wurde auf Sprecher:innen zurückgegriffen, die österreichischen Standard sprechen. Darüber hinaus wäre es jedoch zusätzlich interessant, wie sympathisch bundesdeutscher Standard von dieser Altersgruppe in den Gymnasien im Vergleich zu österreichischem Standard eingeschätzt wird und dies sowohl regional als auch das Geschlecht sowie die Rolle betreffend darzustellen. Diese Fragestellung könnte man gleichzeitig auf dialektale Varietäten anwenden. So wurden in dieser Studie Guises verwendet, die sich einer Varietät des Salzburger Zentralraums bedienen. Es wäre durchaus interessant, die Studie mit unterschiedlichen dialektalen Varietäten zu wiederholen, um eine detailliertere Erkenntnis über Spracheinstellungen zu gewinnen.

- F<sub>7</sub>: Gibt es einen Unterschied in der Sympathieeinschätzung von Schüler:innen mit Deutsch als Erstsprache, Deutsch und einer weiteren Erstsprache und einer anderen Erstsprache als Deutsch in Bezug auf dialektsprechende Personen?
- H<sub>7</sub>: Schüler:innen mit Deutsch als Erstsprache schätzen dialektale Sprechweise sympathischer ein als Schüler:innen mit einer anderen Erstsprache als Deutsch. Schüler:innen mit Deutsch und einer weiteren Erstsprache liegen mit ihrer Einschätzung zwischen den beiden anderen Gruppen.
- H<sub>7n</sub>: Es existiert kein Unterschied in der Sympathieeinschätzung von Schüler:innen mit Deutsch als Erstsprache, Deutsch und einer weiteren Erstsprache und einer anderen Erstsprache als Deutsch in Bezug auf dialektsprechende Personen.

Die Überprüfung der Auswirkung unterschiedlicher Erstsprachen auf die Sympathieeinschätzung von dialektsprechenden Personen bestätigte die Vermutung in H<sub>7</sub>. Zwischen den Rater:innen mit Deutsch als L1 und L2 wurden hochsignifikante Unterschiede bei der Einschätzung der Dialektsprecher:innen gefunden. H<sub>7n</sub> kann daher verworfen werden. Bemerkenswert in dieser Untersuchung ist, dass auch in dieser Kategorie deutliche Geschlechterunterschiede gefunden wurden. So werden Dialektsprecherinnen von den unterschiedlichen erstsprachlichen Herkunftsgruppen sehr ähnlich eingeschätzt. Gleichzeitig schätzen Rater:innen mit Deutsch als L2 Dialektsprecher und Dialektsprecherinnen ebenfalls ähnlich ein. Die Einstellungsunterschiede zwischen den Geschlechtern scheinen also vor allem von den Rater:innen mit Deutsch als L1

herzurühren und die Dialektsprecher zu betreffen. Man könnte daher noch näher eruieren, wie diese Einstellungsunterschiede in der Bevölkerung verteilt sind. Bei einer dementsprechend großen Stichprobe könnten Korrelationen mit anderen Kategorien wie Alter bzw. Bildungshintergrund interessante Einblicke ermöglichen.

- F<sub>8</sub>: Schätzen Schüler:innen mit unterschiedlichen Bildungshintergründen dialekt-sprechende Personen auf Kompetenzebene unterschiedlich ein?
- H<sub>8</sub>: Schüler:innen mit niedrigem Bildungshintergrund schätzen dialekt-sprechende Personen kompetenter ein als Schüler:innen mit hohem Bildungshintergrund.
- H<sub>8n</sub>: Es existiert kein Unterschied in der Kompetenzeinschätzung von Schüler:innen mit niedrigem oder hohem Bildungshintergrund in Bezug auf dialekt-sprechende Personen.

Die Untersuchung des Bildungshintergrunds erbrachte alles in allem keine eindeutigen Ergebnisse, was mit der Stichprobe zu tun haben könnte. Die geringe Anzahl der Personen, die eine der beiden niedrigsten Kategorien angaben (0–10 und 11–25 Bücher), weist darauf hin, dass auch heutzutage Gymnasialschüler:innen einer 10. Schulstufe eher aus bildungsnahen Elternhäusern stammen. Was die Kompetenzeinschätzung dialekt-sprechender Personen betrifft, konnten keine Unterschiede festgestellt werden. H<sub>8n</sub> kann also nicht verworfen werden.

- F<sub>9</sub>: Schätzen Schüler:innen mit unterschiedlichen Bildungshintergründen standardsprechende Personen auf Sympathieebene unterschiedlich ein?
- H<sub>9</sub>: Schüler:innen mit hohem Bildungshintergrund schätzen standardsprechende Personen sympathischer ein als Schüler:innen mit niedrigem Bildungshintergrund.
- H<sub>9n</sub>: Es existiert kein Unterschied in der Sympathieeinschätzung von Schüler:innen mit niedrigem oder hohem Bildungshintergrund in Bezug auf standardsprechende Personen.

Die Sympathieeinschätzung standardsprechender Personen brachte zwar signifikante Unterschiede den Bildungshintergrund betreffend zum Vorschein, allerdings nicht in der Form, die in H<sub>9</sub> erwartet wurde. Die Gruppe mit dem zweithöchsten Bildungshintergrund (100–200 Bücher) schätzt standardsprechende Personen signifikant weniger sympathisch ein als die beiden angrenzenden Gruppen. Für diesen Umstand wurde auch keine Erklärung gefunden. Man könnte dem noch näher nachgehen, indem man die Gruppen mit unterschiedlichem Bildungshintergrund zusätzlich in weitere Kategorien (Rater:innengeschlecht, Erstsprache etc.) teilt. Die Stichprobengröße dieser Untersuchung ist dafür jedoch nicht ausreichend. Zentraler wäre noch, diese Untersuchung mit einer weniger spezifischen Stichprobe (im Idealfall einer repräsentativen) durchzuführen.

F<sub>10</sub>: Lassen sich Spracheinstellungen in Bezug auf die Kompetenzeinschätzung dialektsprechender Personen verändern?

H<sub>10</sub>: Die Intervention sorgt dafür, dass dialektsprechende Personen nach der Intervention von den Schüler:innen kompetenter eingeschätzt werden.

H<sub>10n</sub>: Schüler:innen schätzen nach erfolgter Intervention dialektsprechende Personen auf Kompetenzebene gleich ein wie zuvor.

Was die Veränderung von Spracheinstellungen durch schulische Interventionen betrifft, ist es nicht möglich, auf eine empirische Forschungstradition zurückzugreifen. Einzig Johnson et al. (2017) untersuchten die Wirkung eines CLA-Programms (Critical Language Awareness) auf das Schreibverhalten in der Primarstufe, das sich vor allem auf kontrastiven Unterricht stützte (s. Abschnitt 6.2), die Spracheinstellungen wurden in dieser Studie jedoch nicht überprüft. Es wurden daher nur die zwei sehr allgemeinen Hypothesen H<sub>10</sub> und H<sub>11</sub> formuliert. Dabei konnte nachgewiesen werden, dass Dialektsprecher:innen nach der Intervention von den Rater:innen deutlich kompetenter eingeschätzt wurden. H<sub>10n</sub> kann damit eindeutig verworfen werden.

Die explorative Betrachtung der Prä-Posttest-Ergebnisse offenbarte jedoch noch weitere Unterschiede. So wurden nach der Intervention vor allem die Dialektsprecher:innen kompetenter eingeschätzt, die Gruppe also, die die niedrigsten Werte im Prätest erhielt. Interessant ist auch, dass vor allem die Raterinnen ihre Meinung zu den Dialektsprecher:innen geändert haben. Die Intervention scheint damit vor allem den Abbau stereotyper Vorstellungen in Bezug auf das eigene Geschlecht zu bewirken, da auch bei den Ratern die gleiche Tendenz wie bei den Raterinnen bemerkt werden konnte. Die Veränderung bei den Ratern hinsichtlich der Sprecher ist jedoch weniger stark ausgeprägt, da auch die Prätestwerte viel höher als diejenigen der Sprecherinnen liegen und dies eine größere Veränderung auch nicht zuließ.

