



# Sprachniveau in Online- und TV-Nachrichten: Eine quantitative computergestützte Textanalyse der Online- und TV-Berichterstattung von ARD, ZDF, Sat.1 und RTL

Benjamin P. Lange\*, Aylin Bayirli und Frank Schwab

Julius-Maximilians-Universität Würzburg, Fakultät für Humanwissenschaften, Institut Mensch-Computer-Medien, Arbeitsbereich Medienpsychologie, Germany

\*Corresponding author: [benjamin.lange@uni-wuerzburg.de](mailto:benjamin.lange@uni-wuerzburg.de)

## Abstract

Nachrichten stellen ein beliebtes Forschungsfeld der Kommunikationsforschung dar. In der entsprechenden deutschsprachigen Forschung werden vielfach die Hauptnachrichten der beiden deutschen Rundfunksysteme (öffentlich-rechtlich vs. privat) miteinander verglichen – zuletzt z. B. hinsichtlich ihres Sprachniveaus. Allerdings existiert wenig Forschung, die sich dem Online-Nachrichtenangebot widmet. Daher lag unser Ziel in der Analyse des Sprachniveaus von deutschen Online-Nachrichtenangeboten beider Systeme sowie einzelner Sender (ARD, ZDF, RTL SAT.1) mit Hilfe von vier Kategorien des Textanalyseprogramms Linguistic Inquiry and Word Count sowie von zwei weiteren Massen (Umgangssprache, Fleisch-Index), die unterschiedliche Dimensionen des Sprachniveaus erfassen. In Form einer künstlichen Nachrichtenwoche wurden insgesamt 84 Online-Nachrichtentexte hinsichtlich ihres Sprachniveaus analysiert. Online-Nachrichten der Privatsender wiesen durchschnittlich längere Texte und mehr Wörter pro Satz auf. Ein Unterschied hinsichtlich Wortschatzdiversität und Wortkomplexität konnte nicht gefunden werden. Die Ergebnisse zum Fleisch-Index zeigen, dass Online-Nachrichten beider Systeme ähnlich schwer verständlich sind. Diese Ergebnisse wurden aktuellen Daten zum Sprachniveau klassischer TV-Nachrichten gegenübergestellt. Dabei wurde evident, dass Online-Nachrichten höhere Werte bei einigen Markern des Sprachniveaus aufwiesen, was u. a. mit Hilfe des Kontinuums Mündlichkeit-Schriftlichkeit kontextualisierbar ist, und dass sich das Sprachniveau in Online-Nachrichten zwischen den Systemen und Sendern, im Unterschied zu TV-Nachrichten, eher ähnelte, was im ersten Fall als Konvergenz, im zweiten als Konkurrenz interpretiert werden kann.

## Keywords

Online-Nachrichten, Sprachniveau, sprachliche Gewandtheit, Eloquenz, quantitative computergestützte Textanalyse

## 1 Einleitung

Nachrichten erfüllen in modernen Informationsgesellschaften eine zentrale Aufgabe bei der Vermittlung von tagesaktuellen Informationen (z. B. Schwiesau & Ohler, 2016; Unz & Schwab, 2004). Eine spezielle Rolle wird hierbei dem Fernsehen zugeschrieben, da es sich um das am meisten genutzte Medium handelt, durch welches die Mehrheit der Bürger<sup>1</sup>

in Deutschland adressiert werden kann (Meckel & Kamps, 1998; Unz & Schwab, 2004; Zubayr & Gerhard, 2018). Das Fernsehen kann damit als zentrale Informationsquelle der Gesellschaft angesehen werden (Meckel & Kamps, 1998; Zubayr & Gerhard, 2018).

Nachrichten, insbesondere TV-Nachrichten, stellen daher einen beliebten und viel untersuchten Gegenstandsbereich

und der einfacheren Lesbarkeit und impliziert gleichermassen die weibliche Form («Bürgerin»). Es sind stets Personen jeden Geschlechts gleichermassen gemeint.

