

Die Erforschung des Sprechens mittels Nachrichtenkorpora – Die Nachrichtenarche der ARD

Sven Grawunder (Leipzig)

Die Nachrichten-Arche der ARD (hier kurz NArche) trägt der Tatsache Rechnung, dass es bis dato keine flächendeckende, langzeitorientierte und systematische Sammlung von Nachrichtensendungen in Deutschland gab (vgl. Schwiesau et al. i.d.B.). Die NArche liefert seit 2003 einen jährlichen Querschnitt über die Rundfunklandschaft des öffentlich-rechtlichen Radios in Deutschland und ermöglicht einen Einblick in die Entwicklung des Submediums Nachrichten, aber auch “benachbarter” Textsorten (z.B. Bericht, Wetter, Sport).

Anhand mehrerer Teilstudien soll die Nutzbarmachung einer Nachrichtensammlung, wie die NArche sie darstellt, für medienorientierte Forschungen der Sprechwissenschaft, aber auch für linguistisch orientierte Fragen der Phonetik verdeutlicht werden. So kann demonstriert werden, wie homogen die sprecherische Gestalt innerhalb des eng umschriebenen Gestaltungsrahmens der Textsorte verbleibt (2.1 und 2.2). Dabei wird eruiert, welche möglichen Faktoren durch Sender, Sprecher oder Modus von Bedeutung sind, wenn es darum geht, Generalisierungen über das gesamte Medium hinweg zu treffen. Es zeigt sich, dass sich die Spezifik der Textsorte Nachrichten auch in einzelnen Parametern wie etwa der mittleren Grundfrequenz oder der Sprechgeschwindigkeit darstellen lässt. Damit stellt das Korpus an sich ein *tertium comparationis* dar. Darüber hinaus eignen sich Nachrichtensprecher in ihrer Eigenschaft als standardsprachliche Modellsprecher ideal auch als Referenz für suprasegmental, intersegmental und segmental orientierte Analysen zur Standardaussprache (3.1 und 3.2). Letztlich lassen sich auch Fragen zur Umsetzung sprechorientierten Schreibens aus einem vorhandenen Skript, im Sinne von gewählten Strategien, bearbeiten (3.2) und zu experimentellen Zwecken erhobene ‘nachrichtenartige’ Stimuli validieren (2.3).

1 Was macht Nachrichten, was macht die NArche interessant?

1.1 Geschlossenheit und Differenziertheit

Gesprochene Nachrichten (besonders im Radio) sind schon seit geraumer Zeit Gegenstand phonetischer Untersuchungen (vgl. z.B. Sendlmeier 2005, Grawunder et al. 2006) und wurden umgekehrt benutzt, um phonetische Fragestellungen (Benton et al. 2007, Klaaß 2008) zu beantworten. Was macht nun eine Sammlung von gelesenen Nachrichten für Phonetik und Sprechwissenschaft interessant, die eigentlich aktuell vor allem auf gesprochene Spontansprache ausgerichtet sind? An erster Stelle ist hier die Geschlossenheit des Korpus zu nennen. Schließlich handelt es sich bei Nachrichtensendungen im öffentlich-rechtlichen Radio um eine klar definierte und in hohem Maße institutionalisierte sprachliche Situation.

Neben der Geschlossenheit verfügt das Korpus aber auch über eine interne Differenziertheit. Sie betrifft die geographische Verortung der jeweiligen Sender und ihrer Sprecher, die Programmformate (Info-Radio, regionale vs. überregionale Programme), die innerhalb der Nachrichtensendung verwendeten Textsorten bzw. Darstellungsformen (Meldungen, Sport, Wetter, Verkehr, Berichte etc.) sowie die sprecherspezifischen Eigenschaften (Geschlecht, Alter etc.). Dabei ist gerade die Vielzahl der Sprecher pro Querschnitt interessant sowie die Menge an Material, das unter so kontrollierten Bedingungen gewonnen wurde. Gerade wegen der Rigidität und Reglementiertheit von Nachrichtensendungen ist diese interne Varianz innerhalb des Korpus besonders interessant, denn Hörer verfügen offensichtlich nur über eine geringe Toleranz gegenüber Abweichungen von dieser Norm; sie nehmen selbst geringe Abweichungen wahr und empfinden sie u.U. als störend (s. z.B. Hollmach 1996).

1.2. Nachrichten als Referenz

Eine Sammlung von gesprochenen Nachrichten stellt einen Referenzpool für Standardsprecher dar. Dieser Ansatz wurde u.a. bei den Untersuchungen zum Wörterbuch der Deutschen Aussprache genutzt (GWdA 1982, DAWB 2009). Die NArche verfügt gegenüber ähnlichen Korpora (Verbmobil, Kielkorpus, Tu

B) insofern über einen Vorteil, als sie Aufnahmen mit einer regionalen Streuung über das gesamte binnendeutsche Sprachgebiet und zu geringem Teil auch darüber hinaus enthält. Die Sprecher sind – institutionell begründet – Vertreter einer akzeptierten Interpretation von Standardaussprachenormen, die je nach Reichweite (vgl. z.B. Deutschlandfunk vs. MDR1 Radio Sachsen-Anhalt) überregional definiert ist.

Eine weitere Perspektive ist die auf gesprochene / gelesene Nachrichten als eine kanonische Form bzw. einer kanonischen oder Standard-Sprach-Form folgende Textsorte. Die Professionalität, sprich Routiniertheit, der ‘Leser’-Sprecher ist dabei ein erwünschter Faktor, zumindest was sprechmotorische Abläufe auf Wort- und Phrasenebene anbelangt. Interessant, weil auch aus der Alltagsbeobachtung heraus auffällig, bleiben die spezifischen Merkmale der prosodischen Gestalt der Nachrichten. Dabei ist klar, dass eine Reihe von Merkmalen zur erwarteten Form der Offizialität und Neutralität beitragen, während andere die sprecherspezifischen Strategien der Umsetzung des geschriebenen Textes widerspiegeln. Welche sind durch die Schriftlichkeit der Vorlage unvermeidbar? Wo kann optimiert werden? Ein Beispiel sind die an anderer Stelle schon vielfach betrachteten Relativsätze bzw. deren prosodische Realisierung, die für die Nachrichten (Modus 1) untersucht werden sollen (vgl. 3.2).

1.3 Was kann man erheben? Was bietet das NArche-Korpus?

Die bereits im Beitrag von Schwiesau et al. (i.d.B.) angesprochenen thematischen und formulatorischen Überlappungen lassen sich z.T. dadurch erklären, dass die Agenturmeldungen *verbatim* in den Nachrichtenmeldungstext übernommen werden. Welchen Anteil an solchen Quasi-Paralleltexten finden wir pro Jahr vor? Dabei ist die Grundannahme, dass eine Top-Meldung, also eine Meldung mit einer Vorkommenshäufigkeit von über 50%, auch das Aufkommen paralleler Passagen, also identisch verwendeten Texts, begünstigt. Um dies zu exemplifizieren, wird eine Such-Iteration auf orthographischer Satzebene durch die Jahrgänge 2005 und 2008 verwendet. D.h., es wird nach einer vollständigen Übereinstimmung für einen ganzen orthographischen Satz gesucht.

Für NA05 finden sich 44 Sätze, die durchschnittlich zwei- bis dreimal vorkommen. Bei NA08 sind es 101 identische Sätze mit einer Häufigkeit von zwei- bis zu 11mal. Demgegenüber zeigt sich, dass in beiden Jahrgängen eine höhere Trefferquote von Übereinstimmungen von kürzeren Sätzen zu beobachten ist, die zudem noch als Lead-Satz fungieren. So kommt der Satz „*Hartz war im Zuge der VW-Affäre zurückgetreten*“ (NA05) fünfmal vor, der Lead-Satz „*Der Schauspieler Horst Jüssen ist tot*“ (NA08) sogar bei zehn verschiedenen Sprechern. Dieses Ergebnis kann durchaus als repräsentativ hinsichtlich des Gesamtkorpus angesehen werden, so dass die Wahrscheinlichkeit, einen identischen Text (Satz) zu finden, auf Überzufallsniveau liegt. Allerdings ist die Quote von Überlappungen auf Unter-Satzebene (Phrase, n-Wortgruppe) wesentlich höher und kann weiter bis auf Graphemfolgen (Bi-, Trigraphen) „reduziert“ ermittelt werden. Das Vorkommen von syntaktischen Umstellungen und Umformungen („*Bundestagsdebatte*“ vs. „*Debatte im Bundestag*“) macht jedoch die Fragestellung nach einer engeren Phrasengleichheit nicht trivial.