Der Wirkungsweise eines CLA auf die Spracheinstellungen dem eigenen Geschlecht gegenüber noch näher nachzugehen, würde ein spannendes (neues) Betätigungsfeld am Schnittpunkt psycholinguistischer, soziolinguistischer und didaktischer Forschung darstellen, das in weiterer Folge auch auf nicht sprachspezifische Fragestellungen übertragen werden könnte. In diesem Zusammenhang könnte ebenfalls der Frage nachgegangen werden, ob sich ein CLA dazu eignet, einen *Confirmation Bias* abzubauen, wie ich vermute. Würde sich dies bestätigen, könnte es weitreichende Folgen für den Einsatz dieser Methode in der schulischen Didaktik mit sich bringen.

Ein weiteres Forschungsdesiderat in Bezug auf die Veränderung der Kompetenzeinschätzung stellt die neuerliche Überprüfung der Untersuchung bei Personen mit einer anderen Erstsprache als Deutsch oder Deutsch und einer weiteren Erstsprache dar. Die Intervention scheint ja speziell bei diesen beiden Gruppen starke Veränderungen ausgelöst zu haben. Leider konnte dies durch die kleinen Stichprobengrößen inferenzstatistisch nicht abgesichert werden. Eine Überprüfung mit entsprechender Stichprobengröße wäre also wünschenswert.

Ein ähnliches Desiderat betrifft die Stadt-Land-Unterschiede. So zeigte sich, dass die Intervention bei den Rater:innen am Land die größten Veränderungen bewirkte.

Da jedoch eine andere dialektale Varietät als die regional verwendete zur Überprüfung der Spracheinstellungen am Land verwendet wurde, wäre es auch in Bezug auf die Wirkungsweise der Intervention sinnvoll, die Studie mit Guises der gesprochenen Alltagsvarietät des Pinzgaus zu replizieren. Damit könnte der Frage nachgegangen werden, ob sich die höhere Kompetenzeinschätzung der Dialektsprecher:innen im Pinzgau auf die „eigene“ Varietät oder auf diejenige des Salzburger Zentralraums bezieht. Eine Untersuchung, die Guises beider dialektaler Varietäten einschätzen lässt, könnte eine Antwort auf diese Frage liefern.

F<sub>11</sub>: Lassen sich Spracheinstellungen in Bezug auf die Sympathieeinschätzung standardsprechender Personen verändern?

H<sub>11</sub>: Die Intervention sorgt dafür, dass standardsprechende Personen nach der Intervention von den Schüler:innen sympathischer eingeschätzt werden.

H<sub>11n</sub>: Schüler:innen schätzen nach erfolgter Intervention standardsprechende Personen auf Sympathieebene gleich ein wie zuvor.

Die letzte Hypothese geht der Frage nach der Veränderbarkeit der Sympathieeinschätzung von Standardsprecher:innen durch die Intervention nach. Auch auf dieser Ebene zeigte sich, dass Rater:innen nach der Intervention Standardsprecher:innen signifikant sympathischer einschätzen. Damit kann H<sub>11n</sub> verworfen werden. Durch weiterführende explorative Betrachtungen offenbarte sich wiederum ein deutlicher Geschlechterunterschied. So wurden speziell die standardsprechenden Sprecher und die dialektsprechenden Sprecherinnen nach der Intervention sympathischer eingeschätzt, wobei die Veränderungen schwächer ausfallen als auf Kompetenzebene. Dabei sollte in weiterführender Forschung vor allem jenes unerwartete Ergebnis noch genauer untersucht werden, dass Standardsprecherinnen im Untersuchungsgebiet von den Rater:innen nicht weniger sympathisch eingeschätzt werden als Dialektsprecherinnen. Auf Ebene der persönlichen Präferenz zeigt sich sogar, dass Standardsprecherinnen positiver eingeschätzt werden als Dialektsprecherinnen. Diesem Umstand sollte auf jeden Fall nachgegangen werden, da er bisher in der Forschung noch nicht berichtet wurde. Zum einen gehört überprüft, ob sich dieses Ergebnis auch in anderen Regionen Österreichs oder auch generell im deutschsprachigen Raum zeigt. Zum anderen müsste man sich die Einschätzungen anderer Stichproben ansehen, ob es sich hier um ein Muster einer bestimmten Bildungsgruppe oder doch ein allgemeines Muster handelt.

Gleichzeitig existieren auch einige Desiderata zum methodischen Vorgehen. Es wären generell Replikationsstudien (in der gleichen Region, aber auch in anderen Regionen) wünschenswert, die die Wirkung solcher Materialien ebenfalls empirisch untersuchen, um die Ergebnisse zu validieren. Darüber hinaus läge ein weiterer wichtiger Untersuchungsschritt darin, Langzeiteffekte zu erheben. Für die Stichprobe dieser Studie ist dieser Schritt angedacht und die erforderlichen Erhebungen wurden schon in die Wege geleitet. Die Erhebung von Spracheinstellungen mit einer Likert-Skala wurde ebenfalls in dieser Studie zum Problem. So zeigen die Einstellungen, die auf einer siebenteiligen Skala von „Frowning-Face-Smiley“ bis „Happy-Face-Smiley“ erhoben wur-

den, häufig eine starke linksschiefe Verteilung. Einschätzungen im negativen Bereich wurden von den meisten Rater:innen also vermieden. Dabei wäre es aus statistischen Gründen ideal, wenn die Mitte der Einstellungsskala auch mit der Mitte der Einschätzungen einhergeht. Ob und wie dies mit anderen Skalenbeschriftungen, Skalenabstufungen, Schieberegler usw. erreicht werden könnte, wäre eine Untersuchung wert. Gleichzeitig könnte es der Fall sein, dass die Nutzung einer Skala mit ungeraden Zahlen und dem damit einhergehenden Vorliegen einer „Mitte“ manche Rater:innen speziell in den Kontrollgruppen beim Posttest dazu verleitet hat, aus fehlender kognitiver Auseinandersetzung mit der Aufgabe (Langeweile, Desinteresse usw.) diese Option zu wählen. Diesem Ratingverhalten sollte ebenfalls nachgegangen werden.

Viele weitere Ideen könnte man an dieser Stelle anführen. Das größte Desiderat von meiner Seite liegt jedoch darin, dass die Ergebnisse an die Bildungsinstitutionen und Entscheidungsträger im Bildungsbereich herangetragen werden und im regulären Deutschunterricht Einzug halten. Es wäre eine sehr einfache und kostengünstige Möglichkeit, wie Bewusstheit und Verständnis unterschiedlichen Sprachgruppen gegenüber im Deutschunterricht aufgebaut werden kann.



## 14 Schluss

Trotz der schwierigen Ausgangslage, in der die Untersuchung stattfand – im Wintersemester 2021 war die Lage wegen der anhaltenden Coronapandemie sehr angespannt, die Schulen standen kurz vor der Schließung – hat sich gezeigt, dass Unterrichtsmaterialien schon in einem Ausmaß von zehn Unterrichtseinheiten stereotype Spracheinstellungen von Schülerinnen und Schülern zu den vielgebrauchten Alltagsvarietäten verändern. Speziell auf Kompetenzebene sind diese Einstellungsveränderungen deutlich sichtbar. Das herausragende Ergebnis ist, dass dialektsprechende Sprecher:innen nach der Intervention kompetenter eingeschätzt werden als vorher. Somit zeigt sich, welche Rolle der Deutschunterricht bei der Reflexion von Sprachideologien leisten kann. Diese reflektierte Herangehensweise besitzt das Potenzial, die Chancengleichheit von Sprecher:innen unterschiedlicher Herkunftsvarietäten (und auch Sprachen) zu erhöhen, indem unreflektierte und zum Teil diskriminierende Stereotype bewusst gemacht, reflektiert und gegebenenfalls abgebaut werden können.

Es wäre auch mit überschaubarem Aufwand möglich, die entwickelten Materialien für andere Regionen des deutschsprachigen Raums umzugestalten, um auf diese Weise einen reflektierteren Umgang unterschiedlicher sozialer Gruppen miteinander noch breiter zu forcieren.