1 Die vorliegende Arbeit verwendet das generische Maskulinum (z. B. «Bürger») aus Gründen der sprachlichen Vereinfachung



der Medien- und Kommunikationsforschung dar (Meckel & Kamps, 1998; Unz & Schwab, 2004; Wittwen, 1995). Vielfach wurden in entsprechenden deutschen Forschungsarbeiten die Nachrichtenformate der beiden deutschen Rundfunksysteme (öffentlich-rechtliche vs. private Sender; im Folgenden meist kurz: System) miteinander verglichen (z. B. Krüger, 1998; Lange, Kouros & Schwab, 2019; Weiss, 2003; Wittwen, 1995).

Unter anderem wurden TV-Nachrichten dabei hinsichtlich ihres Sprachniveaus untersucht (z. B. Krüger, 2009; Lange et al., 2019). Dabei wurde u. a. deutlich, dass sich die TV-Nachrichten der öffentlich-rechtlichen Nachrichtensender (im Folgenden: ÖRS) durch ein in vielerlei Hinsicht höheres Sprachniveau auszeichnen als die der Privatsender (im Folgenden PS), wobei insbesondere die Tagesschau durch ihr hohes Sprachniveau herausstach.

Nachrichten werden mittlerweile und zunehmend aber auch online rezipiert (z. B. Arnold, 2016). Während jedoch Forschung zur sprachlichen Gestaltung klassischer TV-Nachrichten vorliegt (z. B. Lange et al., 2019), mangelt es an derartiger Forschung zu Online-Nachrichten. An dieser Forschungslücke setzt der vorliegende Beitrag an.

Zunächst legen wir den Forschungsstand zu Online- und TV-Nachrichten dar. Diesbezüglich beginnen wir mit einer Darlegung des Selbstverständnisses der öffentlich-rechtlichen verglichen mit dem der privaten Nachrichtensender. Im Anschluss gehen wir detailliert auf bisherige Ansätze und Forschungsergebnisse zur sprachlichen Gestaltung von Nachrichten ein. Daran anknüpfend thematisieren wir den Unterschied zwischen Information und Unterhaltung in Nachrichtenformaten und wie sich dieser Unterschied sprachlich ausdrückt. Schliesslich gehen wir auf Online-Nachrichten ein, um aus dem Dargelegten die Forschungsfrage, ob sich Unterschiede im Sprachniveau zwischen den beiden Rundfunksystemen und zwischen den einzelnen Sendern in Online-Nachrichten finden lassen, sowie eine Hypothese abzuleiten, die dann empirisch untersucht wurden.

## 2 Forschungsstand zu Online- und TV-Nachrichten

In diesem Kapitel legen wir den für unsere Arbeit relevanten Forschungsstand zu Online- und TV-Nachrichten dar. Schwerpunkte sind dabei grundsätzliche Unterschiede zwischen den beiden Rundfunksystemen und einzelnen Sendern sowie die sprachliche Gestaltung verschiedener Nachrichtenangebote.

### 2.1 Das Selbstverständnis der öffentlich-rechtlichen und privaten Nachrichtensender

Die TV-Nachrichtenangebote der beiden Systeme können hinsichtlich ihres Selbstverständnisses bezüglich des Informations- und Unterhaltungsgehalts differenziert werden. Die Programmrichtlinien der ÖRS unterliegen in Deutschland, anders als die der PS, dem Rundfunkstaatsvertrag. Ersteren kann deshalb eine gewisse Pflicht unterstellt werden, ihren Zuschauern möglichst neutrale Informationen darzubieten, die einen gleichfalls möglichst objektiven Überblick über das Weltgeschehen vermitteln, zur Meinungsbildung beitragen und damit zur Teilnahme am politischen Geschehen befähigen (Halff, 1998; Unz & Schwab, 2004). Trotz einiger Kritik an den ÖRS haben Zuschauer ein hohes Vertrauen in deren Sendungen und sprechen diesen eine hohe Glaubwürdigkeit zu, was sich u. a. an der gleichbleibend grossen Reichweite dieser Nachrichtenformate zeigt (Krüger & Zampf-Schramm, 2018; Newman, Fletcher, Kalogeropoulos, Levy & Nielsen, 2018; Unz & Schwab, 2004; Zubayr & Gerhard, 2018).