2 Die phonetische Beschreibung des Sprechstils Nachrichten

Im Folgenden wird demonstriert, welche Fragen sich hinsichtlich einer phonetischen Beschreibung des NARche-Korpus bereits beantworten lassen. Die statistischen Analysen wurden mit Hilfe der freien statistischen Softwarepakete *R* (R Development Core Team 2009) erstellt. Alle Angaben zu einzelnen Paketen oder Funktionen beziehen sich auf diese Umgebung. Die unmittelbaren Messungen und Parametrisierungen erfolgten mit Hilfe der akustischen Analysesoftware PRAAT (Boersma 2011).

2.1 Globale Parameter

Für die parameterbasierte Beschreibung des Korpus und seiner Untereinheiten (Jahrgänge, Sender, Modi, Gender etc.) lassen sich relativ unaufwändig folgende Charakterisierungen des Sprechstils mittels akustischer Korrelate von mittlerer Sprechstimmhöhe, Sprechtonhöhenvariabilität, dynamischer Variabilität sowie Sprechgeschwindigkeit vornehmen. Für die Messung der mittleren F0 wie auch der anderen Parameter wird auf den Medianwert innerhalb einer Meldung zu-

rückgegriffen. Diese Medianwerte pro Meldung werden dann weiter pro Sprecher und Jahrgang zusammengefasst. Exemplarisch werden hier die Jahrgänge 2008 / 2009 betrachtet, sie können mit bereits vorliegenden Ergebnissen zu den Jahrgängen 2003 / 2004 / 2006 verglichen werden (Grawunder et al. 2006 und 2008).

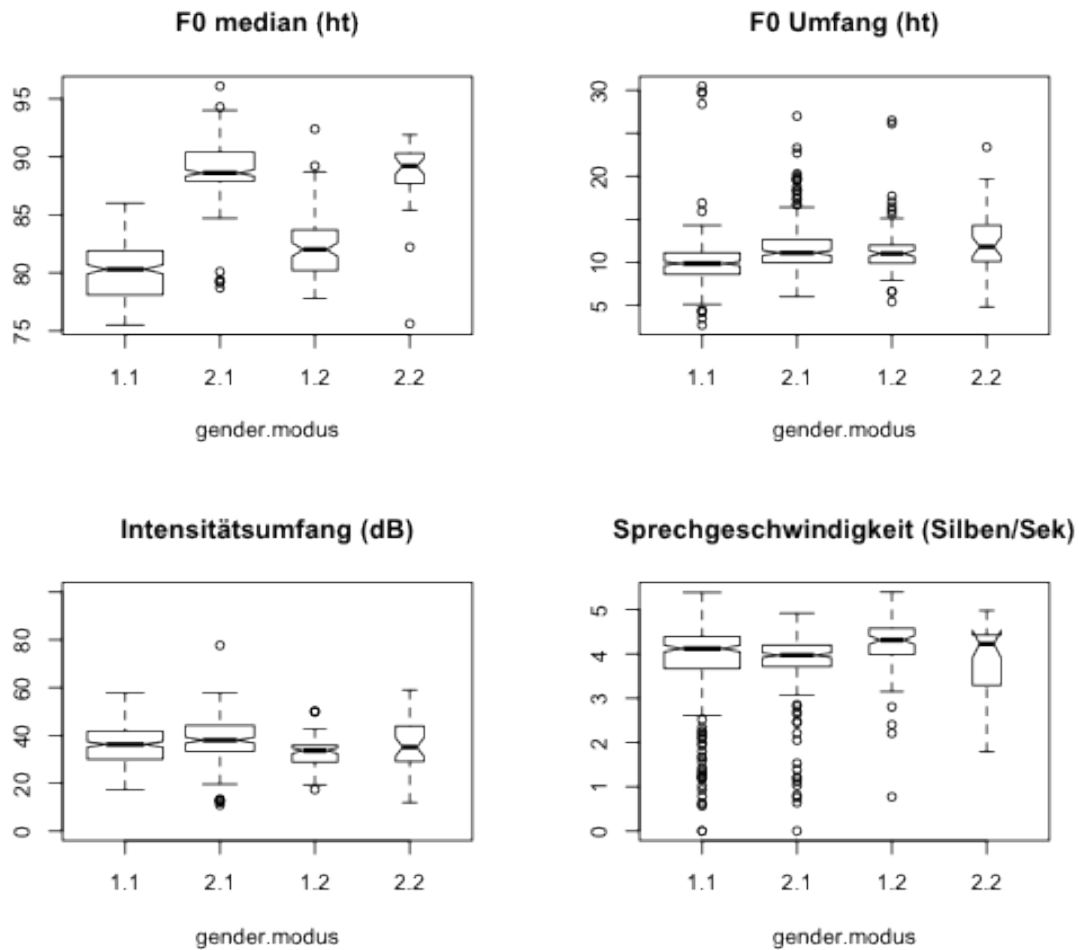


Abb. 1: Nachrichtenarche 2008/09 – Vergleich Modus 1 („klassische“ Nachricht) und Modus 2 („Bericht, Report“) sowie Geschlecht bzw. Gender (1 = männlich; 2 = weiblich)

Systematisch zeigen sich bei Modus 2 („Bericht“) im Durchschnitt höhere F0-Werte (ca. 5 Halbtöne), ein höherer F0-Umfang, jedoch geringere durchschnittliche Intensitätswerte und eine leicht höhere Sprechgeschwindigkeit als beim Modus 1 („klassische“ Nachricht). D.h., der Abstand zwischen den Modi fällt jahrgangswise verschieden aus, bleibt jedoch konstant in seiner Tendenz.

Wenn man sich hierzu das Profil der F0-Werte in einer Art Histogramm der Werte anschaut, in dem die Häufigkeit eines F0-Wertes innerhalb eines Messintervalls abgebildet wird, so zeigen sich für einen Teil der Samples moduspezifische Verteilungen (s. Abbildung 2), die somit eine zusätzliche Charakteristik zu den Mittel-, Median- und Modalwerten beitragen. So kann man für Modus 2 einen breiteren Peak und eine breitere, flacher auslaufende hohe (rechte) Flanke als Tendenz ausmachen, die es systematisch zu prüfen gilt.

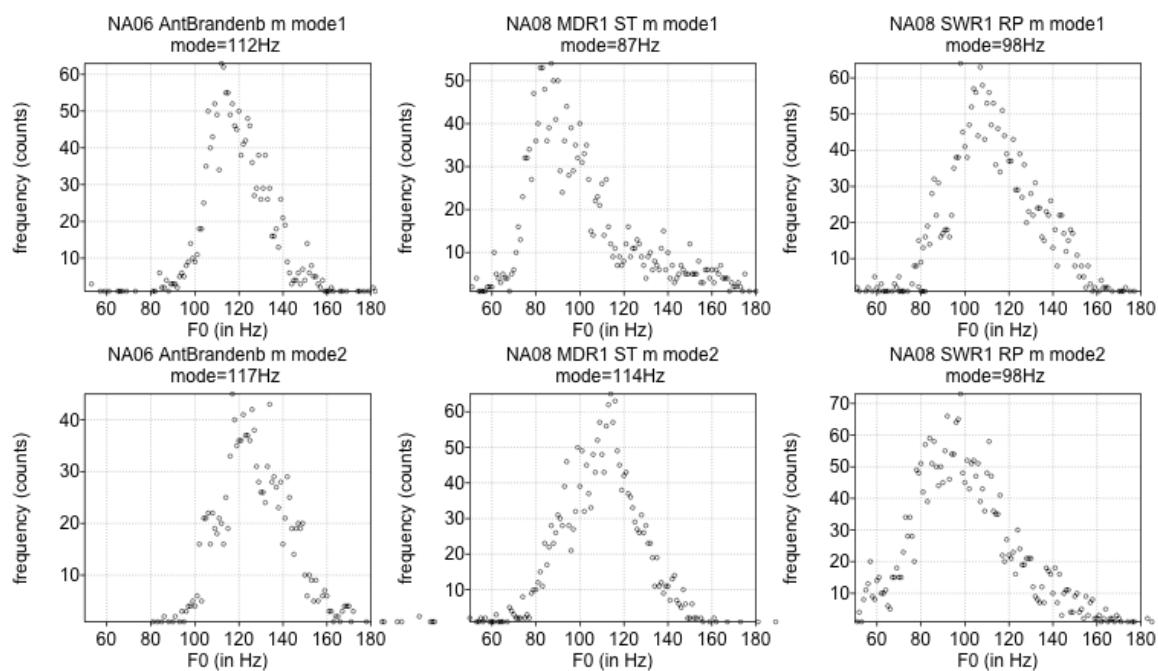


Abb. 2: F0-Profile für Modus 1 und Modus 2 von 3 verschiedenen Sendern; 6 männliche Sprecher

