# WIE SPRECHEN (SCHWEIZER) DIALEKTSPRECHER*INNEN HOCHDEUTSCH?

## INTONATIONSVERGLEICH DER DEUTSCHEN STANDARDVARIETÄTEN IN MÜNCHEN UND ZÜRICH

Phonetik einer Einzelsprache: Schweizerdeutsch, 24. Januar 2022

Ein Vortrag von
Jasmin Rimpler, Institut für Phonetik und Sprachverarbeitung, LMU München

# REMINDER: VERHÄLTNIS VON DIALEKT & STANDARDSPRACHE
(VGL. ULBRICH, 2005)

## BUNDESREPUBLIK DEUTSCHLAND

- Große Bandbreite an Dialekten (Ober-, Mittel- und Niederdeutsch)
- Mitte & Norden: öfter Dialektvermeidung wegen sozialer Markiertheit, Süden: größere Dialektakzeptanz
- Unterteilung in nord- (NSG) und süddeutschen Standard

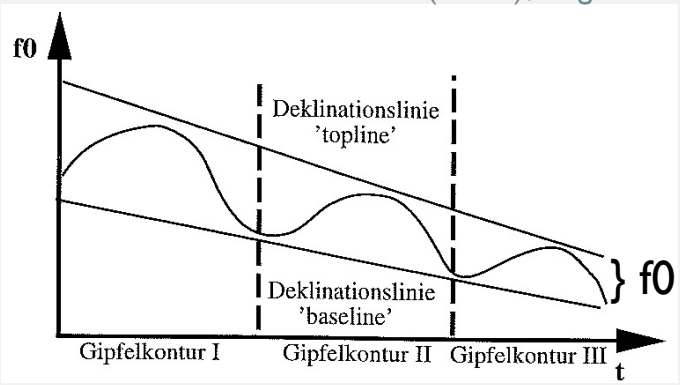
## DEUTSCHSPRACHIGE SCHWEIZ

- Überwiegend alemannischer Dialektraum
- Kein Dialektkontinuum, sondern Diglossie: Standard & Dialekt existieren gleichberechtigt nebeneinander (Gebrauch kontextbedingt)
- Verschiedene Standardregister

# HEUTE BENÖTIGTE BEGRIFFE
## (Z.B. PETERS, 2014)

- Jeder Äußerungsabschnitt, der eine Intonationskontur trägt, bildet eine **Intonationsphrase (IP)**. → s. Phrasierung
- Ein **nuklearer Akzent** ist der einzige obligatorische Akzent einer IP (es gibt folglich nur einen pro IP). In der Regel ist der nukleare Akzent der letzte Tonhöhenakzent einer IP.
- **Pränukleare Akzente** sind alle Akzente einer IP, die vor dem nuklearen Akzent auftreten.
- Der **Tonhöhenumfang (f0 range)** ergibt sich aus der Differenz des f0 Maximums und des f0 Minimums.

Abb. aus Peters & Kohler (2004), ergänzt

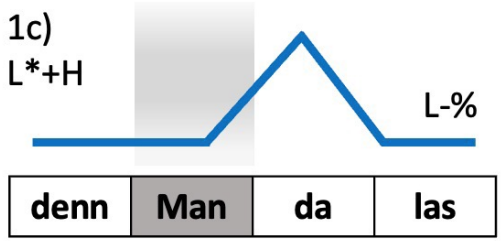


# INTONATION IM SCHWEIZERDEUTSCHEN I. DIALEKT (BERNDEUTSCH)

- **Fitzpatrick-Cole (1999)** → s. Referat
  - Pränuklearer Akzentton: L*+H (gefolgt von L*+H)
  - Nuklearer Akzentton: L*+H (gefolgt von Grenzton L%)

→ Tiefklingende akzentuierte Silbe L* gefolgt von spätem Hochton +H (Gipfel oft nach betonter Silbe)

→ Abb. s. letzte VL




→ Vgl. **Leemann (2012)**: Späte Gipfel als DAS prototypische Merkmal von Berndeutsch

Alternative Repräsentation (nur sinnvoll im Vgl. zu NSG):

- H*+L als zugrundeliegender Akzentton
- Phonetische Implementierung (= Realisierung/Äußerung durch Berner Sprecher*innen): (sehr) spät alignierter Gipfel

→ Hier kein Bedeutungsunterschied durch spätere Alignierung! ≠ Ironie / rhetorische Fragen

# INTONATION IM SCHWEIZERDEUTSCHEN I. DIALEKT (ZÜRICHDEUTSCH)

- **Fleischer & Schmid (2006)**
  - Größerer Tonhöhenumfang und größere Anzahl an Tonhöhenbewegungen