Demgegenüber steht das Selbstverständnis der PS, das sich vor allem über die Nähe zum Zuschauer definieren lässt (Krüger, 1998); sie zielen eher darauf ab, bestimmte Grundbedürfnisse der Bürger anzusprechen, sowie Nachrichten für alle Zuschauer verständlich zu vermitteln (Krüger, 1998). So stellen PS bewusst Themen aus dem Alltagsgeschehen, von denen Menschen persönlich betroffen sind, sowie zwischenmenschliche Beziehungen in den Mittelpunkt der Berichterstattung (Kamps, 1998; Krüger, 1998;

Krüger & Zampf-Schramm, 2018; Unz & Schwab, 2004). Dies soll den Zuschauern vermutlich ein breites «Empathieangebot» (Kamps, 1998, S. 43) unterbreiten, so dass sie sich leichter mit den gezeigten Inhalten identifizieren können und somit emotional involviert werden (Krüger & Zampf-Schramm, 2018). Unz und Schwab (2004) sprechen in diesem Kontext von einer Einflussnahme auf das Rezeptionserleben durch eine emotionale und moralische Adressierung des Rezipienten. Insgesamt kann insbesondere bei den Nachrichten der PS von einer Unterhaltungsorientierung gesprochen werden, die vielfach mit den Begriffen des Infotainments und der Boulevardisierung in Verbindung gebracht wird (Bartel, 1997; Bruck & Stocker, 1996; Donsbach & Büttner, 2005; Wittwen, 1995).

## 2.2 Nachrichten und ihre sprachliche Gestaltung

Die unterschiedlichen Philosophien der beiden Systeme wirken sich darauf aus, wie versucht wird, auf das Rezeptionserleben der Zuschauer Einfluss zu nehmen, d. h. darauf, wie der Zuschauer adressiert wird (Unz & Schwab, 2004). Dies kann auf visuell-bildlicher, aber auch auf sprachlicher Ebene geschehen (z. B. Brosius, 1998; Burger & Luginbühl, 2014; Machill, Köhler, Waldhauser, 2006; McQuail, 2001; Paetzold, 1973; Semetko & Valkenburg, 2000; White, 1950).

Hinsichtlich der sprachlichen Ebene hat die vorherige Forschung (Lange et al., 2019) u. a. das sog. Sprachniveau von TV-Nachrichten in den Blick genommen, d. h. die Frage verfolgt, wie sprachlich gewandt bzw. eloquent sich die Nachrichtensprache darstellt. Sprachniveau wird hierbei als wertfreier Begriff verstanden, der sich als (partiell) synonym zu sprachlicher Gewandtheit sowie Eloquenz auffassen und sich u. a. über Wortschatzdiversität (Anteil unterschiedlicher Wörter), Wortkomplexität (grob: Länge der Wörter) und Satzkomplexität (grob: Länge der Sätze) definieren lässt (s. Lange, 2012; Lange, 2016; Lange, Hennighausen, Brill & Schwab, 2016; Lange, Bögemann & Zaretsky, 2017; s. dazu auch weiter u.). Die bei-

den dabei wesentlich verfolgten Fragen lauteten: Ist die Nachrichtensprache eher elaboriert oder eher einfach und schlicht? Und lassen sich diesbezüglich Unterschiede einerseits zwischen den Angeboten der beiden Systeme und andererseits der einzelnen Sender (d. h. ARD, ZDF, RTL und Sat.1; im Folgenden meist kurz: Sender) feststellen?