2.2 Der Rhythmus der Nachrichten – die Homogenität des PVI

Neben der Erhebung von solch globalen akustischen Eigenschaften wie den auf eine Sequenz bezogenen F0-Parameterausprägungen ist auch eine Charakterisierung der temporalen Struktur sinnvoll, da sie auf rhythmisch-metrische Eigenschaften hinweisen kann. Pausen auf der einen Seite und segmentale wie prosodische Variabilität auf der anderen tragen zur Wahrnehmung rhythmischer Eigenschaften gesprochener Sprache bei. Die jüngste phonetische Forschung hat sich auf die Dauernvariabilität von aufeinanderfolgenden Segmenten konzentriert, weil hier ein enger Zusammenhang mit der Markierung prominenter Silben beobachtet wurde. Gerade zur Prüfung der klassischen Rhythmusklassen- bzw. Isochroniehypothese (Pike 1945, Abercrombie 1967) wurden mehrere Maße zur Beschreibung der Variabilität entwickelt – darunter vielleicht am promi-

nentesten der *pairwise variability index* (PVI) von Low (Grabe / Low 2002). Dieser ‘paarige Variabilitätsindex’, der in seiner Formulierung sehr stark an solche Perturbationsmaße wie Jitter oder Shimmer angelehnt ist, wird sowohl auf konsonantische als auch vokalische ‘Segmente’ angewendet und hat eine sogenannte “raw” Version: $rPVI = \sum_{k=1}^{m-1} |d_k - d_{k+1}| / (m-1)$, die die Differenzen direkt abbildet, als auch eine über den Durchschnittswert der Differenzen normalisierte Version: $nPVI = 100 \times \left[\sum_{k=1}^{m-1} \frac{|d_k - d_{k+1}|}{(d_k + d_{k+1})/2} \right] / (m-1)$, wobei d hier die Dauern der aufeinanderfolgenden ($k, k+1, \dots$) Segmente und m deren Anzahl darstellt (vgl. Grabe / Low 2002). Dass dieses normalisierte Maß nicht ausreicht, um die Effekte verschiedener Artikulations- und Sprechgeschwindigkeiten zu kompensieren, ist inzwischen hinlänglich belegt (s. z.B. Dellwo / Wagner 2003).

Um die Sensibilität dieses PVI-Maßes zu testen, wurden die Jahrgänge 2003, 2004, 2005 und 2006 der NArche verwendet. Dazu wurden 40 Sendungen mit Sprechmodus 1 und 2 ausgewählt und die abgrenzbaren Einheiten sowohl manuell segmentiert als auch nach ihrer Konsonant-/Vokalartigkeit etikettiert (C=Konoid und V=Vokoid). Postvokalische Nasale und Liquide wurden in dieser Version dem Vokal zugeschlagen, weil sie als Längung des vokoiden Segments angesehen werden können. Ebenso wurden silbische Nasale und Liquide als Vokoid gelabelt, da sie zumeist einen ausgefallenen Vokal ersetzen bzw. dessen rhythmische Rolle erfüllen (vgl. Abbildung 3).

Die Samples wurden hinsichtlich Sprechgeschwindigkeit und Dauer kontrolliert, so dass nur Samples mit ähnlicher Sprechgeschwindigkeit von einer Länge um 25sek verwendet wurden. Die Ergebnisse für alle manuell segmentierten Aufnahmen (N=40) zeigen ähnliche nPVI(C) in beiden Modi, jedoch niedrigere nPVI(V) Werte in Modus 1, d.h. die Dauervariation in den vokalischen Segmenten ist für Nachrichten niedriger. Während die die nPVI(V/C) im Modus 2 höher sind, zeigen männliche Sprecher generell ein höheres nPVI(V/C)-Niveau. Dies ist ein zunächst unerwarteter Gender-Effekt, der weiter geprüft werden muss. In vergleichbaren Untersuchungen (Benton et al. 2007) von englischen und Mandarin chinesischen Nachrichten und Interviewsituationen wurden ebenfalls hö-

here nPVI(V)-Werte gefunden, allerdings für weibliche Sprecher. Für die englischen und chinesischen Nachrichten sind die nPVI-Werte gleichsam niedriger. Dies scheint ein weiterer Beleg für die bereits beobachteten unterschiedlichen Artikulationsgeschwindigkeiten bei Männern und Frauen zu sein.

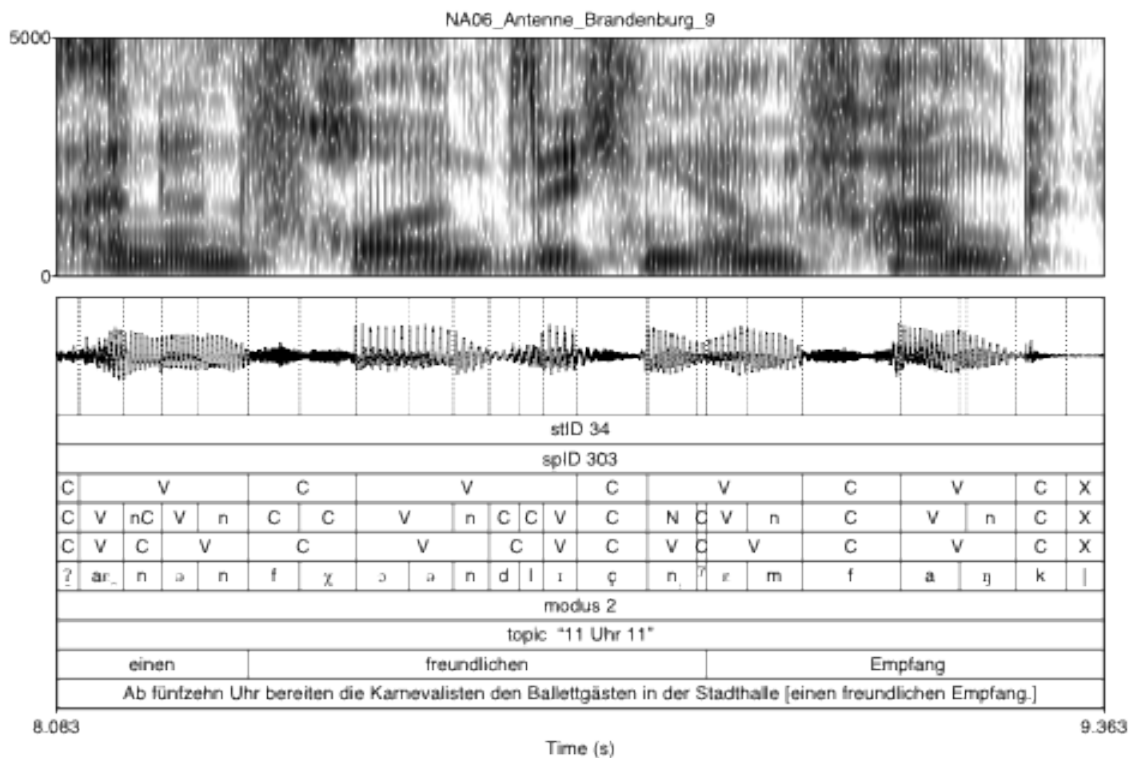


Abb. 3: Annotationszeilen (tiers); die Zeilen 3 bis 6 sind hier spezifisch für die rhythmusorientierte Variabilitätsanalyse der Dauern von Segmenten (s. 4.2) bestimmt; 1 = Sender ID; 2 = Sprecher ID; 7 = (Sprech-)Modus; 8 = Meldungsthema; 9 = Wortebene; 10 = orthographische Satzebene