# Literaturverzeichnis

- Adler, A. & Plewina, A. (2019): Die Macht der großen Zahlen. Aktuelle Spracheinstellungen in Deutschland. In: Eichinger, L. & Plewina, A. (Hrsg.): *Neues vom heutigen Deutsch. Empirisch – methodisch – theoretisch*. Berlin [u. a.]: De Gruyter, 141–162. DOI: <https://doi.org/10.1515/9783110622591-008>.
- Agha, A. (2003): The social life of cultural value. In: *Language and Communication* 23, 231–273. DOI: [https://doi.org/10.1016/S0271-5309\(03\)00012-0](https://doi.org/10.1016/S0271-5309(03)00012-0).
- Agha, A. (2004): Registers of Language. In: Duranti, A. (Hrsg.): *A Companion to Linguistic Anthropology*. Malden, MA/Oxford/Carlton, Victoria: Blackwell, 23–45.
- Ammon, U. (1972): Dialekt als sprachliche Barriere. In: *Muttersprache* 82.4, 224–237.
- Ammon, U. (2003): Dialektschwund, Dialekt-Standard-Kontinuum, Diglossie. Drei Typen des Verhältnisses Dialekt – Standardvarietät im deutschen Sprachgebiet. In: Androutsopoulos, J. & Ziegler, E. (Hrsg.): *„Standardfragen“. Soziolinguistische Perspektiven auf Sprachgeschichte, Sprachkontakt und Sprachvariation*. Frankfurt a. M. [u. a.]: Peter Lang, 163–171.
- Ammon, U. & Kellermeier, B. (1997): Dialekt als Sprachbarriere passé? 25 Jahre danach: Versuch eines Diskussions-Erweckungsküsschens. In: *Deutsche Sprache: Zeitschrift für Theorie, Praxis, Dokumentation* 25.1, 21–38.
- Androutsopoulos, J. (2019): Ethnolekt im Diskurs: Geschichte und Verfahren der Registrierung ethnisch geprägter Sprechweisen in Deutschland. In: Gerd A., Niehr, T. & Spitzmüller J. (Hrsg.): *Handbuch Sprache im Urteil der Öffentlichkeit*. Berlin, Boston: De Gruyter (= Handbücher Sprachwissen), 353–382.
- Antoniou, K., Grohmann, K. K. & Kambanaros, M. [et al.] (2016): The effect of childhood bilingualism and multilingualism on executive control. In: *Cognition* 149, 18–30. DOI: <https://doi.org/10.1016/j.cognition.2015.12.002>.
- Auer, P. (1986): Kontextualisierung. In: Auer, P. (2013): *Sprachliche Interaktion. Eine Einführung anhand von 22 Klassikern*. Berlin/Boston: De Gruyter, 22–47.
- Bangel, M. (2012): Vom Mega-Event zum Concerto Grosso. Fremdwörter in der Musik- und Jugendsprache. In: *Praxis Deutsch* 235, 35–43.
- Baßler, H. & Spiekermann, H. (2001): Dialekt und Standardsprache im DaF-Unterricht. Wie Schüler urteilen – wie Lehrer urteilen. In: *Linguistik online* 9.2, 1–15. DOI: <https://doi.org/10.13092/lo.9.966>.
- Bates, D., Mächler, M. & Bolker, B. (2015): Fitting Linear Mixed Effects Models lme4. In: *Journal of Statistical Software* 67, 1–48. DOI: <https://doi.org/10.18637/jss.v067.i01>.
- Bausinger, H. (1983): Dialekt als Unterrichtsgegenstand. In: *Deutschunterricht: Beiträge zu seiner Praxis und wissenschaftlichen Grundlegung* 35, 98–110.
- Bayerisches Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst (2015): *Dialekte in Bayern. Handreichung für den Unterricht. Mit 2 DVDs*.
- Bayerisches Staatsministerium für Unterricht und Kultus (2019): *MundART WERTvoll. Lebendige Dialekte an bayerischen Schulen. Handreichung*.

- Bekes, P. & Neuland, E. (2006): Norm und Variation in Lehrwerken und im muttersprachlichen Unterricht. In: Neuland, E. (Hrsg.): *Variation im heutigen Deutsch: Perspektiven für den Sprachunterricht*. Bd. Sprache – Kommunikation – Kultur. Soziolinguistische Beiträge. Frankfurt a. M. [u. a.]: Peter Lang, 507–524. URL: <https://permalink.obvsg.at/AC05803132>.
- Bell, A. (1984): Language Style as Audience Design. In: *Language in Society* 13.2, 145–204.
- Bellamy, J. (2012): Language Attitudes in England and Austria. A Sociolinguistic Investigation into Perceptions of High and Low-Prestige Varieties in Manchester and Vienna. *Zeitschrift für Dialektologie und Linguistik – Beihefte*. Stuttgart: Steiner.
- Bernstein, B. (1964): Elaborated and Restricted Codes: Their Social Origins and Some Consequences. In: *American anthropologist* 66.6, 55–69.
- Berthele, R. (2008): Dialekt-Standard Situationen als embryonale Mehrsprachigkeit. Erkenntnisse zum interlingualen Potenzial des Provinzlerdaseins. In: *Sociolinguistica* 22: 87–107. DOI: <https://doi.org/10.1515/9783484605299.87>.
- Berthele, R. (2010): Dialekt als Problem oder Potenzial? Überlegungen zur Hochdeutschoffensive in der deutschen Schweiz aus Sicht der Mehrsprachigkeitsforschung. In: Bitter Bättig, F. & Tanner, A. (Hrsg.): *Sprachen lernen – durch Sprache lernen*. Zürich: Seismo, 37–52.
- Berthele, R. (2016 [PrePrint Online]): Sprachliche Heterogenität und Schule. Eine Kritik an der Linguistisierung sozialer Probleme. In: *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete* 85.3, 1–14. DOI: <https://doi.org/10.2378/vhn2016.art45d>.
- Besch, W. (1975): Dialekt als Barriere bei der Erlernung der Standardsprache. In: Moser, H. (Hrsg.): *Sprachwissenschaft und Sprachdidaktik: Jahrbuch 1974 des Instituts für deutsche Sprache. Sprache der Gegenwart*. Düsseldorf: Schwann (= Sprache der Gegenwart), 150–165. URL: <https://ids-pub.bsz-bw.de/frontdoor/index/index/docId/1319%20http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:bsz:mh39-13195>, [abgerufen am 8. Oktober 2024].
- Besch, W. (2003): Die Entstehung und Ausformung der neuhochdeutschen Schriftsprache/Standardsprache. In: Besch, W., Betten, A. & Reichmann, O. [et al.] (Hrsg.): *Sprachgeschichte. Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung*. Berlin/New York: De Gruyter (= Handbücher zur Sprach- und Kommunikationswissenschaft), 2252–2296. URL: <https://permalink.obvsg.at/AC03076960>.
- Bialystok, E. (2001): Metalinguistic Aspects of Bilingual Processing. In: *Annual Review of Applied Linguistics* 21, 169–181.
- Bialystok, E., Craik, F. I. M. & Freedman, M. (2007): Bilingualism as a protection against the onset of symptoms of dementia. In: *Neuropsychologia* 45, 459–464.
- Bos, W., Lankes, E.-M. & Prenzel, M. [et al.] (Hrsg.) (2003): *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich*. Münster: Waxmann.
- Böttger, H. (2016): Mehr als eine Sprache sprechen. In: Böttger, H. (Hrsg.): *Neurodidaktik des frühen Sprachenlernens*. Bad Heilbrunn/Stuttgart: Julius Klinkhardt/UTB, 177–188.
- Brouwer, D. (1989): *Gender Variation in Dutch. A Sociolinguistic Study of Amsterdam Speech. Topics in Sociolinguistics*. Dordrecht/Providence, R. I.: FORIS PUBLICATIONS.

