

# EINFÜHRUNG IN DIE SOZIOPHONETISCHEN UNTERSUCHUNGSMETHODEN: AUF STUDENTISCHEM NIVEAU

---

Megumi Terada

68. Stuts

21.11.2020

Humboldt-Universität zu Berlin (MA Linguistik)

teradame@hu-berlin.de

# Zum Anfang

- Zielgruppen
  - Alle Interessenten für Phonetik/Soziolinguistik
  - Alle Nicht-Interessenten für Phonetik/Soziolinguistik 😊
- Ziel
  - Theoretische Grundlagen für die Soziophonetik
  - Unterschiedliche Arten von empirischen Forschungsmethoden
  - Worauf zu achten: Eine studentische Arbeit?
- Vorsicht
  - Generischer Maskulin in den Folien: Selbstverständlich werden alle Geschlechter inkludiert

# Vorhaben

- Soziophonetik?
- Konzipierung der Studie
- Datenerhebung
- Annotation
- Datenauswertung
- Übung
- Literaturen

# SOZIOPHONETIK?

---

# Begrifflichkeiten

- Schnittstelle zwischen Soziolinguistik und Phonetik
  - „Socially structured variation“ auf phonetischer Ebene
    - Jeder Mensch passt sich sprachlich an die Gesprächssituationen an
  - Von Natur aus empirisch
    - Hypothesen werden durch akustische Daten überprüft
    - Natürlichkeit der Daten ist der Schlüssel
  - Beschäftigt sich mit Fragen wie
    - Was ruft die Variationen hervor?
    - Welche akustischen Merkmale?
    - Kommunikative Bedeutung?
- Auch in Kombination mit anderen theoretischen Rahmen
  - Fremdspracherwerbsforschung
  - Psycholinguistik

# Arten und Einflussfaktoren der Variationen

- Inter-individuelle Variationen
    - Soziale „Klassen“, Netzwerke
    - Alter und Lebensumstände
    - Geschlecht
    - Regionale Variation
    - Ethnizität, Bilingualismus
  - Intra-Individuelle Variation (Register/Style)
    - Gesprächspartnern
    - Physische Umstände
    - Themen
    - Formalität/Höflichkeit
- Könnte parallel vorkommen, z.B. *observer effect*

# Was konkret?

- Inter-individuelle Variationen
  - Soziale „Klassen“, Netzwerke
    - Kiezdeutsch
  - Alter und Lebensumstände
    - Ausprägung der dialektalen Merkmale
  - Geschlecht
    - Rollenzuschreibung und Norm
  - Regionale Variation
    - Münchner Dialekt
  - Ethnizität, Bilingualismus
    - Pennsylvaniadeutsch

# Was konkret?

- Intra-individuelle Variation (Register/Style)
  - Gesprächspartner
    - Foreigner talk
    - Motherese, fatherese, pet-rese
  - Physische Umstände
    - Nervös vs. neutral
    - Betrunken vs. nüchtern
  - Themen
    - Wissenschaft
    - Babys
  - Formalität/Höflichkeit
    - Machtperson vs. Freunde
    - Vortrag vs. Feier

Interaktion: (inter-individuell + intra-individuell)



# Akustische Merkmale / Parameter

- Um Eigenschaft einzelner Variationen zu beschreiben
  - Unterschiedlich stark vertreten
  - Gut recherchieren, um die sinnvollsten Merkmale auszuwählen
- „Objektiver Maß“, um Aussagen treffen zu können
  - Ohrenphonetische Arbeit: Nicht plausibel genug, denn subjektiver Eindruck einzelner Personen
  - Einheitliche Messwerte: Richtlinie erstellen und daran halten, bewusst auf die Gefahr auf die ungewollte Fälschung der Resultate
- Kann auf unterschiedlichen Ebenen gemessen werden
  - Segmentale & suprasegmentale Ebene
  - Sonstige Phänomene gesprochener Sprache
- Auch als akustische Korrelate genannt

# Akustische Merkmale / Parameter

- Segmentale Ebene
  - Allophone von /ʁ/: /r/, /R/
  - Vokalqualität: /ɛ:/ und /a:/
  - Vokalraum: Realisierung von „Eckvokalen“
  - Phonologische Prozesse: Reduktionen
- Suprasegmentale (=prosodische) Ebene
  - Intonationsmuster
  - Tonhöhe/pitch
  - Stimmqualität
- Sonst noch
  - Stille Pausen
  - Sprechgeschwindigkeit
  - Gefüllte Pausen
  - Versprecher

- Zu Kategorien zugeordnet
- } Gemessen an Formanten [Hz]
- Gemessen an Dauer der Segmenten [ms]
- } Gemessen an Grundfrequenz (fo) [Hz]
- Gemessen an Jitter/Shimmer\*  
\*Regelmäßigkeit und Amplitude der Stimmlippenschwingungen
- Gemessen an Dauer [ms]
- Gemessen an Segmentdauer [ms] +  
Silbenanzahl (Silben pro Sek./Min.)