Abb. ergänzt

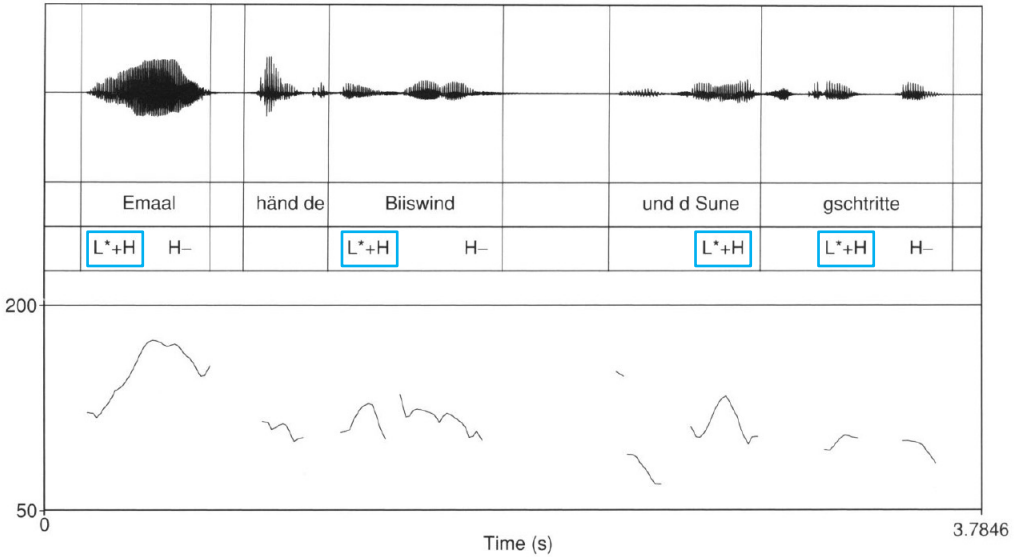



**Figure 1** Low-rising pitch accents (from sentence 1 of the recorded passage).

# INTONATION IM SCHWEIZERDEUTSCHEN II. SCHWEIZERHOCHDEUTSCH
(VGL. ULBRICH, 2005; FLEISCHER & SCHMID, 2006)

- Schweizer Standardvarietät weist im Vergleich zur in Deutschland gesprochenen Standardsprache u.a. phonetische Besonderheiten auf → dialektaler Einfluss

→ Eigenständige und gleichberechtigte Standardsprache

- **Ulbrich (2005)**
  Vergleich der Standardvarietäten in D, AT und CH
  - 28 Nachrichtensprecher*innen lesen Nachrichten & Märchen
  - Auditive und akustische Analyse globaler und lokaler (= Akzentrealisierung) prosodischer Merkmale

# INTONATION IM SCHWEIZERDEUTSCHEN II. SCHWEIZERHOCHDEUTSCH

- **Ulbrich (2005)** → Ergebnisse (CH vs. D)
  - Anzahl betonter Silben pro Äußerung: CH > D
  - Schweizer Sprecher*innen zeigen größeren f0-Umfang im Bereich betonter Silben.
  - Verteilung der Betonungsmuster:
    - D v.a. HL und LH_L
    - CH v.a. LH und LH_L → überwiegend tiefe Akzentsilben

# INTONATION IM SCHWEIZERDEUTSCHEN II. SCHWEIZERHOCHDEUTSCH

- **Ulbrich (2005)** → Ergebnisse (CH vs. D) Nuklearakzente mit…

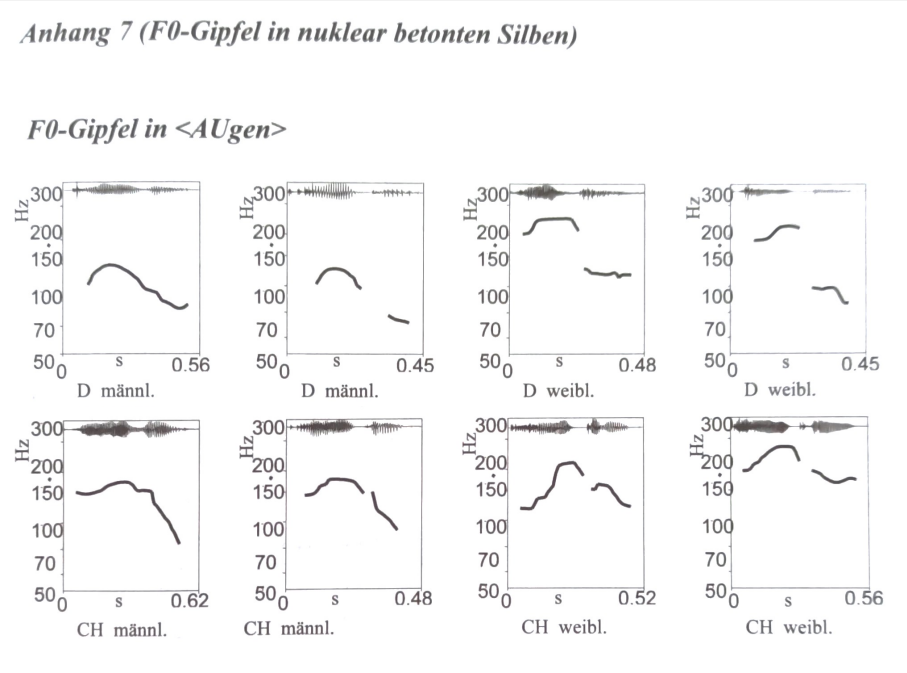
  …**einer unbetonter Silbe** im postnuklearen Verlauf, z.B. *Was hast du für große AUgen/HÄNde?*