- Buchberger, W., Eigler, N. & Kühberger, C. (2020): *Mit Concept Cartoons politisches Denken anregen. Ein methodischer Zugang zum subjektorientierten politischen Lernen*. Frankfurt am Main: Wochenschau.
- Bücherl, R. (1993): Dialekt als Chance. In: Klotz, P. & Sieber, P. (Hrsg.): *Vielerlei Deutsch. Umgang mit Sprachvarietäten in der Schule*. Stuttgart: Ernst Klett Schulbuchverlag, 68–77.
- Bucholtz, M., Lopez, A. & Mojarro, A. [et al.] (2014): Sociolinguistic Justice in the Schools: Student Researchers as Linguistic Experts. In: *Language and Linguistics Compass* 8.4, 144–157.
- Bülow, L., Wittibschlager, A. & Lenz, A. N. (2023): Variation and change of relativizers in Austria's German varieties. In: *Sprachwissenschaft* 48.3, 243–280.
- Carifio, J. & Perla, R. J. (2007): Ten Common Misunderstandings, Misconceptions, Persistent Myths and Urban Legends about Likert Scales and Likert Response Formats and their Antidotes. In: *Journal of Social Sciences* 3.3, 106–116.
- Charity, A., Scarborough, H. & Griffin, D. (2004): Familiarity with School English in African American children and its relation to early reading achievement. In: *Child Development* 75.5, 1340–1356. DOI: <https://doi.org/10.1111/j.1467-8624.2004.00744.x>.
- Cheshire, J. (2008): Sex and gender in variationist research. In: Chambers, J. K., Trudgill, P. & Schilling-Estes, N. (Hrsg.): *The Handbook of Language Variation and Change*. Malden, Massachusetts/Oxford/Carlton, Victoria: Wiley-Blackwell, 423–443.
- de Cillia, R., Fink, I. & Ransmayr, J. (2013): Österreichisches Deutsch als Unterrichts- und Bildungssprache an österreichischen Schulen. In: *ÖDaF-Mitteilungen* 2, 34–47.
- de Cillia, R., Fink, I. & Ransmayr, J. (2017): Varietäten des Deutschen an österreichischen Schulen. Ergebnisse des Forschungsprojekts „Österreichisches Deutsch als Unterrichts- und Bildungssprache“. In: Davies, W., Häcki Buhofer, A. & Schmidlin, R. [et al.] (Hrsg.): *Standardsprache zwischen Norm und Praxis. Theoretische Betrachtungen, empirische Studien und sprachdidaktische Ausblicke*. Tübingen: Francke (= Basler Studien zur deutschen Sprache und Literatur), 207–234.
- de Cillia, R. & Ransmayr, J. (2019): *Österreichisches Deutsch macht Schule. Bildung und Deutschunterricht im Spannungsfeld von sprachlicher Variation und Norm*. Wien/Köln/Weimar: Böhlau.
- Clark, R., Fairclough, N. & Ivanič, R. [et al.] (1990): Critical Language Awareness Part I: A Critical Review of Three Current Approaches to Language Awareness. In: *Language and Education* 4.4, 249–260. DOI: <https://doi.org/10.1080/09500789009541291>.
- Clark, R., Fairclough, N. & Ivanič, R. [et al.] (1991): Critical language awareness part II: Towards critical alternatives. In: *Language and Education* 5.1, 41–54.
- Craig, H., Kolenic, G. & Hensel, S. (2014): African American English-Speaking Students: a Longitudinal Examination of Style Shifting From Kindergarten Through Second Grade. In: *Journal of Speech, Language, and Hearing Research* 57.1, 143–157. DOI: <https://doi.org/10.1044/1092-4388>.

- Davies, W. (2007): Normbewusstsein, Normkenntnis und Normtoleranz von Deutschlehrkräften. In: Neuland, E. (Hrsg.): *Variation im heutigen Deutsch: Perspektiven für den Sprachunterricht*. Bd. Sprache – Kommunikation – Kultur. Soziolinguistische Beiträge. Frankfurt a. M. [u. a.]: Peter Lang, 483–491.
- Deák, J. (2007): African-American Language and American Linguistic Cultures: An Analysis of Language Policies in Education. In: *Working Papers in Educational Linguistics* 22.1, 105–134.
- Dittmar, N. (1997): *Grundlagen der Soziolinguistik – Ein Arbeitsbuch mit Aufgaben. Konzepte der Sprach- und Literaturwissenschaft*. Tübingen: Max Niemeyer.
- Dopsch, H. (2004): Das Zeugnis der Ortsnamen. In: EuRegio Salzburg – Berchtesgadener Land – Traunstein (Hrsg.): *Heimat mit Geschichte und Zukunft*. Trostberg: Erdl Druck Medien, 23–30.
- Dragojevic, M. (2017): Language Attitudes. In: *Oxford Research Encyclopedias, Communication*, 1–30. DOI: <https://doi.org/10.1093/acrefore/9780190228613.013.437>.
- Eagly, A. & Chaiken, S. (1993): *The Psychology of Attitudes*. New York: Harcourt Brace Jovanovic College Publishers.
- Eckert, P. (2008): Variation and the indexical field. In: *Journal of Sociolinguistics* 12.4, 453–476.
- Eckert, P. (2012): Three Waves of Variation Study: The Emergence of Meaning in the Study of Sociolinguistic Variation. In: *Annual Reviews of Anthropology* 41, 87–100. URL: <https://www.jstor.org/stable/23270700>.
- Edwards, J. (2012): *Language and Identity. An introduction. Key Topics in Sociolinguistics*. Cambridge: Cambridge University Press.
- Ender, A. (2020): Zum Zusammenhang von Dialektkompetenz und Dialektbewertung in Erst- und Zweitsprache. In: Hundt, M., Kleene, A. & Plewnia, A. [et al.] (Hrsg.): *Regionalekte – objektive Sprachdaten und subjektive Sprachwahrnehmung*. Tübingen: Narr.
- Ender, A. & Kaiser, I. (2009): Zum Stellenwert von Dialekt und Standard im österreichischen und Schweizer Alltag – Ergebnisse einer Umfrage. In: *Zeitschrift für germanistische Linguistik* 37.2, 266–295.
- Ender, A. & Kaiser, I. (2014): Diglossie oder Dialekt-Standard-Kontinuum? Zwischen kollektiver, individueller, wahrgenommener und tatsächlicher Sprachvariation in Vorarlberg und im bairischsprachigen Österreich. In: Huck, Dominique (Hrsg.): *Alemannische Dialektologie: Dialekte im Kontakt. Beiträge zur 17. Arbeitstagung für alemannische Dialektologie in Straßburg vom 26.-28.10.2011*. Zeitschrift für Dialektologie und Linguistik. Beihefte. Stuttgart: Franz Steiner Verlag, 131–146.
- Ender, A., Kasberger, G. & Kaiser, I. (2017): Wahrnehmung und Bewertung von Dialekt und Standard durch Jugendliche mit Deutsch als Erst- und Zweitsprache. In: *ÖDaF-Mitteilungen* 33.1, 97–110.
- Erlach, D. & Schurf, B. (2011): *Sprachwandel und Sprachvarietäten: Tendenzen der deutschen Gegenwartssprache. Kursthemen Deutsch*. Berlin: Cornelsen.