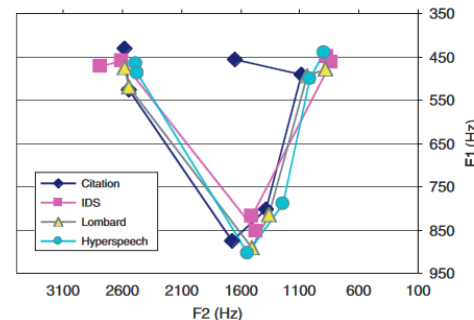


Abb.1: Eckvokale und Vergrößerung des Vokalraums je nach Gesprächseinstellungen (Wassink et al. 2007: 371)

# Welche akustischen Merkmale?

- Inter-individuelle Variationen
  - Soziale „Klassen“, Netzwerke
    - Kiezdeutsch: Alternation von /ç/ zu /ʃ/ (Formanten)
  - Alter und Lebensumstände
    - Ausprägung dialektaler Merkmale: <ä> in unterschiedlichen Generationen (Formanten)
  - Geschlecht
    - Rollenzuschreibung und Norm: Niedrigere Stimmlage von Frauen (fo)
  - Regionale Variation
    - Münchner Dialekt: Realisierung von /ʁ/ als /r/ (Lautkategorien)
  - Ethnizität, Bilingualismus
    - Pennsylvaniadeutsch: /ɛ:/ und /e:/ statt /aɪ/ (Formanten)

# Welche akustischen Merkmale?

- Intra-Individuelle Variation
  - Gesprächspartner
    - Foreigner talk: Reduzierte Sprechgeschwindigkeit (Segmentdauer)
    - Motherese, fatherese, pet-ese: Vergrößerter Vokalraum (Formanten)
  - Physische Umstände
    - Nervös vs. neutral: Mehr Versprecher (Häufigkeit)
    - Betrunken vs. nüchtern: Ausgeprägte Reduktionen (Segmentdauer)
  - Themen
    - Wissenschaft: Überartikulation (Segmentdauer, Formanten)
    - Babys: Erhöhte Stimmlage (fo)
  - Formalität/Höflichkeit
    - Machtperson Freunde: Creaky Voice (Jitter/Shimmer)
    - Vortrag vs. Feier: Mehr gefüllte Pausen (Häufigkeit)

# KONZIPIERUNG DER STUDIE

---

# Konzipierung der Studie

- Themenfelder erkunden
  - „Abschlussarbeiten“ auf der Institutsseite
  - Konkrete Vorstellungen durch Vorträge, Konferenzen, Kolloquien
    - Welche akustischen Parameter für welche sprachliche Variation am relevantesten?
    - Wie werden sie beschrieben und analysiert?
    - Alles sehr hilfreich, traut euch!
- Recherche
  - Welche Untersuchungsmethoden für welche Ziele?
  - Welche akustischen Parameter sollen was darstellen?
- Forschungsfrage formulieren
- Hypothese aufstellen

# Konzipierung der Studie

- Forschungsdesign
  - Theoretische Arbeit
  - Empirische Arbeit
    - **Produktionsstudie**: Akustische Analyse von Dialogen, Leseexperiment
    - Perzeptionsstudie: Rating-Experiment (wahrgenommene Angemessenheit, Verständlichkeit, Prestige, Emotion usw.)
    - Viele Papers über die Methodik („experimental design“)
- Methodik auswählen
  - Korpusstudie
  - Eigene Daten erheben