- Schweizer und deutsche Sprecher*innen realisieren f0-Gipfel im Bereich des betonten Vokals
- Vielfältige Variation der Gipfelanbindung (= Alignment)

  …**mehreren unbetonten Silben** im post-nuklearen Verlauf, z.B. *…FRESsen/PAKken kann.*

- Kurzvokale in betonter Silbe → Schweizer*innen realisieren f0-Gipfel öfter in erster postnuklearer Silbe
- Einige produzieren f0-Anstieg auf *kann*

Abb. gekürzt




→ NB: Kein systematischer Vergleich der Standardvarietäten bzgl. Alignment!

# REGIONALE ALIGNMENT-UNTERSCHIEDE

- Dialektaler Hintergrund, Akzenttyp und Silbenstruktur (u.a.) beeinflussen in Standardvarietäten des Deutschen die phonetische Alignierung
- Sprecher*innen des süddeutschen Standards alignieren f0-Gipfel (H) von steigenden Akzenttönen in nicht-kontrastiven, pränuklearen Akzenten später als NSG Sprecher*innen
  (→ Atterer & Ladd, 2004: Bayern vs. Nord-West Deutschland; Mücke et al., 2008: Wien vs. Düsseldorf → Abb.)

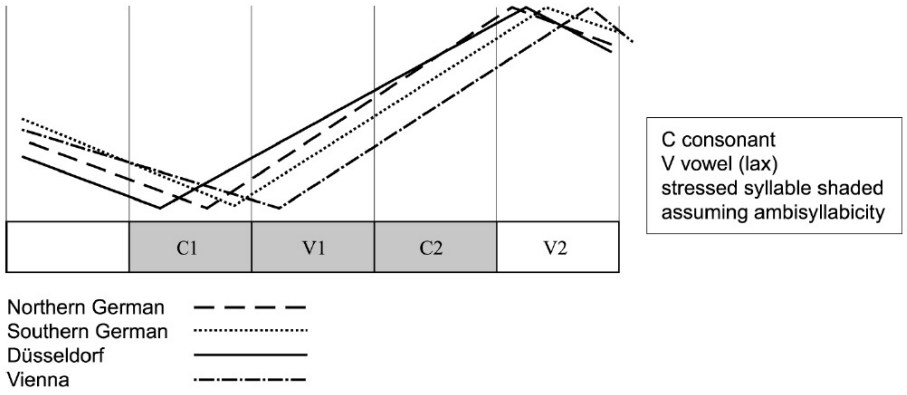



Figure 4: *Summary of alignment properties of prenuclear LH accents. Illustration of German varieties adapted from Atterer & Ladd 2004; Düsseldorf and Vienna added.*

*Die **MINN**esänger von Nürnberg waren sehr berühmt.*

| | Nicht-kontrastive Akzente | Kontrastive Akzente |
|---|---|---|
| Pränukleare Position | ✓ | |
| Nukleare Position | | |

# REGIONALE ALIGNMENT-UNTERSCHIEDE

| | Nicht-kontrastive Akzente | Kontrastive Akzente |
|---|---|---|
| Pränukleare Position | ✓ | ✗ |
| Nukleare Position | ? | ✗ |

- In kontrastiven Kontexten (s. u.) gibt es weder in pränuklearen noch in nuklearen Akzenten regionale Alignmentunterschiede (Braun, 2007: München vs. Münster; Mücke et al., 2009: Wien vs. Düsseldorf → Abb.)

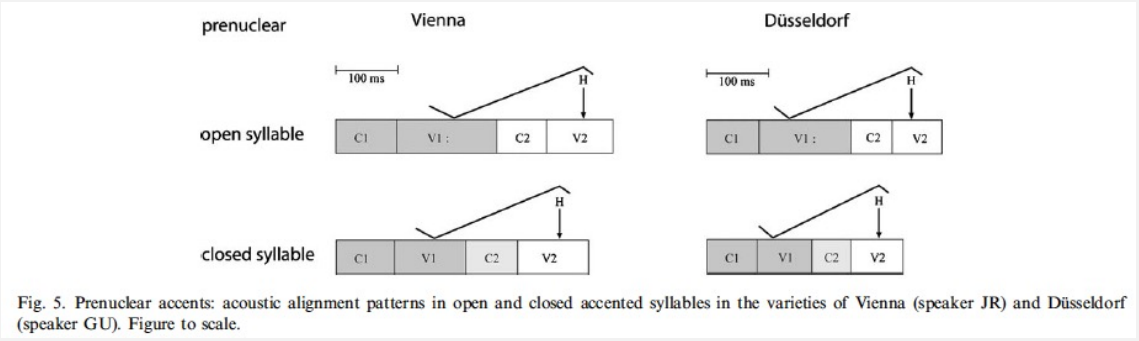



Fig. 5. Prenuclear accents: acoustic alignment patterns in open and closed accented syllables in the varieties of Vienna (speaker JR) and Düsseldorf (speaker GU). Figure to scale.