- European Commission (1995): White Paper on Education and Training – Teaching and Learning – Towards the Learning Society. COM (95) 590 final, 29 November 1995. URL: <https://op.europa.eu/en/publication-detail/-/publication/d0a8aa7a-5311-4eee-904c-98fa541108d8/language-en>, [abgerufen am 8. Oktober 2024].
- Fasold, R. W. (2001): Ebonic need not be English. In: Alatis, J. E. & Tan, A.-H. (Hrsg.): *Language in Our Time: Bilingual Education and Official English, Ebonics and Standard English, Immigration and the Unz Initiative*. Georgetown University Round Table on Languages and Linguistics 1999. Washington, D. C.: Georgetown University Press, 262–280.
- Fink, I. (2014): „Wimmerl is eigentlich wirklich a Mundortwort, oba ‚Pickel‘ is a net schöner.“. Ergebnisse einer Studie über das Korrekturverhalten von österreichischen Lehrenden. In: *Informationen zur Deutschdidaktik. Zeitschrift für den Deutschunterricht in Wissenschaft und Schule* 38.3, 78–87.
- Fink, I., Ransmayr, J. & de Cillia, R. (2017): »Also grammatisch würd ich fast sagen, dass die Österreicher inkorrekt sind, aber sonst eigentlich gar nicht.« Wahrnehmung von und Einstellungen gegenüber Varietäten des Deutschen bei österreichischen LehrerInnen und SchülerInnen. In: *ÖDaF-Mitteilungen* 33.1, 79–96.
- Fischer, H. (Hrsg.) (2023): *Sprachvariation@Schule*. Bearbeitet von Mareike Krause und Ella Wissenbach. Unter Mitarbeit von Lisa Dücker, Robert Engsterhold, Felix Klaner und Clemens Wissel. Forschungszentrum Deutscher Sprachatlas: Marburg. URL: <https://schule.dsa.info>, [abgerufen am 8. Oktober 2024].
- Fuchs, E. & Elspaß, S. (2019): *Innere und äußere Mehrsprachigkeit an österreichischen Schulen*. Electronic Article. URL: <https://bimm.at/wp-content/uploads/2019/09/fuchselspa2019innereundaueremehrsprachigkeitanosterreichischenschuleneinprojektberichtzuwahrnehmungenundeinstellungenteili.pdf>, [abgerufen am 8. Oktober 2024].
- Garbin, G., Sanjuan, A. & Forn, C. [et al.] (2010): Bridging language and attention: Brain basis of the impact of bilingualism on cognitive control. In: *NeuroImage* 53, 1272–1278. DOI: <https://doi.org/10.1016/j.neuroimage.2010.05.078>.
- Garrett, P. (2010): *Attitudes to Language. Key Topics in Sociolinguistics*. Cambridge [u. a.]: Cambridge University Press.
- Giles, H. & Johnson, P. (1987): Ethnolinguistic identity theory: a social psychological approach to language maintenance. In: *Int'l J. Soc. Lang.* 69, 69–99.
- Giles, H. & Powesland, P. F. (1975): *Speech Style and Social Evaluation. European Monographs in Social Psychology*. London/New York/San Francisco: Academic Press.
- Glantschnig, M. (2011): Möglichkeiten und Grenzen eines dialektorientierten Grammatikunterrichts. In: Köpcke, K.-M. & Ziegler, A. (Hrsg.): *Grammatik – Lehren, Lernen, Verstehen. Zugänge zur Grammatik des Gegenwartssprachen*. Berlin/Boston: De Gruyter (= Reihe Germanistische Linguistik), 223–244.
- Godley, A. J. & Reaser, J. (2018): *Critical Language Pedagogy: Interrogating Language, Dialects, and Power in Teacher Education. Social Justice Across Contexts in Education*. New York: Peter Lang.

- Grogger, J., Steinmayr, A. & Winter, J. (2020): The wage penalty of regional accents. In: *NBER WORKING PAPER SERIES* Working Paper, 1–36. URL: <http://www.nber.org/papers/w26719>, [abgerufen am 8. Oktober 2024].
- Gumperz, J. J. (1982): *Discourse Strategies. Studies in Interactional Sociolinguistics*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511611834>.
- Häcki Buhofer, A. & Studer, T. (1993): Zur Entwicklung von Sprachdifferenzbewusstsein und Einstellungen zu den Varianten des Deutschen in der Deutschen Schweiz. In: Werlen, I. (Hrsg.): *Schweizer Soziolinguistik: Soziolinguistik der Schweiz*. Neuchâtel: Université de Neuchâtel, Institut de linguistique, 179–199.
- Healy, K. (2019): *Data Visualization. A Practical Introduction*. Princeton: Princeton University Press.
- Holmes, J. (2013): *An Introduction to Sociolinguistics. Learning about Language*. Abingdon, OX/New York: Routledge.
- Hornsey, M. J. (2008): Social Identity Theory and Self-categorization Theory: A Historical Review. In: *Social and Personality Psychology Compass* 1, 204–222. DOI: <https://doi.org/10.1111/j.1751-9004.2007.00066.x>.
- Huesmann, A. (1998): *Zwischen Dialekt und Standard. Empirische Untersuchung zur Soziolinguistik des Varietätenspektrums im Deutschen*. Tübingen: Max Niemeyer Verlag (= Reihe Germanistische Linguistik).
- Irvine, J. T. & Gal, S. (2000): Language Ideology and Linguistic Differentiation. In: Kroskrity, P. V. (Hrsg.): *Regimes of Language: Ideologies, Politics, and Identities*. Santa Fe: School of American Research Press, 35–84.
- Jakobson, R. O. & Halle, M. (1960): *Grundlagen der Sprache*. Berlin: Akademie Verlag.
- Johnson, L., Terry, N. P. & McDonald Connor, C. [et al.] (2017): The effects of dialect awareness instruction on nonmainstream American English speakers. In: *Reading and Writing* 30, 2009–2038. DOI: <https://doi.org/10.1007/s11145-017-9764-y>.
- Kaiser, I. (2006): Bundesdeutsch aus österreichischer Sicht. *Eine Untersuchung zu Spracheinstellungen, Wahrnehmungen und Stereotypen*. amades. Arbeitspapiere und Materialien zur deutschen Sprache. Mannheim: Leibniz-Institut für Deutsche Sprache.
- Kaiser, I. & Ender, A. (2015): Das Spektrum der Sprachvariation im alemannischsprachigen Vorarlberg und im übrigen Österreich: Realisierungen und Kategorisierungen. In: Lenz, A. N., Ahlers, T. & Glauning, M. M. (Hrsg.): *Dimensionen des Deutschen in Österreich. Variation und Varietäten im sozialen Kontext*. Frankfurt a. M.: Peter Lang (= Schriften zur deutschen Sprache in Österreich), 11–30.
- Kaiser, I. & Ender, A. (2020): Innere Mehrsprachigkeit als Potenzial für den Deutschunterricht. In: Langlotz, M. (Hrsg.): *Grammatikdidaktik – theoretische und empirische Zugänge zu sprachlicher Heterogenität*. Baltmannsweiler: Schneider Verlag Hohengehren, 237–271.
- Kaiser, I. & Kasberger, G. (2021): Children's sociolinguistic preferences: the acquisition of language attitudes within the Austrian standard-dialect-continuum. In: Chevrot, J.-P., Ghimenton, A. & Nardy, A. (Hrsg.): *Sociolinguistic variation and language acquisition across the lifespan*. Amsterdam: John Benjamins, 130–160.