In der Gesamtschau kann von einem höheren Sprachniveau der ÖRS die Rede sein (Lange et al., 2019; s. auch Wittwen, 1995). Lange et al. (2019) konnten mit einer quantitativen computergestützten Textanalyse der Sprache in den Hauptnachrichtensendungen der vier Sender mittels der Textanalyse-Software Linguistic Inquiry and Word Count (LIWC; s. Wolf, Horn, Mehl, Haug, Pennebaker & Kordy, 2008) zeigen, dass die ÖRS moderat effektstark komplexere Wörter und hoch effektstark komplexere Sätze verwendeten als die PS. Der erstgenannte Befund kam dabei fast ausschliesslich durch die *Tagesschau* zustande, d. h., dass *heute* diesbezüglich eher auf dem Niveau der PS lag (vgl. Wahdat, 2009). Lediglich bei der Wortschatzdiversität lagen die PS über den ÖRS, wobei dieser Unterschied überwiegend durch die *SAT.1 Nachrichten* zustande zu kommen schien. Die Autoren vermuteten als Erklärung für das höhere Sprachniveau der PS bei diesem speziellen Mass u. a. einen höheren Anteil an Umgangssprache in den *SAT.1 Nachrichten*.

## 2.3 Information und Unterhaltung in Nachrichtenformaten und ihre sprachlichen Korrelate

In einer sprachlichen Analyse von TV-Nachrichten identifizierte Wittwen (1995) verbal-auditive Unterhaltungselemente des Infotainments, wie z. B. der Einsatz von Umgangssprache (s. dazu auch Mangold, 2004). Damit im Einklang konnte er in den Nachrichtenangeboten der PS einfache Sprache nachweisen, die sich u. a. durch den Einsatz von Umgangssprache und kurzen Sätzen auszeichnet (s. dazu auch Landmeier & Daschmann, 2011). Auch Arnold (2016) zeigte, dass Boulevardmedien einfache Sprache, bis hin zur Umgangssprache, verstärkt einsetzen. Damit

scheint der Einsatz von Umgangssprache ein weiteres untersuchungswürdiges Mass für Sprachniveau in Nachrichten zu sein.

Die bisherigen Befunde zeigen somit, dass das Sprachniveau mit der inhaltlichen Ausrichtung der diversen Formate und ihrer Philosophien nicht unkorreliert zu sein scheint: Die *Tagesschau* als hochseriös und glaubwürdig wahrgenommene Instanz der deutschen Nachrichtenwelt (Bartel, 1997; Matzen, 2009; Meckel & Kamps, 1998; Wahdat, 2009) weist bei zwei zentralen Massen des Sprachniveaus (Wort- und Satzkomplexität) die höchsten Werte auf (Lange et al., 2019). Die Sendung *heute* liegt nur bei der Satzkomplexität auf dem Niveau der *Tagesschau*, ansonsten aber eher auf dem Niveau der PS (Lange et al., 2019). Dies lässt sich damit erklären, dass sich *heute* insofern quasi in einer Zwischenwelt zwischen der *Tagesschau* und den Nachrichten der PS befindet, als sie u. a. auch unterhaltungsorientiert ausgerichtet ist (Bartel, 1997).

#### 2.4 Online-Angebote der Nachrichtensender

Die bisher referierten Befunde beziehen sich allesamt auf klassische TV-Nachrichten. Das Einholen aktueller Informationen erfolgt jedoch mittlerweile in starkem Masse auch Online; Nachrichten werden mittlerweile auch Online «in Echtzeit» (Arnold, 2016, S. 11) rezipiert. Der technische Fortschritt mit dem Entstehen und der Verbreitung des Internets konnte nicht folgenlos für den Nachrichtenjournalismus sein (Arnold, 2016; Meckel & Kamps, 1998; Rankl, 2014; Unz & Schwab, 2004). Neben Herausforderungen bieten sich auch neue Möglichkeiten der Verbreitung von Nachrichten (s. z. B. Arnold, 2016; Hölig & Hasebrink, 2018; Meckel & Kamps, 1998; Neuberger, 2018; Rankl, 2014; Stanyer, 2008).