*Mit der **NAN**ni nahm sie den SECHS-Uhr-Bus.*
*Mit der **NAH**ni nahm sie den EIN-Uhr-Bus.*

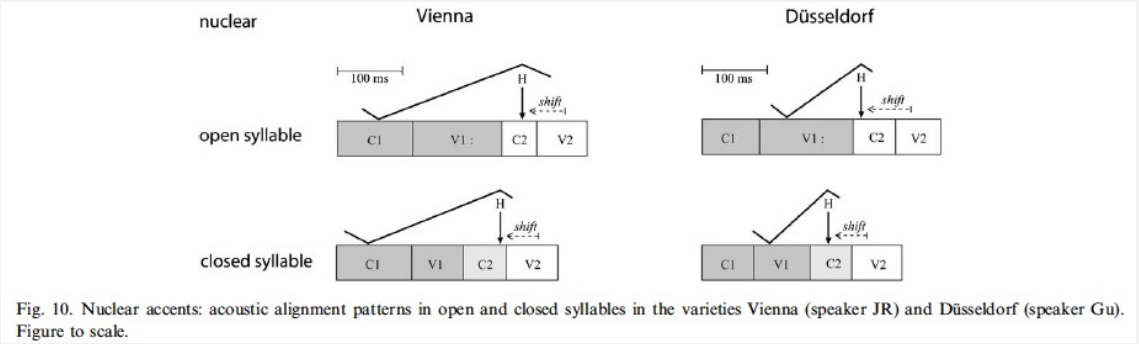



Fig. 10. Nuclear accents: acoustic alignment patterns in open and closed syllables in the varieties Vienna (speaker JR) and Düsseldorf (speaker Gu). Figure to scale.

F: *Hat sie die NANni oder die NAHni bestohlen?*
A: *Sie hat die **NAH**ni bestohlen.*

- Spätere Gipfelalignierung in geschlossenen Silben mit Kurzvokalen als in offenen Silben mit Langvokalen

- Keine dieser Studien untersucht nicht-kontrastive Akzente in nuklearer Position

# FORSCHUNGSLÜCKE + HAUPTZIEL

- Untersuchung möglicher regionaler Unterschiede in nicht-kontrastiven Nuklearakzenten zwischen deutschen Standardsprecher*innen aus München (MSG) und Zürich (ZSG)

*Was ist neu?*

- Fokus auf nicht-kontrastive Nuklearakzente
- Vergleich zweier spätalignierender Standardvarietäten des Deutschen (gesprochen in einer mittelbairischen und einer alemannischen Dialektregion)

# HYPOTHESEN

I. ZSG Sprecher*innen alignieren Gipfel später als MSG Sprecher*innen (H1)
II. Spätere Gipfelalignierung in geschlossenen Silben mit Kurzvokalen als in offenen Silben mit Langvokalen (H2)

# METHODE

## VERSUCHSPERSONEN

- 12 MSG Sprecher*innen
  - 6 weiblich + 6 männlich
  - Alter: ⌀23.8 J. (19-30 J.)
- 10 ZSG Sprecher*innen
  - 5 weiblich + 5 männlich
  - Alter: ⌀23 J. (18-29 J.)

## MATERIALIEN

- 4 Zielwörter → 2 Minimalpaare im Hinblick auf Silbenstruktur
  *Höhle-Hölle, Diener-Dinner*
- Zweisilbige Trochäen (Wortstruktur: $C_1\mathbf{V_1}\mathbf{S_2}V_2$), eingebettet in Trägersätze
  z.B. *Er hat die Höhle erforscht.*
- Zielsequenz:
  akzentuierter Vokal (V1)
  + nachfolgender Sonorant
  (S2: Lateral oder Nasal)

Analyse der Grundfrequenz (f0), also der Anzahl der Stimmlippenschwingungen / Sekunde
→ **Stimmhaftigkeit!**

# METHODE
## AUFNAHME, SEGMENTIERUNG UND LABELLING

- Akustische Aufnahme mit SpeechRecorder (Draxler & Jänsch, 2004), 5 Wdh. mit normaler Sprechgeschwindigkeit
- Automatische Segmentierung der Trägersätze mit MAUS (Kisler et al., 2016)
- Händische Korrekturen der MAUS-Segmente und **Markierung der f0-Gipfel (H)** in EMU SDMS (Winkelmann et al., 2017; Jochim, 2017)