- Kasberger, G. & Kaiser, I. (2019): „I red normal“ – eine Untersuchung der varietätenspezifischen Sprachbewusstheit und -bewertung von österreichischen Kindern. In: Bülow, L., Fischer, A. K. & Herbert, K. (Hrsg.): *Dimensions of Linguistic Space: Variation – Multilingualism – Conceptualisations. Dimensionen sprachlichen Raums: Variation – Mehrsprachigkeit – Konzeptualisierung*. Bern: Peter Lang, 319–340.
- Kellermeier-Rehbein, B. (2017): Alte und neue Sprachbarrieren. In: *Der Deutschunterricht* 4, 16–27.
- Klassert, A., Kauschke, C. & Gagarina, N. (2009): Lexikalische Fähigkeiten bilingualer Kinder. In: Heide, J., Hanna, S. & Brandt, O.-C. [et al.] (Hrsg.): *Spektrum Patholinguistik. Ein Kopf – Zwei Sprachen. Mehrsprachigkeit in Forschung und Therapie*. Bd. 2. Potsdam: Universitätsverlag Potsdam, 113–119.
- Klayman, J. (1995): Varieties of confirmation bias. In: *Psychology of learning and motivation* 32, 385–418.
- Kliesch, M. & Pfenninger, S. (2021): Cognitive and Socioaffective Predictors of L2 Microdevelopment in Late Adulthood: A Longitudinal Intervention Study. In: *The Modern Language Journal* 105.1, 237–266.
- Koch, P. & Oesterreicher, W. (1985): Sprache der Nähe – Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. In: *Romanistisches Jahrbuch* 36, 15–43.
- Koch, P. & Oesterreicher, W. (1994): Funktionale Aspekte der Schriftkultur. Functional Aspects of Literacy. In: Günther, H. & Ludwig, O. (Hrsg.): *Handbuch Schrift und Schriftlichkeit*. Band I. Berlin: De Gruyter, 587–604.
- Krafl, G. (2020): *Spracheinstellungen von Eltern zu Dialekt und Standardsprache*. Thesis.
- Krumpal, I. (2013): Determinants of social desirability bias in sensitive surveys: A literature review. In: *Quality & Quantity: International Journal of Methodology* 47.4, 2025–2047.
- Labov, W. (1969): *A Study of Nonstandard English*. Champaign, Ill.: National Council of Teachers of English.
- Labov, W. (2006): *The Social Stratification of English in New York City*. Cambridge [u. a.]: Cambridge University Press.
- Lambert, W. E., Hodgson, R. C. & Gardner, R. C. [et al.] (1960): Evaluational reactions to spoken languages. In: *Journal of Abnormal and Social Psychology* 60.1, 44–51.
- Lenz, A. N. (2014): Sprachvariation und Sprachwandel aus der Perspektive von Deutschlehrerinnen und Deutschlehrern. Einstellungsdaten aus Österreich, Deutschland und der Schweiz. In: Plewnia, A. & Witt, A. (Hrsg.): *Sprachverfall? Dynamik – Wandel – Variation*. Institut für deutsche Sprache. Jahrbuch 2013. Berlin/Boston: De Gruyter, 323–352.
- Lenz, A. N., Soukup, B. & Koppensteiner, W. (2022): Standard German in Austria from the folk perspective: Conceptualizations, attitudes, perceptions. In: Lenz, A. N., Soukup, B. & Koppensteiner, W.: *Standard Languages in Germanic-speaking Europe: Attitudes and Perception*. Oslo: Novus (SLICE Book Series 4), 23–57. URL: <https://omp.novus.no/index.php/novus/catalog/view/20/30/1063>, [abgerufen am 29. Mai 2024].
- Levshina, N. (2015): *How to do Linguistics with R. Data Exploration and Statistical Analysis*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

- Llamas, C., Mullany, L. & Stockwell, P. (2006): *The Routledge Companion to Sociolinguistics*. Taylor & Francis e-Library.
- Löbner, S. (2015): *Semantik. Eine Einführung*. Berlin/Boston: De Gruyter.
- Löffler, H. (1972): Mundart als Sprachbarriere. In: Pädagogischer Verlag Schwann (Hrsg.): *Wirkendes Wort. Deutsche Sprache in Wort und Lehre*. Düsseldorf: Pädagogischer Verlag Schwann, 23–39.
- Maitz, P. (2015): Sprachvariation, sprachliche Ideologien und Schule. In: *Zeitschrift für Dialektologie und Linguistik* 82.2, 206–227.
- Maitz, P. & Elspaß, S. (2011): Dialektfreies Sprechen – leicht gemacht! Sprachliche Diskriminierung von deutschen Muttersprachlern in Deutschland. In: *Der Deutschunterricht* 6, 7–17.
- Mallinson, C. & Charity Hudley, A. H. (2010): Communicating about Communication: Multidisciplinary Approaches to Educating Educators about Language Variation. In: *Language and Linguistics Compass* 4.4, 245–257.
- Mattheier, K. (1990): Dialekt und Standardsprache. Über das Varietätensystem des Deutschen in der Bundesrepublik. In: *International Journal of the Sociology of Language* 83, 59–81.
- Mattheier, K. (1991): Standardsprache als Sozialsymbol. In: Wimmer, R. (Hrsg.): *Das 19. Jahrhundert: Sprachgeschichtliche Wurzeln des heutigen Deutsch*. Berlin: De Gruyter, 41–72.
- Mattheier, K. (2000): Die Durchsetzung der deutschen Hochsprache im 19. und beginnenden 20. Jahrhundert: sprachgeographisch, sprachsoziologisch. In: Besch, W., Betten, A. & Reichmann, O. [et al.] (Hrsg.): *Sprachgeschichte. Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung*. 2. Teilband. Berlin: De Gruyter (= Handbücher zur Sprach- und Kommunikationswissenschaft), 1951–1966.
- Mesthrie, R., Swann, J. & Deumert, A. [et al.] (2009): *Introducing Sociolinguistics*. Edinburgh: Edinburgh University Press.
- Meyerhoff, M. (2006): *Introducing Sociolinguistics*. Abingdon, OX/New York: Routledge.
- Milroy, L. (1989): *Language and social networks. Language in Society*. Oxford: Basil Blackwell.
- Moosmüller, S. (1990): Einschätzung von Sprachvarietäten in Österreich. In: *International Journal of the Sociology of Language* 83, 105–120.
- Mufwene, S. S. (2001): Ebonics and Standard English in the Classroom: Some Issues. In: Alatis, J. E. & Tan, A.-H. (Hrsg.): *Language in Our Time: Bilingual Education and Official English, Ebonics and Standard English, Immigration and the Unz Initiative*. Georgetown University Round Table on Languages and Linguistics 1999. Washington, D. C.: Georgetown University Press, 253–261.
- Murray, D. E. (2001): Whose „standard“? What the Ebonics tells us about language, power and pedagogy. In: Alatis, J. E. & Tan, A.-H. (Hrsg.): *Language in Our Time: Bilingual Education and Official English, Ebonics and Standard English, Immigration and the Unz Initiative*. Georgetown University Round Table on Languages and Linguistics 1999. Washington, D. C.: Georgetown University Press, 281–291.

- Neuland, E. (1993): Sprachbewußtsein und Sprachvariation. Zur Entwicklung und Förderung eines Sprachdifferenzbewußtseins. In: Klotz, P. & Sieber, P. (Hrsg.): *Vielerlei Deutsch. Umgang mit Sprachvarietäten in der Schule*. Stuttgart: Ernst Klett Schulbuchverlag, 173–191.
- Noack, C. (2003): „Usem“ – „ausm“ – „aus dem“: Schriftsprachliche Probleme von Dialekt- und Hochdeutschsprechern. Ergebnisse eines Projekts. In: *Alfa-Forum* 53, 17–19.
- Oskolkov, N. (2020a): *How Linear Mixed Model Works. And how to understand LMM through Bayesian lenses*. Electronic Article: URL: <https://towardsdatascience.com/how-linear-mixed-model-works-350950a82911>, [abgerufen am 29. Mai 2024].
- Oskolkov, N. (2020b): *Linear Mixed Model from Scratch*. Electronic Article: URL: <https://towardsdatascience.com/how-linear-mixed-model-works-350950a82911>, [abgerufen am 29. Mai 2024].
- Papapavlou, A. & Phili, A. (2009): Childhood Bidialectism and Metalinguistic Awareness. In: *Selected Papers on Theoretical and Applied Linguistics* 18 (= Selected Papers from the 18th International Symposium on Theoretical and Applied Linguistics, Thessaloniki 4–6 May 2007), 311–320.
- Paulus, C. (2009): Die „Bücheraufgabe“ zur Bestimmung des kulturellen Kapitals bei Grundschulern. DOI: <https://doi.org/10.23668/psycharchives.10915>.
- Peter, K. (2015): Sprachliche Normvorstellungen in Österreich, Deutschland und der Schweiz. In: Lenz, A. N., Ahlers, T. & Glauninger, M. M. (Hrsg.): *Dimensionen des Deutschen in Österreich. Variation und Varietäten im sozialen Kontext*. Schriften zur deutschen Sprache in Österreich. Frankfurt a. M.: Peter Lang (= Schriften zur deutschen Sprache in Österreich), 123–147.
- Pfrehm, J. W. (2007): *An empirical study of the pluricentricity of German: Comparing German and Austrian nationals' perceptions of the use, pleasantness, and standardness of Austrian standard and German standard lexical items*. Thesis.
- Poarch, G. J., Vanhove, J. & Berthele, R. (2019): The effect of bidialectalism on executive function. In: *International Journal of Bilingualism* 23.2, 612–628.
- Preston, D. R. (2006): Language with an Attitude. In: Chambers, J. K., Trudgill, P. & Schilling-Estes, N. (Hrsg.): *The Handbook of Language Variation and Change*. Malden, MA/Oxford/Carlton, Victoria: Blackwell (= Blackwell Handbooks in Linguistics), 40–66.
- Preston, D. R. (2010): Language, People, Salience, Space: Perceptual Dialectology and Language Regard. In: *Dialectologia* 5, 87–131.
- Preston, D. R. (2017): The cognitive foundations of language regard. In: *Poznań Studies in Contemporary Linguistics* 53.1, 17–42.
- Preston, D. R. (2019): Folk Linguistics and the Perception of Language Variety. In: Antos, G., Niehr, T. & Spitzmüller, J. (Hrsg.): *Handbuch Sprache im Urteil der Öffentlichkeit*. Berlin/Boston: De Gruyter (= Handbücher Sprachwissen), 140–164.
- Pribyl-Resch, C. (2021): *Dialektsprechen und Bilingualismus – Potenzial oder Hindernis? Einschätzungen von Lehrpersonen*. Thesis.
- R-Core-Team (2017): *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. URL: <https://www.r-project.org>, [abgerufen am 8. Oktober 2024].