Der Einfluss des Internets lässt sich in der Tat als ein Wandel in der Nachrichtenlandschaft auffassen (Meckel & Kamps, 1998; Rankl, 2014; Unz & Schwab, 2004). Durch das Internet bieten sich für den Nachrichtenjournalismus neue Möglichkeiten der Veröffentlichung und der Adressierung der Rezipienten (Rankl, 2014). Weiss (2003) konnte etwa zeigen, dass

sich besonders Nachrichtensender die entsprechende Ausweitung der medialen Kanäle zunutze machen, was sich in der Ausdifferenzierung ihres Online-Angebots feststellen liess: TV-Nachrichtenanbieter bieten sowohl auf Senderebene (z. B. ard.de, zdf.de) als auch auf Sendungsebene (z. B. tagesschau.de, heute.de) ein spezifisches Internetangebot zur Informationsvermittlung an. Das Internet stellt somit gerade für Nachrichtensender wie auch rezipienten ein wichtiges Zusatzangebot dar und ist insbesondere für die Adressierung eines jüngeren Publikums der wichtigste Informationskanal (Hölig & Hasebrink, 2018; Schneller, 2017; Weiss, 2003).

#### 2.5 Forschungslücke und vorliegende Studie

Trotz einiger Studien zur Nachrichtensprache lässt sich eine Forschungslücke identifizieren: Erstens liegen zum jetzigen Zeitpunkt kaum quantitativ-empirische Analysen des Online-Informationsangebots von Fernsehnachrichten i.A. vor, die die bereits grosse Zahl an Studien mit qualitativem Ansatz zur Fernsehnachrichtenanalyse (z. B. Bruns & Marcinkowski, 1997; Krüger & Zampf-Schramm, 2018; Uribe & Gunter, 2007; Weiss, 2003) sinnvoll ergänzen würden. Das heisst, es besteht insbesondere hinsichtlich des Sprachniveaus des Online-Nachrichtenangebots der beiden Systeme eine Forschungslücke. Unseres Wissens existiert keine empirische Forschung, die die Sprache an sich und das Sprachniveau im Besonderen von Online-Nachrichten quantitativ analysiert und hinsichtlich dieser Analyseebene klassischen Fernsehnachrichten gegenübergestellt hat. Dabei könnte der (crossmediale) Produktionsprozess in Abhängigkeit vom Medium eine unterschiedliche sprachliche Aufbereitung einfordern: Online-Nachrichten könnten – als anderes neues Medium – andere Charakteristika aufweisen und eigenen Regeln folgen als ihre TV-Pendants (vgl. Lange et al., 2019); und dies könnte sich auch auf die sprachliche Gestaltung in Form des Sprachniveaus auswirken.

In Form einer aktuellen Publikation (Lange et al., 2019) liegen aktuelle Da-

ten zum Sprachniveau von klassischen TV-Nachrichten vor, die ein teils deutlich höheres Sprachniveau der Nachrichten der ÖRS dokumentiert (s.o.). Auf Basis dieser Befunde lässt sich fragen, ob sich die teils auffallend grossen Unterschiede im Sprachniveau von TV-Nachrichten zwischen den beiden Systemen bzw. zwischen den einzelnen Sendern auch in den Online-Pendants finden lassen (Lange et al., 2019): Werden die sprachlichen Unterschiede also von der Offline- in die Online-Welt transferiert? Zeigt sich z. B. die Sonderstellung der *Tagesschau* nicht nur Offline, sondern auch Online? Zwar mögen die Online-Redaktionen nicht deckungsgleich mit den Redaktionen sein, die die TV-Inhalte liefern; wenn die sprachliche Gestaltung in Form des Sprachniveaus jedoch als Ausdruck einer bestimmten journalistischen Kultur und Philosophie eines Senders oder als wiedererkennbares Markenmerkmal verstanden wird (vgl. Lange et al., 2019), wäre ein solcher Transfer durchaus denkbar. Ist also die Selbstinszenierung der Anbieter bzgl. des Sprachniveaus crossmedial gleichbleibend?

Oder stellte das Internet sozusagen «Alles auf Null» und hat damit im Zuge des Wandels von «Offline Only» zu «Online Too» zu einer Konvergenz zwischen den beiden Systemen und den verschiedenen Sendern geführt (vgl. Neuberger, 2018; Weiss, 2003)? Werden also aufgrund der relativen Neuartigkeit und der eigenen Regeln der Online-Welt derartige System- und auch Senderunterschiede nivelliert? Die Fragen lauten also, inwiefern sich das Sprachniveau je nach Medium (Online vs. TV) unterscheidet und ob und wenn ja, wie versucht wird, ein bestimmtes Sprachniveau, z. B. als Art Markenmerkmal, beim crossmedialen Produzieren wiedererkennbar zu gestalten.