# METHODE
## VORBEREITUNG DES DATENSATZES

1) Ausschluss aller Aufnahmen…

…ohne f0-Anstieg (auf Zielsequenz / von vorangehender Silbe zur Zielsilbe)
…mit Zielwörtern ohne Nuklearakzent
…mit früher Gipfelrealisierung
…mit steigendem Grenzton (H%)
…mit technischen Fehlern (unvollständige Aufnahmen)
…von Sprecher*innen, die <25% mit der Zielkontur realisieren (pro Kategorie: Kurz-/Langvokal)

1) F0-Extraktion über die akustisch festgelegte Zielsequenz (**V1S2**) mit R (Bombien et al., 2021)

2) F0-Normalisierung (gleichentfernte Punkte auf einem Zeitintervall von 0 bis 1)

3) Entfernung von Datenpunkten, bei denen f0-Wert gleich null (NAs) (z.B. Shigemori & Pouplier, 2020)
→ SH_ZH_0028_Hoehle_normal_01

**SD_0006_Diener_normal_02**

| sequence | V1_start | C2_end |
|---|---|---|
| i:->n | 935,725624 | 1091,55329 |

| times_orig | times_rel | times_norm | Hz | scaledF0 |
|---|---|---|---|---|
| 939,626 | 0,000 | 0,000 | 123,88 | -0,22 |
| 944,637 | 5,011 | 0,033 | 128,08 | 0,04 |
| 949,649 | 10,023 | 0,067 | 130,14 | 0,17 |
| 954,660 | 15,034 | 0,100 | 131,77 | 0,27 |
| 959,671 | 20,045 | 0,133 | 134,68 | 0,45 |
| 964,683 | 25,057 | 0,167 | 137,37 | 0,62 |
| 969,694 | 30,068 | 0,200 | 138,44 | 0,69 |
| 974,705 | 35,079 | 0,233 | 139,25 | 0,74 |
| 979,717 | 40,091 | 0,267 | 139,68 | 0,77 |
| 984,728 | 45,102 | 0,300 | 139,34 | 0,74 |
| 989,739 | 50,113 | 0,333 | 140,03 | 0,79 |
| 994,751 | 55,125 | 0,367 | 139,69 | 0,77 |
| 999,762 | 60,136 | 0,400 | 138,76 | 0,71 |
| 1004,773 | 65,147 | 0,433 | 141,24 | 0,86 |
| 1009,785 | 70,159 | 0,467 | 146,33 | 1,18 |
| 1014,796 | 75,170 | 0,500 | 149,81 | 1,40 |
| 1019,807 | 80,181 | 0,533 | 144,83 | 1,09 |
| 1024,819 | 85,193 | 0,567 | 145,38 | 1,12 |
| 1029,830 | 90,204 | 0,600 | 143,53 | 1,00 |
| 1034,841 | 95,215 | 0,633 | 140,04 | 0,79 |
| 1039,853 | 100,227 | 0,667 | 146,82 | 1,21 |
| 1044,864 | 105,238 | 0,700 | 147,62 | 1,26 |
| 1049,875 | 110,249 | 0,733 | 147,72 | 1,27 |
| 1054,887 | 115,261 | 0,767 | 147,95 | 1,28 |
| 1059,898 | 120,272 | 0,800 | 145,50 | 1,13 |
| 1064,909 | 125,283 | 0,833 | 145,96 | 1,16 |
| 1069,921 | 130,295 | 0,867 | 146,06 | 1,16 |
| 1074,932 | 135,306 | 0,900 | 148,76 | 1,33 |
| 1079,943 | 140,317 | 0,933 | 149,55 | 1,38 |
| 1084,955 | 145,329 | 0,967 | 145,66 | 1,14 |
| 1089,966 | 150,340 | 1,000 | 141,66 | 0,89 |