2.3 Validität generierter Test-Nachrichten – Validität der Stilmerkmale

Die sprechwissenschaftliche Forschung zu Nachrichtensprache und -sprechen bedient sich u.a. experimenteller Verfahren – Bose et al. (i.d.B.) verwenden z.B. für Perzeptionsexperimente zur Nachrichtenverständlichkeit Testsendungen als Stimuli. Die Validität der Stimuli kann insofern mit Hilfe der NArche statistisch getestet werden, als angenommen wird, dass eine solche Test-Nachricht Bestandteil der Grundgesamtheit von Nachrichten ist bzw. nicht signifikant davon abweicht. Dazu wurden zwei Jahrgänge (2008/2009) als Referenz-Stichprobe verwendet. Nur für den Modus 1 wurden die Grundfrequenzparameter F0-Median, F0-StAbw, und F0-Umfang sowie die Sprechgeschwindigkeit (Silben

pro Sekunde) betrachtet. Eine äußerst relevante Frage ist hierbei, wie die zu vergleichenden Mittelwerte erhoben werden. Würde man alle Meldungen, also diejenigen Sequenzen, über die ein Messwert erhoben wird, zusammenfassen, ohne Sprecher (abgesehen von deren Gender) als gruppierenden Faktor zu berücksichtigen, sondern Sprecher lediglich als “random effect” fungieren lassen, ließe sich diese Formel in einem gemischten linearen Modell (LME - linear mixed effect model) (Bates et al. 2008; Baayen et al. 2008) testen.

In diesem Fall wurden 1061 Messwerte getestet hinsichtlich ihrer zu erwartenden “fixed effects” Gender (Frauen vs. Männer) und Korpus (NArche vs. Testsendungen) und hinsichtlich der Faktoren Sprecher (67) und Aufnahme (80).

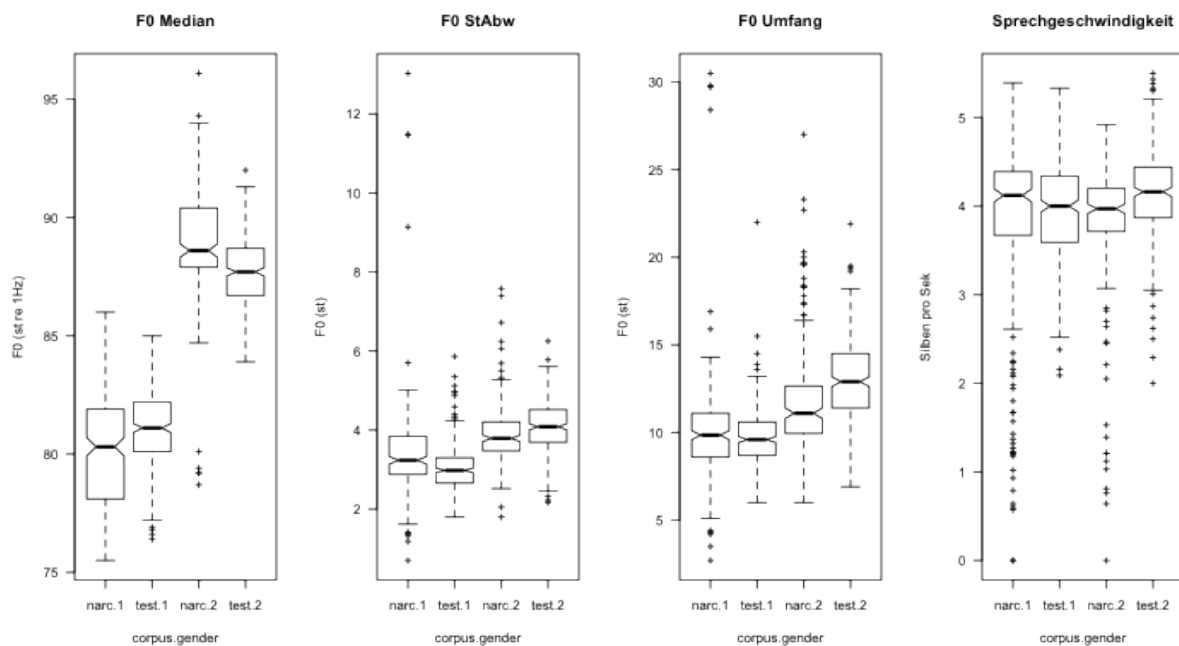


Abb. 4: Test-Sendungen vs. Nachrichten-Arche

Tabelle 1 enthält für die Parameter F0, F0-StAbw, F0-Umfang und Sprechgeschwindigkeit (Silben pro Sekunde) die p-Werte des Haupteffekts (NArche vs. Testsendungen) sowie der Interaktion (Korpus vs. Gender) und das Ergebnis des Differenztests (ANOVA) mit einem randomisierten Null-Modell (Freiheitsgrade werden üblicherweise nicht berichtet.). Dabei ist erkennbar, dass sich die Werte der beiden Korpora nicht signifikant unterscheiden. Nach diesem Vorgehen

handelt es sich bei den fiktiven Testnachrichten phonetisch betrachtet um realitätsnahe Nachrichten.

Lediglich für den F0-Median scheint eine Interaktion mit Gender nahe, was allerdings auch zu erwarten wäre, weil Männer und Frau bekanntermaßen verschiedene Sprechstimmlagen haben. Doch ist die Variation innerhalb der Sprecher und Sprechergruppen so groß, dass auch Überschneidungen zum anderen Geschlecht auftreten. Dies bestätigt eine frühere Beobachtung, bei der zum Teil recht hohe mittlere Sprechstimmlagen für Männer und *vice versa* tiefe Lagen für Frauen auftraten (Grawunder / Bose 2008).

Allerdings würde es sich bei der oben geschilderten Herangehensweise strenggenommen um einen Fall von Pseudoreplikation (Hurlbert 1984) handeln, weil es sich dabei nur teilweise um voneinander unabhängige Messungen handelt und die Abhängigkeit von der Variable Sprecher ignoriert würde. Trotzdem wäre eine solche Herangehensweise aus Erwägungen der vollen Datenexploration doch unter Berücksichtigung der Einschränkungen einer Faktorreduktion zu rechtfertigen.

Tab. 1: Test-Statistik des LME-Modells mit Sprecher als *random effect*

Parameter	Haupteffekt [Korpus] (pMCMC)	Interaktion [Gender] (pMCMC)	Null-Modell (χ^2)
F0 Median	0.4528	0.0776	< 2.2e-16
F0 StAbw	0.1334	0.0948	0.003416
F0 Umfang	0.5318	0.1590	0.001906
SprGeschw	0.8742	0.4964	0.9095

Fasst man allerdings die Werte pro Sprecher als Mittelwerte zusammen, ergeben sich lediglich zwei Werte pro Gender für das Test-Korpus. Damit sind übliche statistische Tests zwar nicht mehr möglich, doch können für jede Gender-Gruppe die Grenzwerte des NArche-Korpus des Konfidenzintervalls berechnet werden. Dies erfolgte mit Hilfe des Wilcoxon-Tests, da laut Shapiro-Wilk-Test lediglich drei von zehn Wertegruppen (2 Gender \times 5 Parameter) eine Normalverteilung aufweisen.

Augenscheinlich (s. Abbildung 5) sind es besonders die Testnachrichten-Sprecherinnen, die durch eine höhere mittlere Grundfrequenz, einen leicht höheren Grundfrequenzumfang, höhere F0-Varianz und höhere Sprechgeschwindigkeit von ihren ‚Kolleginnen‘ in den NArche-Jahrgängen 08/09 abweichen. Die männlichen Testnachrichten-Sprecher sind bis auf eine geringere Varianz der Grundfrequenz als Teilmenge einer vom NArche-Korpus abgeleiteten Grundgesamtheit anzusehen.