# Workflow: Theoretische Arbeit

- Studie konzipieren
- Daten erheben
- Daten annotieren
- Daten auswerten



# Workflow: Empirische Arbeit mit Korpus

- Studie konzipieren
- Daten erheben
- Daten annotieren
- Daten auswerten

# Workflow: Empirische Arbeit ohne Korpus

- Studie konzipieren
- Daten erheben
- Daten annotieren
- Daten auswerten

# Sammeln wir eigene Ideen!

- Schreibt eure Ideen gern im Etherpad

# DATENERHEBUNG

---

# Eigene Daten oder Korpus

- Schon aufgenommene Daten verwenden
  - Die glücklichste Variante! 😊
  - I.d.R. sehr viele frei verfügbare Akustikdaten am Institut
    - Korpus mit gesprochenen Daten
    - Oft bereits transkribiert und segmentiert
    - Oft gute Tonqualität (wichtige Voraussetzung für die akustische Analyse)
    - Kontrolliertheit
    - (Noch) nicht veröffentlichte Daten
  - Gern nachfragen!
- Bestehende Korpora: Annis, DGD (Datenbank für gesprochenes Deutsch) ...

# Eigene Daten oder Korpus

- Daten selbst erheben
  - Allgemeiner Aufwand
    - Versuchspersonen und Raum finden (ohne Aufwandentschädigung)
    - Einverständniserklärung
  - Aufnahmequalität
    - Räumlichkeiten
    - Aufnahmegeräte (Ausleihe im Labor)
  - Kontrolliertheit der Daten
    - Sprechereigenschaft (Alter, Geschlecht, sprachliche Hintergründe)
    - Größe des Datensatzes (statistische Gültigkeit vs. Machbarkeit)
    - Gesprächseinstellung (Monolog oder Dialog / experimentelle vs. natürliche Daten)
- Mit Korpora können all diese Schritte gespart werden

# Natürlichkeit vs. Kontrolliertheit

- Konflikt zwischen Natürlichkeit und Kontrolliertheit: „Sozio“-Phonetik
  - Entspricht der Realität? Akustische Analyse möglich?
  - Sprechdaten in natürlichen Gesprächskontexten, dafür weniger Kontrolle
    - + Authentizität: Soziale Eigenschaft der Gesprächssituation miteinbezogen
    - Aufnahmequalität: Hintergrundgeräusche, zu wenig Ziellaute
  - Experimente mit kontrollierten Bedingungen, dafür unnatürlich
    - + Aufnahmequalität: Kontrollierter phonologischer Kontext
    - Authentizität: Künstliche Gesprächssituation trotz Fokus auf soziale Faktoren

# Natürlichkeit vs. Kontrolliertheit

- Unterschiedliche Erhebungsmethoden
  - Gespräche in „echten“ Situationen: Auf der Straße, im Klassenraum
  - Freie Dialoge mit vorgegebenen Themen / Diapixe
  - Imaginäres Sprechen: Stimmnachricht, Rollenspiel
  - ✓ Kompromiss eingehen: „perfekte“ Aufnahmen gibt es nicht
    - Sinnvolle Einflussfaktoren
    - Machbarkeit: Studentische Arbeit!
- Kontrolle durch Vergleich zwischen den beiden Einstellungen
  - Tatsächlich unterschiedliche Ergebnisse: Imaginär vs. Echt





Datenschutzhinweise gemäß Art. 13  
der EU-Datenschutzgrundverordnung (DSGVO)

<b>Verantwortliche Stelle</b>	ELC, Festival of Light International Productions GmbH Königsplatz 9 10117 Berlin
<b>Datenschutzbeauftragter</b>	LEON SHARF Rechtsanwalt Königsplatz 9, 1. Etage 10117 Berlin E-Mail: <a href="mailto:datenschutz@elc.de">datenschutz@elc.de</a> Tel: +49 30 89 20 10 75 0 Fax: +49 30 89 20 10 75 9
<b>Zweck und Rechtsgrundlage der Datenverarbeitung</b>	Wahrung des Hausrechts (Art. 8 Abs. 1 Nr. 2 DSGVO)
<b>Berechtigtes Interesse</b>	<ul style="list-style-type: none"><li>• Erfüllung Veranstaltungszwecke (Hausrechtsschutz)</li><li>• Schutz des Eigentums</li><li>• Verhinderung und Aufklärung von Straftaten insbesondere Diebstahl und Vandalismus</li></ul>
<b>Speicherdauer</b>	48 Stunden zur Wahrung von Straftaten und sonstigen im Zusammenhang mit dem Hausrechtsschutz
<b>Berechtigtes Interesse an den Daten</b>	—
<b>Zusteller</b>	Besucher der Veranstaltung ELC Königsplatz 9 10117 Berlin Tel: +49 30 89 20 10 75 0 Fax: +49 30 89 20 10 75 9 E-Mail: <a href="mailto:info@elc.de">info@elc.de</a>