# METHODE
## VORBEREITUNG DES DATENSATZES

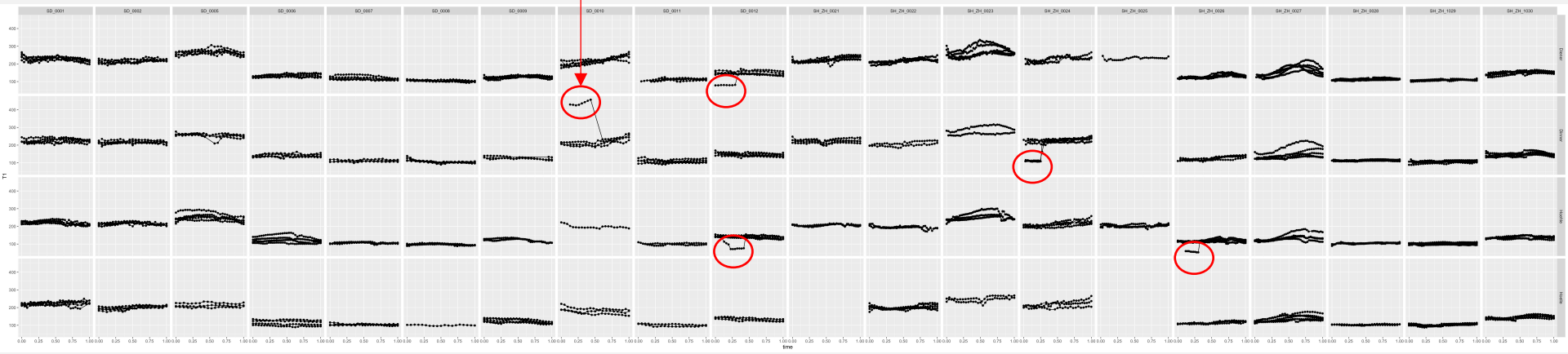
5) Entfernung von f0-Ausreißern (z.B. aufgrund von Knarrstimme)

6) Berechnung von sprecherskalierten f0-Werten (z-scores: Entfernung der einzelnen Datenpunkte vom Mittelwert) über alle stimmhaften Zeitfenster (z.B. Kirby et al., 2020)

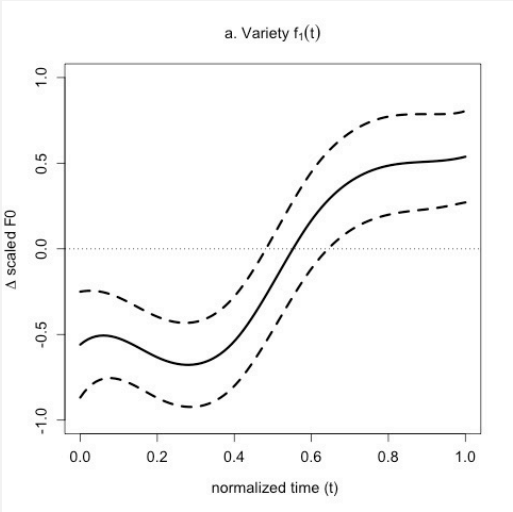
**SD_0006_Diener_normal_02**

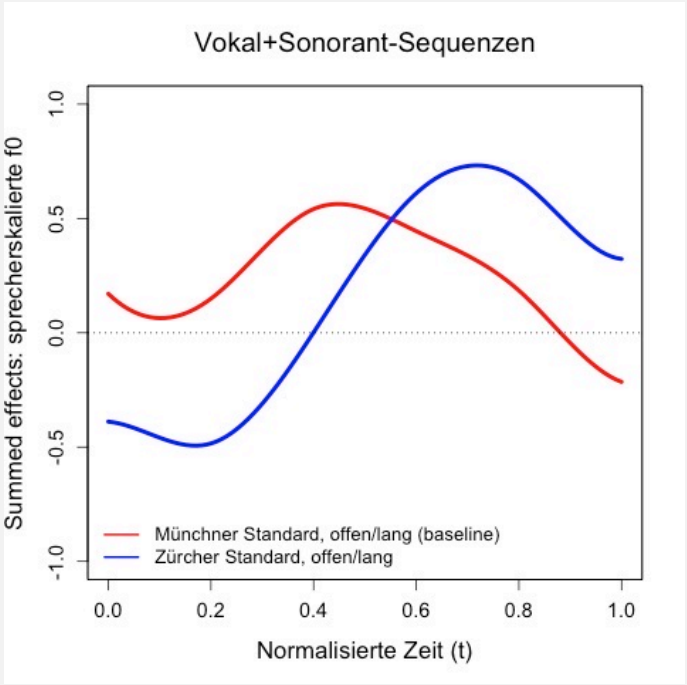
| sequence | V1_start | C2_end |
|---|---|---|
| i:->n | 935,725624 | 1091,55329 |

| times_orig | times_rel | times_norm | Hz | scaledF0 |
|---|---|---|---|---|
| 939,626 | 0,000 | 0,000 | 123,88 | -0,22 |
| 944,637 | 5,011 | 0,033 | 128,08 | 0,04 |
| 949,649 | 10,023 | 0,067 | 130,14 | 0,17 |
| ... | | | | |
| 1039,853 | 100,227 | 0,667 | 146,82 | 1,21 |
| 1044,864 | 105,238 | 0,700 | 147,62 | 1,26 |
| 1049,875 | 110,249 | 0,733 | 147,72 | 1,27 |
| 1054,887 | 115,261 | 0,767 | 147,95 | 1,28 |
| ... | | | | |



# ERGEBNISSE
## VERGLEICH VON MSG UND ZSG, SONORANTENDATEN (V1S2-SEQUENZ)




/øːl/-/œl/
/iːn/-/ɪn/

Statistische Analyse der extrahierten f0-Kurven (**V1S2**-Sequenz)
mit Functional Linear Mixed Models (FLMM; Cederbaum et al., 2016; Pouplier et al., 2017) in R (Cederbaum et al., 2021)

- 440 mögliche Observierungen (22 Sprecher*innen*4 Wöter*5 Wdh.)