- Rastner, E. M. (1997): Sprachvarietäten im Unterricht. Eine Umfrage unter Österreichs LehrerInnen zu Standardsprache – Umgangssprache – Dialekt. In: *Fremde Muttersprachen Deutsch. Informationen zur Deutschdidaktik. Zeitschrift für den Deutschunterricht in Wissenschaft und Schule* 21.3, 80–93.
- Reiffenstein, I. (2020): Aspekte einer Sprachgeschichte des Bayerisch-Österreichischen bis zum Beginn der frühen Neuzeit. In: Wiegand, H. (Hrsg.): *Handbücher zur Sprach- und Kommunikationswissenschaft* 2.3. Berlin/Boston: De Gruyter, 2889–2942.
- Rosch, E. (1978): Principles of Categorization. In: Rosch, E. & Lloyd, B. B. (Hrsg.): *Cognition and Categorization*. Hillsdale NJ: Lawrence Erlbaum, 27–48.
- Rutter Strickling, L. (2012): „This Whole Language Variation Journey“: *Examining Teachers' Changing Beliefs about Language and Linguistically Aware Pedagogical Applications*. Thesis.
- Saxalber, A. & Lanthaler, F. (1993): Deutschunterricht in Südtirol – Sprachdidaktik für den kleinen Raum. In: Klotz, P. & Sieber, P. (Hrsg.): *Vielerlei Deutsch. Umgang mit Sprachvarietäten in der Schule*. Stuttgart: Ernst Klett Schulbuchverlag, 70–83.
- Scheutz, H. (2009a): Der Dialekt: Sprache der Nähe, Ausdruck regionaler Identität – oder defizitäres System? In: Scheutz, H. (Hrsg.): *Drent und Herent. Dialekte im Salzburgerisch-Bayerischen Grenzgebiet*. Salzburg [u. a.]: EuRegio, 7–12.
- Scheutz, H. (2009b): Lautliche Dialekteigenschaften und ihre geographische Verteilung. In: Scheutz, H. (Hrsg.): *Drent und Herent. Dialekte im Salzburgerisch-Bayerischen Grenzgebiet*. Salzburg [u. a.]: EuRegio, 21–56.
- Schoel, C. & Stahlberg, D. (2012): Spracheinstellungen aus sozialpsychologischer Perspektive II: Dialekte. In: Eichinger, L. M., Plewnia, A. & Schoel, C. [et al.] (Hrsg.): *Sprache und Einstellungen. Spracheinstellungen aus sprachwissenschaftlicher und sozialpsychologischer Perspektive*. Tübingen: Narr, 205–225.
- Schweiger, H. (2014): Österreichisches Deutsch in der Literatur. Überlegungen zu einem sprach- und kulturreflexiven Unterricht. In: *Informationen zur Deutschdidaktik* 3, 105–111.
- Siegel, J. (2006): Language ideologies and the education of speakers of marginalized language varieties: Adopting a critical awareness approach. In: *Linguistics and Education* 17.2, 157–174.
- Smith, P., Kumi-Yeboah, A. & Chang, R. [et al.] (2019): Rethinking „(Under)Performance“ for Black English Speakers: Beyond Achievement to Opportunity. In: *Journal of Black Studies*, 1–27.
- Sonderegger, M. (2021): *Regression Modeling for Linguistic Data*. Web Page. DOI: <https://doi.org/10.17605/OSF.IO/PNUMG>.
- Soukup, B. (2009): *Dialect use as interaction strategy: a sociolinguistic study of contextualization, speech perception, and language attitudes in Austria*. Wien: Braumüller.
- Soukup, B. (2015): Zum Phänomen ‚Speaker Design‘ im österreichischen Deutsch. In: Lenz, A. N. & Glauninger, M. M. (Hrsg.): *Standarddeutsch im 21. Jahrhundert – Theoretische und empirische Ansätze mit einem Fokus auf Österreich*. Göttingen: V&R unipress, 59–79.

- Soukup, B. (2019): Sprachreflexion und Kognition: Theorien und Methoden der Sprach-einstellungsforschung. In: Antos, G., Niehr, T. & Spitzmüller, J. (Hrsg.): *Handbuch Sprache im Urteil der Öffentlichkeit*. Berlin/München/Boston: De Gruyter, 83–106. DOI: <https://doi.org/10.1515/9783110296150-005>.
- Spitzmüller, J. (2022): *Soziolinguistik. Eine Einführung*. Berlin: Metzler.
- Staatsinstitut für Schulqualität und Bildungsforschung (ISB): *Sprachliche Kultur und Vielfalt in Bayern*. Webseite. URL: <https://www.dialekte.schule.bayern.de>, [abgerufen am 14. August 2024].
- Steinegger, G. (1998): *Sprachgebrauch und Sprachbeurteilung in Österreich und Südtirol. Ergebnisse einer Umfrage*. Frankfurt a. M. [u. a.]: Peter Lang (= Schriften zur deutschen Sprache in Österreich).
- Stroebe, J., Stroebe, W. & Hewstone, M. (2007): *Sozialpsychologie. Eine Einführung*. Heidelberg: Springer.
- Stroebe, W. (2014): Strategien zur Einstellungs- und Verhaltensänderung. In: Klaus, J., Stroebe, W. & Hewstone, M. (Hrsg.): *Sozialpsychologie. Eine Einführung*. Berlin Heidelberg: Springer, Kap. 7.
- Tajfel, H. & Turner, J. (1979): An Integrative Theory of Intergroup Conflict. In: Austin, W. G. & Worchel, S. (Hrsg.): *The social psychology of intergroup relations*. Monterey, CA: Brooks-Cole, 33–47.
- Tegegner, W. (2015): The Use of Dialects in Education and Its Impacts on Students' Learning and Achievement. In: *Education Journal* 4.5, 263–269. DOI: <https://doi.org/10.11648/j.edu.20150405.22>.
- Tophinke, D. (2019): *Praxis Deutsch 275. Dialekt heute*. Seelze: Friedrich Verlag.
- Trudgill, P. (1972): Sex, covert prestige and linguistic change in the urban British English of Norwich. In: *Language in Society* 1, 179–195.
- Tufte, E. (1983): The Visual Display of Quantitative Information. In: Schneider B. Ernst, C & Wöpping, J. (Hrsg.): *Diagrammatik-Reader: Grundlegende Texte aus Theorie und Geschichte*. Berlin, Boston: De Gruyter, (2016), 219–230. DOI: <https://doi.org/10.1515/9783050093833-035>.
- Turner, J. C., Hogg, M. A. & Oakes, P. J. [et al.] (1987): *Rediscovering the social group: A self-categorization theory*. New York: Basil Blackwell.
- Turner, J. C. & Reynolds, K. J. (2012): Self-Categorization Theory. In: Van Lange, P. A. M., Kruglanski, A. W. & Higgins, T. E. (Hrsg.): *Handbook of Theories of Social Psychology*. London: SAGE Publications Ltd, 399–417. DOI: <http://dx.doi.org/10.4135/9781446249222.n46>.
- Unterberger, E. & Pribyl-Resch, C. (im Druck): Sprachverwendung in der Schule. Ein Schulstufen- und Ländervergleich zwischen Österreich und Bayern. In: Hochstadt, C. & Schweigkofler Kuhn, A. (Hrsg.): *Dialekt in der Lehre – dialektologische, sprachdidaktische und interdisziplinäre Perspektiven*.
- Vangsnes, Ø. A., Söderlund, G. B. W. & Bleckesaune, M. (2017): The effect of bidialectal literacy on school achievement. In: *International Journal of Bilingual Education and Bilingualism* 20.3, 346–361.