**SIERE DATENSCHUTZRECHTE**  
Sie haben das Recht auf Auskunft nach Art. 15 DSGVO, das Recht auf Berichtigung nach Art. 16 DSGVO, das Recht auf Löschung nach Art. 17 DSGVO, das Recht auf Einschränkung der Verarbeitung nach Art. 18 DSGVO sowie das Recht auf Widerspruch nach Art. 17 DSGVO sowie das Recht auf Datenübertragbarkeit nach Art. 18 DSGVO.

mer\_platz  
merplatz  
erplatz.de

ACHTUNG  
KAMERAÜBERWACHUNG



# Einverständniserklärung

- DSGVO (Datenschutz-Grundverordnung)
  - Keine Weitergabe der personenbezogenen Daten an Dritte
    - Wenn dann Anon- und Pseudomisierung je nach Abmachung
  - Könnte zu ernsten Konsequenzen führen
    - Reales, schlechtes Szenario: Datenlöschung, nicht für die Analyse verwendbar
    - Vor der Aufnahme ausgefüllt
- Vorlagen werden oft im Seminar zur Verfügung gestellt
- Ansonsten am Lehrstuhl fragen

# Metadaten

- Sehr relevant für die Analyse
- Sprechereigenschaft
  - Alter
  - Geschlecht
  - Regionale Herkunft
  - Sprachliche Hintergründe
- Sprechereinstellung
  - Eindruck dem Gesprächspartner gegenüber
  - Bewusste Strategien
- Am besten nach dem Experiment

# Hygienekonzept

- Danke, Corona...

# ANNOTATION

---

# Grundkonzepte

- Vorbereitung
  - Transkribieren (i.d.R. manuell, evtl. automatisch bei Leseexperiment)
  - Segmentierung
- Untersuchungsgegenstände annotieren
  - Markieren nach Richtlinie
  - Um Dauer/Grundfrequenz/Formanten zu messen; Silben zu zählen
  - Praat 😊 (Textgrid erstellen)

# Transkription

- Automatische Transkription?
  - Nochmal, DSGVO – „Weitergabe an die dritte Partei“
  - Je nach Bestimmungen untereinander
- Richtlinie erstellen
  - Unterschiedliche Darstellungsmöglichkeiten
    - Aussprachenbasiert: gibt's, so ne, hab
    - Orthographiebasiert: gibt es, so eine, habe
    - Konventionalisierte Transkription wie GAT
  - Was wird wie transkribiert?
    - Schwa-Elision
    - g-Spirantisierung
- Mit Korpora können all diese Schritte gespart werden

# Segmentierung

- Phonetische Transkription
  - Orthographische Darstellung der Segmente
  - /o:tɔgra:fiʃə da:ftɛlʊŋ dɛɛ zɛgmɛntə/ (IPA)
  - /o:tOgra:flS@ d6StElUN dE6 zEgmEnt@/ (SAMPA)
  - ✓ Vokallänge und Gespanntheit / Phoneme und Stimmhaftigkeit werden sichtbar
- Wird entsprechend im Zeitsignal markiert
- **M**unich **A**utomatic **S**egmentation (WebMAUS)
  - Chunk preparation + WebMAUS General oder Pipeline
  - Link: <https://clarin.phonetik.uni-muenchen.de/BASWebServices/interface/WebMAUSGeneral>
- Mit Korpora können all diese Schritte gespart werden



# Annotation

- Allererster Schritt: Richtlinie!
  - Relevante Exponenten auswählen (S.10f), basierend auf Literaturen
  - Für das einheitliche Ergebnis (=Gültigkeit der Studie)
  - Oft veröffentlicht als „Dokumentation“
    - Was wird aus welchen Gründen wie annotiert? (Gut begründen!)
- Untersuchungsgegenstände markieren (=annotieren)
  - Daten durchgehen und finden
  - Shortcuts möglichst früh lernen, bevor ihr anfangt Praat zu hassen
  - Automatisierung der Arbeitsschritte? (evtl. programmieren)

# DATENAUSWERTUNG

---

# Daten auswerten

- Statistik + Abbildungen
  - Wenn nicht vertraut: Excel
    - Daten exportieren als csv-Tabelle in Praat (Tabulate > List)
    - In Excel/Libreoffice einlesen (Daten > Aus Text/CSV > Importieren)
      - Auf Encodierung (UTF8, ASCI ...) und Separator (Komma, tab ...) achten
    - Funktion für t-test
    - Abbildungen manuell erstellen
  - Wenn etwas weiter: R
    - Datenbank konvertieren mit Package *emuR*
    - Statistische Tests mit/ohne Packages
    - Abbildungen mit *ggplot2*