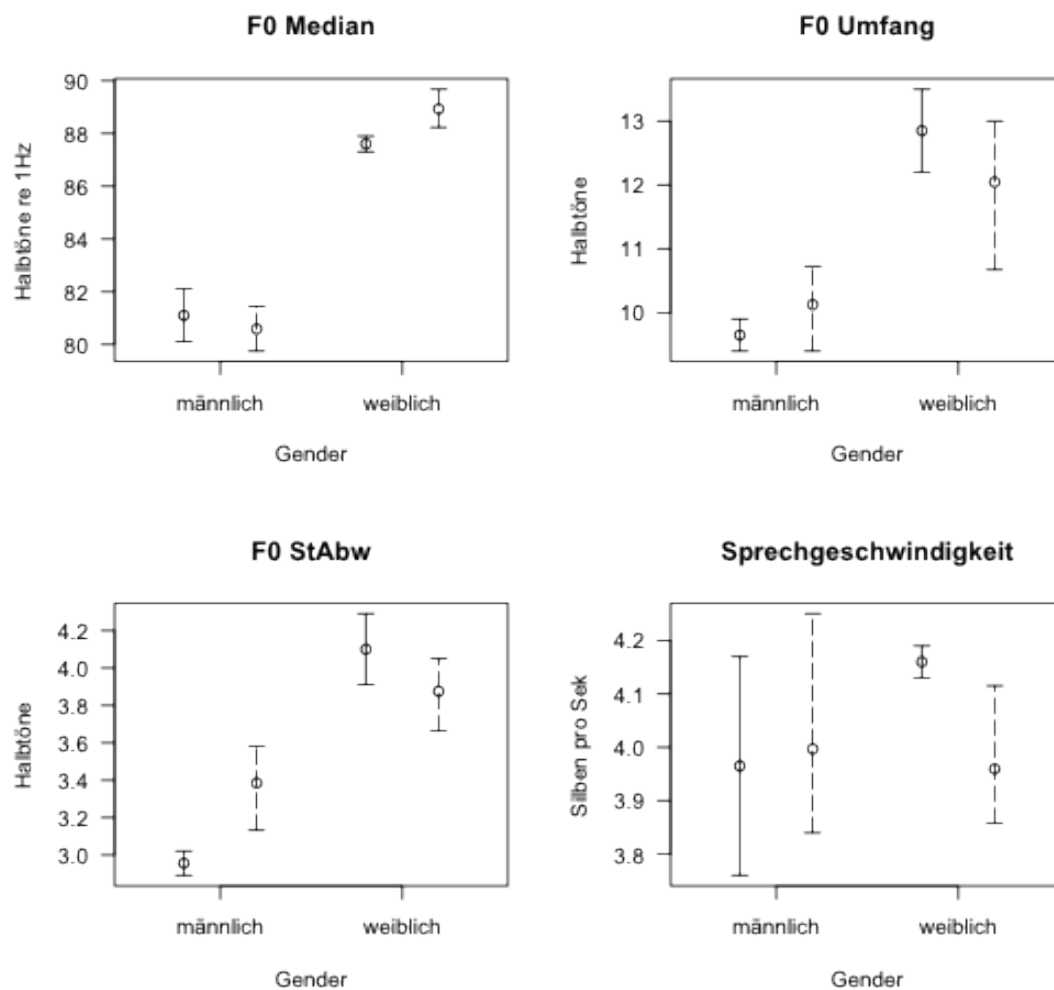


Abb. 5: Testsendungen (links) vs. Nachrichtenarche (rechts, gestrichelt); Mittelwerte mit Konfidenzintervallen

3 Nachrichten als standardsprachliche Referenz

Nachrichten, auch gesprochene Hörfunknachrichten, stellen eine Extremform von Mündlichkeit dar und sind ein klassisches Beispiel für eine „mündlich reali-

sierte schriftkonstituierte Textsorte“ (Gutenberg 2000). Dennoch (oder gerade deswegen) ist die Mündlichkeit in hohem Grade professionell. Dies beinhaltet z.T. hohe Sprechgeschwindigkeiten, komplexe Konstruktionen, hohe Referenzialität (Rekursion) und Abstraktheit, die in dieser Dichte in spontaner Alltagssprache vermutlich kaum vorkommen. Im Allgemeinen fehlt für die Beurteilung solcher Kriterien wie emotionale Neutralität bzw. neutrale Haltung, Lebhaftigkeit, Distanziertheit, Geplantheit vielfach das *tertium comparationis*. Hier kann die institutionalisiert neutrale Präsentation der Nachrichten als alltagsweltliche Referenz dienen, was zumindest uneingeschränkt für die „klassischen“ Nachrichten (Modus 1) gelten kann. Andere Sprechmodi und Subformen der Präsentation im Radio, wie die Berichte und Überleitungen, scheinen einem weniger starken Normierungsdruck von „außen“ wie von „innen“ zu unterliegen.

Letztendlich stellen Nachrichten nicht nur eine allgemein akzeptierte und erwartete Referenz für Standardsprache dar (Hollmach 1996), sie ‚verkörpern‘ in vielen gerade mehrsprachigen Staaten und multidialektalen Gesellschaften Standardsprache schlechthin (Bell 1983). Es ist deshalb nur konsequent, auch für das (Binnen-)Deutsche (öffentlich-rechtliche) Nachrichten bundesdeutscher Programme als Verkörperung oder Realisierung eines Standards anzusehen.

3.1 Regionale Interpretation von Standardaussprache

In Anlehnung an die Konzepte ‚regionaler Standard‘ (Meinhold 1973, Spiekermann 2004) sowie ‚regionaler Gebrauchsstandard‘ (Berend 2004) füge ich hier die Idee einer regionalen Interpretation des überregionalen Standards hinzu, die sich aus der Sprecherperspektive ergibt. Die einzelnen Sprecher unterliegen auch zum Teil senderintern definierten Normen, doch betreffen diese meist explizite Einzelfälle, wie Eigennamen, Fremdwörter o.ä. Selbstverständlich sind orthoepische Auffassung und standardsprachliche Realisierung der professionellen Sprecher von denjenigen von Nicht-Medienprofis zu trennen. Dieser Fokus auf Laien-Auffassungen erfolgt an anderer Stelle, z.B. in solchen Projekten wie dem am Institut für Deutsche Sprache in Mannheim angesiedelten „Deutsch heute“ [<http://www.deutsch-heute.de>] oder dem Augsburger Atlas zur deutschen „Alltagssprache“ [<http://www.philhist.uni-augsburg.de/lehrstuehle/germanistik/>]

sprachwissenschaft/ada/] mit jeweils völlig anderen methodischen Ansätzen als den hier präsentierten.

Sei es im Sinne standardsprachlicher Formstufen (Meinhold 1973, Krech et al. 1982) oder einer auf interpersoneller Distanz basierenden Implementierung des H&H-Modells (Lindblom 1990) oder eines situativ adaptiven *audience design*-Modells (Bell 1984) – es ist nicht überraschend, in den vom Nachrichtenmodus 1 abweichenden Präsentationsformen stärkere Abweichungen von der standardsprachlichen Norm zu finden. Diese dürften zum Teil als Ausdruck der (inszenierten) Spontaneität als auch der Individualität der sprechenden Journalisten zu interpretieren sein (Bell 1982). Außerdem zeigt sich in solch stärkerer Abweichung auch ein Ausdruck der Regionalität, die ja vielfach Teil des zu vermittelnden Senderprofils i.S. einer Identität ist. Diese Identität definiert sich mitunter über eine Abgrenzung zum Hoch-Standard, wie z.B. Bell (1983) für die Frequenz der Reduktionen herausgefunden hat.