- Vergeiner, P. C., Buchner, E. & Fuchs, E. [et al.] (2019): Sprachnormvorstellungen in sekundären und tertiären Bildungseinrichtungen in Österreich. In: *Zeitschrift für Dialektologie und Linguistik* 86.3, 284–330. DOI: <https://doi.org/10.25162/ZDL-2019-0011>.
- Wandruszka, M. (1975): Mehrsprachigkeit: Öffentlicher Vortrag. In: Moser, H. (Hrsg.): *Sprachwissenschaft und Sprachdidaktik: Jahrbuch 1974 des Instituts für deutsche Sprache. Sprache der Gegenwart*. Düsseldorf: Schwann, 321–350.
- Wardhaugh, R. & Fuller, J. M. (2015): *An Introduction to Sociolinguistics*. Blackwell Textbooks in Linguistics. Chichester, West Sussex: John Wiley & Sons.
- Wheeler, R. S. & Swords, R. (2004): Codeswitching: Tools of Language and Culture Transform the Dialectally Diverse Classroom. In: *Language Arts* 81.6, 470–480. URL: <https://cdn.ncte.org/nctefiles/pd/consulting/wheeler1ajuly2004.pdf>, [abgerufen am 8. Oktober 2024].
- Wiese, H., Mayr, K., Krämer, P., Seeger, P., Müller, H.-G. & Freywald, U.: Deutsch ist vielseitig: Aus- und Fortbildungsmodule zur Sprachvariation im urbanen Raum. In: Wiese, H., Mayr, K., Krämer, P., Seeger, P., Müller, H.-G. & Freywald, U.: *Deutsch ist vielseitig: Aus- und Fortbildungsmodule zur Sprachvariation im urbanen Raum*. URL: <https://www.linguistik.hu-berlin.de/de/institut/professuren/multilinguale-kontexte/schulen/deutsch-ist-vielseitig>, [aufgerufen am 14. August 2024].
- Wiese, H. (2006): „Ich mach dich Messer“: Grammatische Produktivität in KiezSprache („Kanak Sprak“). In: *Linguistische Berichte* 207, 245–273.
- Wiese, H. (2015): „This migrants’ babble is not a German dialect!“: The interaction of standard language ideology and ‘us’/‘them’ dichotomies in the public discourse on a multiethnolect. In: *Language in Society* 44, 341–368.
- Wiesinger, P. (1992): Zur Interaktion von Dialekt und Standardsprache in Österreich. In: van Leuvensteijn, J. A. & Berns, J. B. (Hrsg.): *Dialect and Standard Language in the English, Dutch, German and Norwegian Language Areas. Seventeen Studies in English or German*. Verhandelingen der Koninklijke Nederlandse Akademie van Wetenschappen, Afd. Letterkunde; Nieuwe reeks, deel 150. Amsterdam: Royal Netherlands Academy of Arts/Sciences, 290–311.
- Wiesinger, P. (2010): Deutsch in Österreich: Standard, regionale und dialektale Variation. In: Krumm, H.-J., Fandrych, C. & Hufeisen, B. [et al.] (Hrsg.): *Deutsch als Fremd- und Zweitsprache. Ein internationales Handbuch*. Band 1. Berlin: De Gruyter (= Handbücher zur Sprach- und Kommunikationswissenschaft), 360–372.
- Wiesinger, P. (2014a): Die sprachsoziologischen Verhältnisse in Österreich. Ergebnisse einer Umfrage. In: Wiesinger, P. (Hrsg.): *Das österreichische Deutsch in Gegenwart und Geschichte*. Wien: Lit-Verlag, 87–99.
- Wiesinger, P. (2014b): Zur Interaktion von Dialekt und Standardsprache in Österreich. In: Wiesinger, P. (Hrsg.): *Das österreichische Deutsch in Gegenwart und Geschichte*. Wien: Lit-Verlag, 101–112.
- Winter, B. (2020): *Statistics for Linguists: An Introduction Using R*. New York: Routledge.
- Wolfram, W. & Friday, W. C. (1997): The Role of Dialect Differences in Cross-cultural Communication: Proactive Dialect Awareness. In: *Bulletin suisse de linguistique appliquée* 65, 143–154.

- Wölke, A. (2018): *Sprachvarietäten und die Sapir-Whorf-Hypothese in der Diskussion: Gymnasiale Oberstufe*. EinFach Deutsch Unterrichtsmodelle. Braunschweig: Schöningh Verlag im Westermann Schulbuch.
- Yiakoumetti, A. (2007): Choice of classroom language in bidialectal communities: to include or to exclude the dialect? In: *Cambridge Journal of Education* 37.1, 51–66. DOI: <https://doi.org/10.1080/03057640601179046>.
- Yiakoumetti, A. (2012): The Danger of Dialects: Debunking (or Substantiating) the Myths. In: Yiakoumetti, A. (Hrsg.): *Harnessing Linguistic Variation to Improve Education. Rethinking Education*. Oxford [u. a.]: Peter Lang, 293–310. DOI: <https://doi.org/10.3726/978-3-0353-0312-4>.
- Yzerbyt, V., Provost, V. & Corneille, O. (2005): Not Competent but Warm ... Really? Compensatory Stereotypes in the French-speaking World. In: *Group Processes & Intergroup Relations* 8.3, 291–308.
- Zehetner, L. (1977): *Bairisch. Dialekt/Hochsprache – kontrastiv*. Düsseldorf: Schwann.
- Zehetner, L. (1980): Der Lehrer im Spannungsfeld zwischen Hochsprache und Dialekt. In: *Schulreport: Tatsachen u. Meinungen zur aktuellen Bildungspolitik in Bayern* 5, 7–9.
- Ziegler, E. (1999): Deutsch im 19. Jahrhundert. Normierungsprinzipien und Spracheinstellungen. In: Bister-Broosen, H. (Hrsg.): *Beiträge zur historischen Stadtsprachenforschung. Schriften zur diachronen Sprachwissenschaft*. Wien: Praesens, 79–101.



Welche Bedeutung hat sprachliche Variation im Schulalltag? Wie beeinflussen Spracheinstellungen das Miteinander in der Schule? Werden Personen mit Dialekt anders wahrgenommen als Standardsprechende? Diesen Fragen widmet sich der Autor in seiner Dissertation.

Österreich besitzt eine ausgeprägte und über weite Teile lebendige Dialektlandschaft. Der überwiegende Teil der Bevölkerung verwendet im Alltag Dialekt oder eine dialektal gefärbte Umgangssprache. Vor diesem Hintergrund analysiert der Autor die Einstellungen gegenüber Dialekt- und Standardsprechenden.

Im Rahmen des Projekts „Sprachliche Vielfalt verstehen, wertschätzen und ausbauen“ wurden Materialien für den Deutschunterricht entwickelt und deren Wirkung in sechs Partnerschulen getestet. Eine Untersuchung der Spracheinstellungen zeigte etwa geschlechtsspezifische sowie regionale Unterschiede in der Wahrnehmung von Dialekt- und Standardsprechenden. Die Interventionsmaßnahmen führten zu einer signifikanten Verbesserung der Kompetenzeinschätzung von Dialektsprechenden.

Ziel der Publikation ist es, Sprachvariation als wesentlichen Gegenstand des Deutschunterrichts zu betrachten und Möglichkeiten der bewussten Bearbeitung aufzuzeigen.

Ein Schneider Verlag-Titel bei wbv Publikation



**wbv.de**