→ 310 f0-Kurven, Datenpunkte pro Kurve: ⌀29.18 (10-62)

Referenzlinie (baseline: rote Linie)

- Klarer Gipfel zwischen Zeitpunkt 0.4 und 0.5

a. Effekt von *Varietät*

→ ZSG Sprecher*innen (blaue Linie) realisieren Gipfel später als MSG Sprecher*innen (zwischen Zeitpunkt 0.7 und 0.8)

# ZUSAMMENFASSUNG + DISKUSSION

In den vorliegenden Sonorantendaten zeigen sich regionale Unterschiede zwischen den untersuchten Standardvarietäten des Deutschen:

**V1S2-Sequenz**

- ZSG Sprecher*innen alignieren Gipfel in nicht-kontrastiven Nuklearakzenten signifikant später als MSG Sprecher*innen

→Im Einklang mit H1 (Ulbrich, 2005)

→Regionale Unterschiede zeigen sich nur in nicht-kontrastiven Akzenten, unabhängig von der Position innerhalb der IP
(Atterer & Ladd, 2004; Braun, 2007; Mücke et al., 2008; Mücke et al., 2009)

| | Nicht-kontrastive Akzente | Kontrastive Akzente |
|---|---|---|
| Pränukleare Position | ✓ | ✗ |
| Nukleare Position | ✓ (Sonorantendaten) | ✗ |

# DISKUSSION + AUSBLICK

- Heute Fokus auf Langvokalwörtern

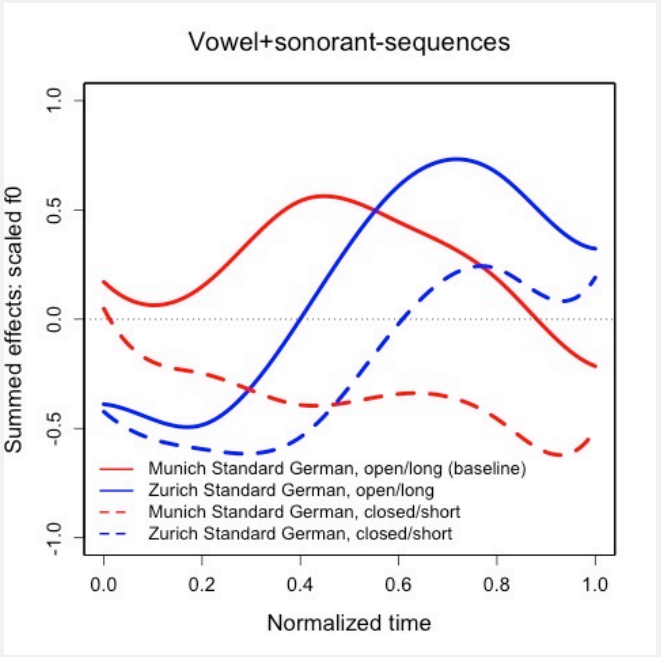
→ Ergebnisse für Kurzvokalwörter: Unterschiede bzgl. *Silbenstruktur* v.a. in MSG, betreffen eher f0 range als Alignment

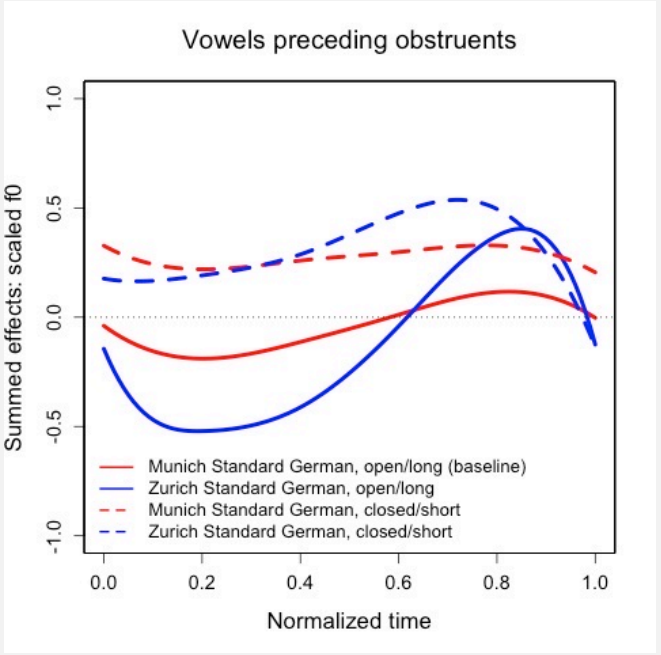
→ Modelle mit Obstruentendaten (z.B. *Bieter-bitter*): regionale Unterschiede können sich auch in f0-Modulation manifestieren

*Was kommt als nächstes?*

I. Einbezug von SG Sprecher*innen aus Wien → höhere Vergleichbarkeit mit früheren Studien, die "magische Momente" statt ganzen f0-Konturen verwenden (Atterer & Ladd, 2004; Braun, 2007; Mücke et al., 2008; Mücke et al., 2009)

II. Analyse von Dialektdaten aus den drei untersuchten Regionen



Forschung gefördert durch DFG-Projekt KL 2697/1- 2

# DANKESCHÖN!!

Kommentare oder Fragen?