Die einzelnen programm- und sprecherspezifischen spektralen Charakteristika der Aufnahmen der NArche sind im Moment noch intransparent. Deshalb soll ein einfach zu erhebender und durch Datenkompression nicht beeinflusster Parameter wie die Stimmeinsatzzeit (*voice onset time* – VOT) betrachtet werden, also die Zeit von der Lösung des Plosivverschlusses bis zum Einsetzen der Stimme im nachfolgenden Vokal. Dabei wird die Auswahl lediglich auf die stimmlosen alveolaren und velaren Plosive im Silben- und Wortanlaut eingeschränkt. Schon von den ersten sprachvergleichenden Untersuchungen (Lisker / Abramson 1964) ist bekannt, dass velare Plosive meist eine längere VOT als alveolare und bilabiale Plosive aufweisen. Allerdings wird die VOT von einer Reihe von Faktoren beeinflusst und hängt unter anderem von der prosodischen Position des Plosivs ab (vgl. Lisker / Abramson 1967 oder Kohler 1995, 157 für Aspiration im Deutschen). Insbesondere ist eine phraseninitiale Verstärkung hier von Bedeutung, u.a. weil sie auch schon an einem Nachrichten-Korpus (Cole et al. 2007) demonstriert wurde. Für die NArche sollte es die weitgehende Homogenität der Sprechgeschwindigkeit innerhalb der Modi (s. oben) möglich machen, sie als Faktor außer Acht zu lassen.

In einem Sample von 55 Textgrids der Jahrgänge 2006/07/08 wurden silbeninitial 487 /t/- und 214 /k/-Fälle, wortinitial 126 /t/- und 205 /k/-Fälle untersucht. Interessanterweise kommt es dabei wortinitial, was in den meisten Fällen des gelabelten NArche-Materials phraseninitial bedeutet, zu einer Verstärkung bzw. Längung (*initial strengthening*) der VOT, aber auch zu einer Art Neutralisierung hinsichtlich des velar-alveolar-Kontrastes (vgl. Abbildung 6). Die Werte wurden dabei in drei quasi-geographische Gruppen (N-Nord, M-Mitte, S-Süd) nach Standort der Sender eingeteilt, doch scheint sich hieraus keinerlei klare Tendenz zu ergeben, lediglich scheint für die ‚süddeutschen‘ Proben die Distanz zwischen den beiden Werten kleiner zu sein. D.h., während für silbeninitiale Plosive VOT eine Rolle spielen kann, kommen wort- und phraseninitial wahrscheinlich eher andere akustische Marken zum Tragen (z.B. Formanttransitionen).

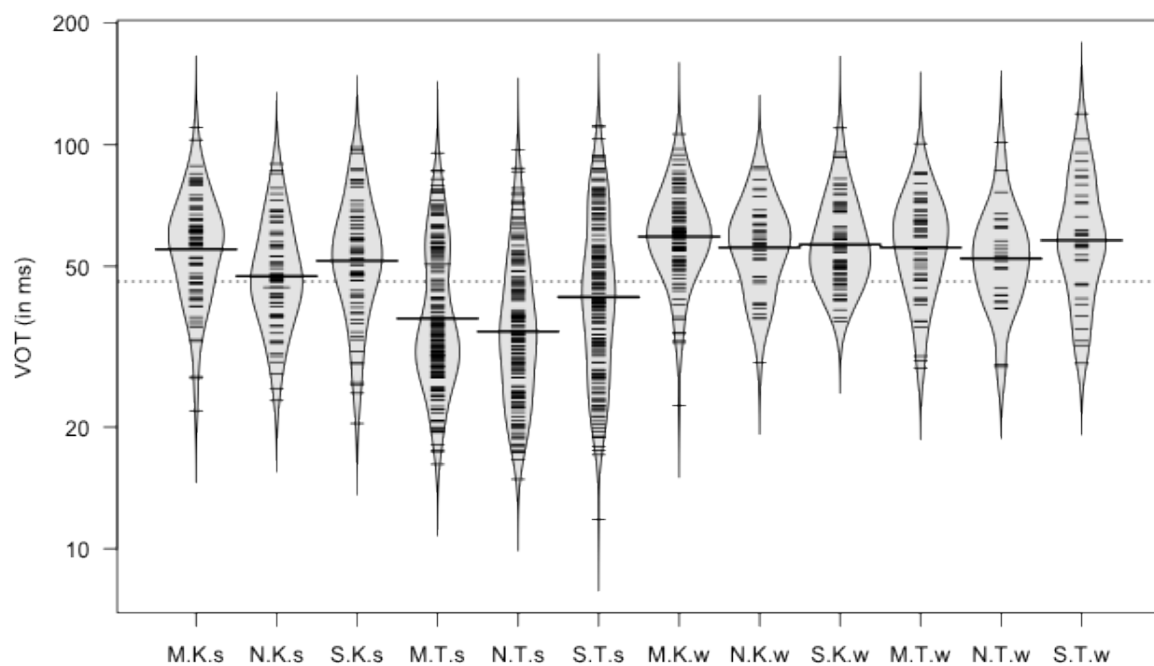


Abb. 6: VOT- Werte für wort(w)- und silben(s)-initiale /t/(T) und /k/(K) gruppiert nach Regionen (N-Nord,M-Mitte,S-Süd)

3.2 Realisierung kanonischer Prosodie am Beispiel von Relativsätzen

Eng verbunden mit der im vorangegangenen Abschnitt betrachteten lautlich-segmentalen Ebene steht auch die suprasegmentale Ebene, insbesondere die Prosodie. Prosodie wird dabei als die Gesamtheit aller suprasegmentalen Mittel

(wie z.B. Tonhöhenbewegung, Lautheitsveränderung, Dauerveränderungen, Stimmqualitätsalternation, Sprechgeschwindigkeit und Pausen) aufgefasst, die jeweils sprachlich relevante Merkmale wie Akzentuierung, Rhythmus, Phrasierung oder Intonation konstituieren. Als Intonation wird im Allgemeinen die Bewegung der Sprechstimmhöhe (Sprechmelodie bzw. deren akustisches Korrelat, die Grundfrequenz F0) verstanden.

Im Rahmen sowohl orthoepischer Vorschläge zur Standardsprache als auch solcher zur Vermittlung von intonatorischen Regeln im Fremdsprachenunterricht des Deutschen (Drach 1937, Trojan 1961, Stock / Zacharias 1982, Stock 1996) sind bereits Aussagen zum Wort- und Satzakzent sowie zur Kontur der Sprechmelodie einfacher Satzarten getroffen worden. Nichtsdestotrotz gibt es bisher keine Kodifizierung einer zumindest für das Binnendeutsche gültigen Standardintonation (vgl. Auer 2003), die auch den Prinzipien einer empirisch gestützten Formulierung folgte. Dabei kann die Realisierung der sprachspezifischen Gestalten regionale Prägungen (Gilles 2005) aufweisen, ungeachtet einer in den meisten Fällen völlig unproblematischen Verständlichkeit oder Grammatikalität. Sind es aber Mittel zur Markierung der Prominenz einzelner Silben oder aber zur Markierung von Weiterweisung, so kann die jeweilige regionale „Färbung“ hervortreten.

Ein sich anknüpfender Aspekt ist die in der Leselehre (vgl. Winkler 1984: 752) zu vermittelnde Realisierung komplexer Sätze (vgl. hierzu z.B. Bose et al. i.d.B.). An dieser Stelle soll beispielhaft die prosodische Realisierung des deutschen Relativsatzes aufgenommen werden. In der bisherigen Forschung (vgl. Birkner 2007), spielt hierzu die wichtige kategorische Unterscheidung aus semantischer Sicht in restriktive und nicht-restriktive (appositive) Relativsätze. Dabei wurde von mehreren Autoren beobachtet, dass tendenziell mit den restriktiven Relativsätzen ein prosodischer Einschnitt einhergeht (vgl. Abbildung 7), während bei den nicht-restriktiven Relativsätzen kein solcher Einschnitt erfolgt, sondern vielmehr eine prosodische Einbettung. Allerdings folgen in dem spontansprachlichen Korpus bei Birkner nur 64% Prozent der nicht-restriktiven Relativsätze dem für diese Kategorie als prototypisch beschriebenen prosodischen

Format, und immerhin 19% zeigen dieses Format auch für die restriktiven Relativsätze. Alle anderen Belege weisen dazwischenliegende Format-Stufen auf.