# Ergebnisse zusammenfassen

- Als eine Hausarbeit, Bachelorarbeit oder Masterarbeit
- Wichtig: Deskriptive Art
  - Keine Bewertung/Urteil
  - Objektivität ist das höchste Gebot
- Publizieren in edoc-Server
  - Link: <https://edoc.hu-berlin.de/>
- Datenveröffentlichung via Repositorium

# Zu guter Letzt...

- Je nach Energie und Leidenschaft
  - Aufwand ist viel höher als man denkt: Eigene Grenze kennen
  - Wichtiger: Grenze setzen!
- An den Vorschriften halten
  - In veröffentlichten Richtlinien stehen nützliche Anweisungen und Vorschläge
  - Immer nachfragen bei Betreuer\*innen, daran halten

# Übung

- Entwickeln wir gemeinsam Studienkonzepte!
    - Vortrag vs. Freie Gespräche (Formalität – Sprechgeschwindigkeit)
    - Nüchtern vs. Betrunknen (physiologischer Zustand – segmentale Merkmale)
    - Flirten vs. Nicht-Flirten (Sprecherrelation – Verschleifung & Tonhöhe & Stimmqualität)
  
  - 1. Recherche: Theoretischer Rahmen, akustische Parameter
  - 2. Einverständniserklärung: Ethische Vereinbarkeit
  - 3. Aufnahme: Echte Situation oder Rollenspiel
  - 4. Datenvorbereitung: Annotationsebenen
  - 5. Akustische Analyse
  - 6. Statistik
- 
- Perfekt für eine HA 😊

# Übung

- Entwickeln wir gemeinsam Studienkonzepte!
  1. Recherche: Theoretischer Rahmen, akustische Parameter
  2. Einverständniserklärung: Ethische Vereinbarkeit
  3. Aufnahme: Echte Situation oder Rollenspiel
  4. Datenvorbereitung: Annotationsebenen
  5. Akustische Analyse
  6. Statistik
- Viel Spaß!

# Quellen

- Literatur

- Albert, Ruth & Marx, Nicole (2017). *Empirisches Arbeiten in Linguistik und Sprachlehrforschung*. Narr Starter. Tübingen.
- Boersma, Paul & Weenink, David (2018). Praat: doing phonetics by computer [Computer program]. Version 6.2.30, abgerufen am 10 November 2020 von <http://www.praat.org/>
- Foulkes, Paul, James M Scobbie & Dominic Watt (2010). Sociophonetics. In William J. Hardcastle, John Laver, und Fiona E. Gibbon (Hg.), *The handbook of phonetic sciences*, Blackwell handbooks in linguistics, S. 703–754. Malden, MA [u. a.]: John Wiley & Sons, Ltd.
- Kendall, Tyler (2013). *Speech rate, pause and sociolinguistic variation Studies in corpus sociophonetics*. Basingstoke: Palgrave Macmillan.
- Kisler, Thomas, Reichel, Uwe D. & Schiel, Florian (2017). Multilingual processing of speech via web services, *Computer Speech & Language*, Volume 45 (September 2017), 326–347.
- Scarborough, Rebecca, Olga Dmitrieva, Lauren Hall-Lew, Yuan Zhao & Jason Brenier (2007). An acoustic study of real and imagined foreigner-directed speech. *Journal of the Acoustical Society of America* 121(5), 3044.
- Thomas, Erik, R. (2011). *Sociophonetics: An introduction*. Basingstoke [u.a.]: Palgrave Macmillan.
- Wassink, Alicia Beckford, Wright, Richard A., & Franklin, Amber D. (2007). Intraspeaker variability in vowel production: An investigation of motherese, hyperspeech, and Lombard speech in Jamaican speakers. *Journal of Phonetics*, 35(3), 363-379.
- Winkelmann, Raphael, Harrington, Jonathan & Jänsch, Klaus (2017). EMU-SDMS: Advanced speech database management and analysis in R. *Computer Speech & Language*, Volume 45 (September 2017), 392-410.

- Bild

- [https://www.irasutoya.com/2013/12/blog-post\\_7916.html](https://www.irasutoya.com/2013/12/blog-post_7916.html) (abgerufen am 20.10.20)