# REFERENZEN


Atterer, M., & Ladd, D. R. (2004). On the phonetics and phonology of "segmental anchoring" of F0: evidence from German. *Journal of Phonetics, 32*(2), 177-197.

Bombien, L., Winkelmann, R., and Scheffers, M. (2021). *wrassp: an R wrapper to the ASSP Library*. R package version 1.0.1.

Braun, B. (2007). *Effects of dialect and context in the realisation of German prenuclear accents*. Paper presented at the XVIth International Congress of Phonetic Sciences (ICPhS), Saarbücken, Germany.

Cederbaum, J., Pouplier, M., Hoole, P., & Greven, S. (2016). Functional linear mixed models for irregularly or sparsely sampled data. *Statistical Modelling, 16*(1), 67-88.

Cederbaum, J., Volkmann, A., & Stöcker, A. (2021). *sparseFLMM: Functional Linear Mixed Models for Irregularly or Sparsely Sampled Data.* R package version 0.4.1. Retrieved from https://CRAN.R-project.org/package=sparseFLMM

Draxler, C. & K. Jänsch. (2004). *SpeechRecorder – A universal platform independent multi-channel audio recording software.* Paper presented at the 4th, International Conference on Language Resources and Evaluation (LREC), Lisbon, Portugal.

Fitzpatrick-Cole, J. (1999). *The alpine intonation of Bern Swiss German.* Paper presented at the XIVth International Congress of Phonetic Sciences (ICPhS), San Francisco, USA.

Fleischer, J., & Schmid, S. (2006). Zurich German. *Journal of the International Phonetic Association, 36*(2), 243-253.

Jochim, M. (2017). *Extending the EMU Speech Database Management System: Cloud Hosting, Team Collaboration, Automatic Revision Control.* Paper presented at the 18th Annual Conference of the International Speech Communication Association (Interspeech), Stockholm, Sweden.

Kirby, J., Kleber, F., Siddins, J., and Harrington, J. (2020). *Effects of prosodic prominence on obstruent-intrinsic F0 and VOT in German.* Paper presented at the Xth International Conference on Speech Prosody, Tokyo, Japan.

Kisler, T., U. Reichel, F. Schiel, C. Draxler, B. Jackl & N. Pörner. (2016). *BAS Speech Science Web Services – an Update on Current Developments.* Paper presented at the 10th International Conference on Language Resources and Evaluation (LREC), Portorož, Slovenia.

Leemann, A. (2012). *Swiss German intonation patterns.* Amsterdam: John Benjamins.

Mücke, D., Grice, M., Becker, J., & Hermes, A. (2009). Sources of variation in tonal alignment: Evidence from acoustic and kinematic data. *Journal of Phonetics, 37*(3), 321-338.

Mücke, D., Grice, M., Hermes, A., & Becker, J. (2008). *Prenuclear Rises in Northern and Southern German.* Paper presented at the 4th International Conference on Speech Prosody (SP), Campinas, Brazil.

# REFERENZEN


Peters, B., & Kohler, K. J. (2004). *Trainingsmaterialien zur prosodischen Etikettierung mit dem Kieler Intonationsmodell KIM.* Retrieved from: https://www.ipds.uni-kiel.de/kjk/pub_exx/bpkk2004_1/TrainerA4.pdf

Peters, J. (2014). *Intonation.* Heidelberg: Winter.

Pouplier, M., Cederbaum, J., Hoole, P., Marin, S., & Greven, S. (2017). Mixed modeling for irregularly sampled and correlated functional data: Speech science applications. *The Journal of the Acoustical Society of America, 142*(2), 935-946.

Shigemori, L.S.B. and Pouplier, M. (2020). *Realization of F0 on vocalic and consonantal nuclei in Slovak.* Paper presented at the 12th International Seminar on Speech Production (ISSP), Providence, USA (online).

Ulbrich, C. (2005). *Phonetische Untersuchungen zur Prosodie der Standardvarietäten des Deutschen in der Bundesrepublik Deutschland, in der Schweiz und in Österreich.* Frankfurt am Main: Peter Lang.

Winkelmann, R., J. Harrington & K. Jänsch. (2017). EMU-SDMS: Advanced speech database management and analysis in R. *Computer Speech & Language, 45,* 392-410.