Von den 73 Relativsätzen, die in den Jahrgängen 2008/09 durch das Vorhandensein der Relativmarker *die, der, das, den, welche, wer, wo, was* erfasst wurden, sind nach einer ersten Zuordnung 27 nicht-restriktiv und 45 restriktiv. Von den restriktiven Relativsätzen zeigen 25 einen klaren prosodischen Einschnitt, während es bei den nicht-restriktiven Relativsätzen 14 sind. D.h. dass sich weder hinsichtlich des Verhältnisses zwischen ‚prosodisch integrierten‘ und ‚nicht-integrierten‘ Realisierungen eine signifikant verschiedene Verteilung der Häufigkeiten ergibt ($\chi^2=0.0239$, $df=1$, $p=0.87$), als auch hinsichtlich ihres Verhältnisses generell ($\chi^2=0.0037$, $df=1$, $p=0.95$). Daher können die für die wesentlich inhomogenere Spontansprache erhobenen Befunde von Birkner für die gesprochenen Nachrichten und Berichte nicht bestätigt werden. Sowohl eine präzisere Untersuchung, die die Stärke des Einschnitts einbezieht, aber auch eine präzisere Typisierung der Relativsätze vornimmt, einschließlich der Länge der Teilsätze, sind derzeit in Vorbereitung.

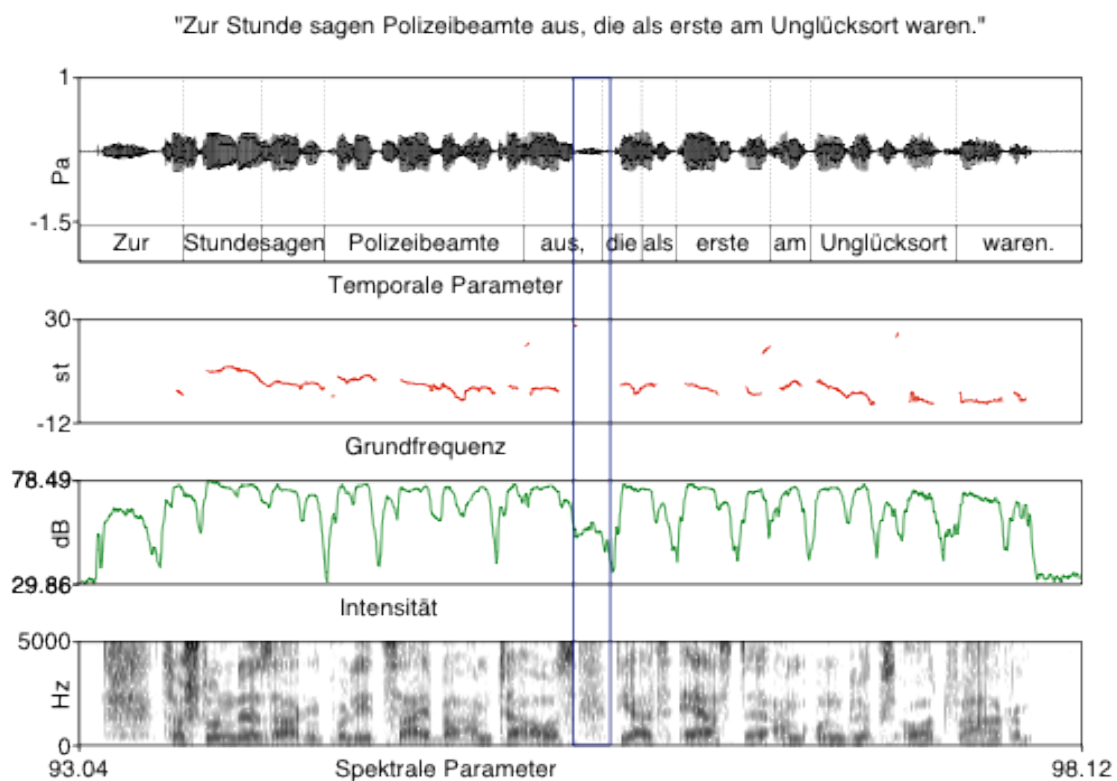


Abb. 7: Beispiel (NA08) für einen Relativsatz mit stärkerem prosodischem Einschnitt

4 Wohin geht die phonetische Forschung an der Nachrichten-Arche?

Das NArche-Korpus stellt im Moment eine wiederholte Querschnittsstudie mit einer starken Tendenz zur Überlappung von Sprecherebene und Sender- bzw. Programmebene dar. D.h., ein Sprecher bzw. eine Sprecherin steht häufig *pars pro toto* für das Programm oder den Sender in jenem Jahrgang. Es stellt sich deshalb die Frage nach einer möglichen höheren Zahl von Datenerhebungen pro Jahr. Damit ist selbstverständlich ein erhöhter Aufwand bei der Erstellung der Metadaten verbunden, doch können diese zumindest nachträglich ergänzt werden, was auf die Primärdaten nicht zutrifft. So bedarf es eines großen Aufwandes, bis die phonetischen Annotationen und Transkriptionen die Tiefe, z.B. Etikettierung bis auf Segmentebene erreichen, wie die eines Kiel-Korpus [<http://www.ipds.uni-kiel.de/forschung/kielcorpus.de.html>]. Hinsichtlich senderbezogener Faktoren scheint die geographische Dichte der Sender zwar durchaus ausreichend, doch können weitere Sender aus der Schweiz, Österreich, Luxemburg, Liechtenstein und Sender anderer deutschsprachiger Regionen das Bild der deutschsprachigen Nachrichten vervollständigen. Kontrastiv dazu stünden Nachrichten in anderen Sprachen, wie sie etwa die Dienste von WDR Funkhaus Europa oder der Deutsche Welle anbieten. Und nicht zuletzt wird das Bild der Nachrichten nicht nur von den öffentlich-rechtlichen, sondern auch von den privaten Rundfunkanstalten geprägt. Deshalb wäre auch ein möglicherweise kleineres Sample aus solchen Sendern eine notwendige Ergänzung, um das Gesamtspektrum der Nachrichtensendungen zu repräsentieren.

Nachrichtensendungen sind eingebettet in ein Gesamtkonzept einer Sendezeit und z.T. stark verflochten mit den davor und danach liegenden Programmelementen. Eng damit verknüpft ist der noch zu untersuchende Einfluss von kompressionsbehafteten Aufnahmeformaten, Voice-Keys und anderen Layout- und Verpackungsfiltren auf die im Vorangegangenen betrachteten Parameter. Voice-Keys sind z.T. sprecherspezifische Klangfilter, die zum einen den gewünschten Senderklang hervorheben, aber auch sprecherische „Eigenheiten“ z.B. der Zischlaute „verdecken“. Bei einem Vergleich etwa spektraler Parameterausprägungen

zwischen den Sendern sowohl von segmentalen wie suprasegmentalen Eigenschaften muss dies zumindest berücksichtigt und geprüft werden.

5 Literaturverzeichnis

- Abercrombie, D. (1967): *Elements of general phonetics*. Chicago.
- Baayen, R. / Davidson, D. / Bates, D. (2008): Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59(4):390-412.
- Bates, D. / Maechler, M. / Dai, B. (2008): *lme4: Linear mixed-effects models using Eigen and Eigen++*. R package version 0.999375-28. <http://www.r-project.org>.
- Bell, A. (1982). Radio: The style of news language. *Journal of Communication*, 32(1):150–164.
- Bell, A. (1984). Style as audience design. *Language in society* 13 (2), 145-204.
- Benton, M. / Dockendorf, L. / Jin, W. / Liu, Y. / Edmondson, J. (2007): The continuum of speech rhythm: computational testing of speech rhythm of large corpora from natural Chinese and English speech. In: *Proceedings of the 16th ICPHS*. Saarbrücken, 1269-1272.
- Birkner, K. (2007): *Relativ(satz)konstruktionen im gesprochenen Deutsch: Syntaktische, prosodische, semantische und pragmatische Aspekte*. Berlin.
- Boersma, P. / Weenink, D. (2011): *Praat: doing phonetics by computer* [Computer program]. Version 5.2.12, retrieved 28 January 2011 from <http://www.praat.org/>.
- Bose, I. / Gutenberg, N. / Ohler, J. / Schwiesau, D. (i.d.B.): *Testmaterial zur Hörverständlichkeit von Radionachrichten: Theoretische und methodische Grundlagen*. In: Bose, I. / Schwiesau, D. (Hg.): *Nachrichten schreiben, sprechen, hören. Forschungen zur Hörverständlichkeit von Radionachrichten*. Berlin.
- Cole, J. / Kim, H. / Choi, H. / Hasegawa-Johnson, M. (2007): Prosodic effects on acoustic cues to stop voicing and place of articulation: Evidence from radio news speech. *Journal of Phonetics*, 35(2):180-209.
- Dellwo, V. / Wagner, P. (2003): Relations between language rhythm and speech rate. In: *Proceedings of the 15th international congress of phonetics sciences*, 471–474, Barcelona.
- Drach, E. (1937): *Grundgedanken der deutschen Satzlehre*. Frankfurt / M.
- Gilles, P. (2005): *Regionale Prosodie im Deutschen: Variabilität in der Intonation von Abschluss und Weiterweisung*. Berlin.
- Grabe, E. / Low, E. (2002): Acoustic correlates of rhythm class. *Laboratory phonology* 7, 515-546.
- Grandke, V. (2009): Der Bericht in Hörfunknachrichten. In: Anders, L. C. / Bose, I. (Hg.): *Aktuelle Forschungsthemen der Sprechwissenschaft (Band I): Sprach-, Sprech- und Stimmstörungen / Sprache und Sprechen von Hörfunknachrichten*. (HSSP 30). Frankfurt / M. u. a., 129-143.
- Grawunder, S. / Bose, I. / Hertha, B. / Trauselt, F. / Anders, L.C. (2006): Perceptive and acoustic measurement of average speaking pitch of female and male speakers in German radio news. *Proceedings of the Interspeech 2006 — ICSLP, Pittsburgh 9/2006*, 885-888.
- Grawunder, S. / Bose, I. (2008): Average speaking pitch vs. average speaker fundamental frequency—reliability, homogeneity, and self report of listener groups. *Proceedings of the International Conference Speech Prosody 2008, Campinas Brazil 5/2008*, 763-766.
- Grawunder, S. / Engert, H. / Kaiser, H. / Unger, S. / Bose, I. (2008): Pitch, speech rate and rhythm as speaking style characterizing parameters in German news presentation. Poster

- auf EASR08 - Workshop on Empirical Approaches to Speech Rhythm am UCL, London. unveröff. Mskript.
- Grehl, M. / Kupietz, M. (2009): Systematisierung und Verwaltung der Nachrichtenarche der ARD. Bachelor-Abschlussarbeit am Seminar für Sprechwissenschaft, Martin-Luther-Universität Halle-Wittenberg. unveröff. Mskript.
- Großes Wörterbuch der deutschen Aussprache (GWdA)(1982): Hg.: Krech, E.-M. / Kurka, E. / Stelzig, H. / Stock, E. / Stötzer, U. / Teske, R., unter Mitwirkung von Jung-Alsen, K.. Leipzig.
- Gutenberg, N. (2000): Mündlich realisierte schriftkonstituierte Textsorten. In: Brinker, K. / Antos, G. / Heinemann, W. / Sager, S. (Hg.): Text- und Gesprächslinguistik (HSK 16.1), 574-587.
- Hollmach, U. (1996): Eine soziophonetische Untersuchung als Grundlage zur Neukodifizierung der deutschen Standardaussprache. In: Lemke, S. / Thiel, S. (Hg.): Sprechen – Reden – Mitteilen. München / Basel, 211-217.
- Hurlbert, S. (1984): Pseudoreplication and the design of ecological field experiments. *Ecological monographs* 54(2), 187-211.
- Klaaß, D. (2008): Untersuchungen zu ausgewählten Aspekten des Konsonantismus bei österreichischen Nachrichtensprechern. Frankfurt / M.
- Kohler, K. J. (1995): Einführung in die Phonetik des Deutschen. (Grundlagen der Germanistik 20). Berlin.
- König, W. (1989): Atlas zur Aussprache des Schriftdeutschen in der Bundesrepublik Deutschland. Ismaning.
- Krech, Eva-Maria u.a. (2009): Deutsches Aussprachewörterbuch (DAWB). Berlin / New York.
- Kuzla, C. / Ernestus, M. (2011): Prosodic conditioning of phonetic detail in German plosives. *Journal of Phonetics*, im Druck, Corrected Proof.
- Lindblom, B. (1990): Explaining phonetic variation: a sketch of the H&H theory. In: Hardcastle, W. J. / Marchal, A. (Hg.): *Speech production and speech modeling*. vol. 55. Dordrecht, 403-439.
- Lisker, L. / Abramson, A. (1964): A cross-language study of voicing in initial stops: acoustic measurements. *Word*, 20(3), 384-422.
- Meinhold, G. (1973): Deutsche Standardsprache – Lautschwächungen und Formstufen. Jena.
- Pike, K. L. (1945): *The intonation of American English*. Michigan.
- R Development Core Team (2009): *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna.
- Schwiesau, D. / Ohler, J. (2003): *Die Nachricht in Presse, Radio, Fernsehen, Nachrichtenagentur und Internet – Ein Handbuch für Ausbildung und Praxis*. München.
- Schwiesau, D. / Grawunder, S. / Bose, I. (i.d.B.): *Die Nachrichtenarche der ARD*. In: Bose, I. / Schwiesau, D. (Hg.): *Nachrichten schreiben, sprechen, hören. Forschungen zur Hörverständlichkeit von Radionachrichten*. Berlin.
- Sendlmeier, W. (Hg.) (2005): *Sprechwirkung – Sprechstile in Funk und Fernsehen*. Berlin. (Mündliche Kommunikation 3).
- Spiekermann, H. (2005): Regionale Standardisierung, nationale Destandardisierung. In: Eichinger, L. / Kallmeyer, W. / Institut für deutsche Sprache (Hg.): *Standardvariation: Wie viel Variation verträgt die deutsche Sprache? IDS Jahrbuch 2004*. Mannheim, 100-125.
- Stock, E. / Zacharias, C. (1982): *Deutsche Satzintonation*. Leipzig.
- Trojan, F. (1961). *Deutsche Satzbetonung, Grundsätze und Übungen: Ein Lehr- und Übungsbuch für In- und Ausländer*. Wilhelm Braumüller: Stuttgart, Wien.
- Winkler, C. (1973): *Die Klanggestalt des Satzes*. In: Duden Bd. 4, Grammatik. Mannheim, Wien, Zürich, 637-666.

Danksagung

Ich möchte mich herzlich für die Möglichkeit der Arbeit an diesem Korpus bei ihrem Initiator und Vertreter der ARD, Dietz Schwiesau, bedanken. Des Weiteren danke ich ausdrücklich für das Beherbergen der digitalen Audio-Kopien sowie für die infrastrukturelle Unterstützung beim Erstellen der Metadaten in der Arbeitsgruppe Linguistik am Max-Planck-Institut für evolutionäre Anthropologie in persona Prof. Bernard Comrie. Ganz besonderer Dank gilt auch den Praktikant/-inn/-en Katharina Sturm, Josephine Rocholl, Rainer Schmitz, Helen Kaiser, Stefanie Unger, Caroline Pfeifer, den Bakkalaureatinnen Maxi Grehl, Michaela Kupietz und Luise Dölle sowie den wissenschaftlichen Hilfskräften Hans Nenoff (Engert) und Stefanie Dietz und weiteren Assistentinnen wie Anna Schwenke, Caroline Schniggenfittig und Natalie Fecher, die bisher alle an der Aufbereitung der Rohdaten für dieses Korpus mitgearbeitet haben.

Dr. phil. Sven Grawunder
Max-Planck-Institut für Evolutionäre Anthropologie
AG Linguistik
Deutscher Platz 6
04103 Leipzig
grawunder@eva.mpg.de