Beim Vergleich von klassischen TV mit Online-Nachrichten sticht darüber hinaus eine wesentliche Unterscheidung ins Auge, nämlich die zwischen Mündlichkeit und Schriftlichkeit (Koch & Oesterreicher, 1985). Zwar liegen nicht nur Online-Nachrichten, sondern auch die Inhalte der TV-Nachrichten zunächst schriftlich vor; diese werden jedoch mündlich vorgetra-

gen. Gleichwohl können diese hinsichtlich des konzeptionellen Aspekts eher der Schriftlichkeit zugeordnet werden, da ihr Duktus einer eher gewählten Ausdrucksweise entspricht: Sie sind zwar mündlich, weisen aber Charakteristika der Schriftlichkeit auf (vgl. Koch & Oesterreicher, 1985, S. 15). Hinsichtlich des medialen Aspekts wird der Unterschied zwischen beiden Nachrichtenformen jedoch deutlich: So sind Online-Nachrichten grafisch, TV-Nachrichten hingegen (überwiegend) phonisch realisiert (vgl. Koch & Oesterreicher, 1985; s. auch Dürscheid, 2003). Dies kann zum Anlass genommen werden, Online-Nachrichten auf dem Kontinuum Mündlichkeit-Schriftlichkeit eher bei Schriftlichkeit, TV-Nachrichten hingegen eher bei Mündlichkeit zu verorten. Wie sehr Online-Nachrichten dabei schriftlich und wie sehr TV-Nachrichten mündlich sind, ist zweitrangig; wichtig ist nicht ihre absolute, sondern ihre »relative Situierung« (Koch & Oesterreicher, 1985, S. 18). Und in diesem Sinne lassen sich Online-Nachrichten als eher schriftlich und TV-Nachrichten als eher mündlich auffassen.

Interessant ist nun, dass Hinweise dafür existieren, dass Mündlichkeit mit einem niedrigen Sprachniveau bzw. mit Markern eines niedrigen Sprachniveaus einhergeht (Koch & Oesterreicher, 1985; s. auch Dürscheid, 2003; Storrer, 2013).

Zusammenfassend ergibt sich aus den o.g. Ausführungen folgende Forschungsfrage:

Existieren Unterschiede hinsichtlich der verschiedenen Dimensionen des Sprachniveaus von Online-Nachrichten zwischen den beiden Rundfunksystemen sowie – feiner aufgelöst – zwischen verschiedenen Sendern? Anders ausgedrückt: Lassen sich die Unterschiede im Sprachniveau klassischer TV-Nachrichten zwischen den Systemen und zwischen den einzelnen Sendern auch in Online-Nachrichten finden?

Die oben angestellten Überlegungen zur Unterscheidung zwischen Mündlichkeit und Schriftlichkeit führten zudem zu folgender Hypothese:

Online-Nachrichten weisen auf den drei wesentlichen Dimensionen des Sprachniveaus (Wortschatzdiversität, Wortkomplexität und Satzkomplexität) höhere Werte sowie auf den zwei Dimensionen Umgangssprache und Verständlichkeit (Flesch-Index) niedrigere Werte auf als klassische TV-Nachrichten.

### 3 Methode

Zunächst beschreiben wir das analysierte Material (Texte der Nachrichtenbeiträge). Es folgt eine Beschreibung der eingesetzten Instrumente (u. a. LIWC) und die Operationalisierung zentraler Konstrukte (Sprachniveau und dessen einzelne Dimensionen). Schliesslich erläutern wir, wie die statistischen Analysen des Materials erfolgten.

#### 3.1 Material

Unsere Auswahl der Nachrichtensender bzw. -sendungen richtete sich nach den Hauptnachrichtensendern mit der grössten Reichweite der beiden Rundfunksysteme. Nach Zubayr und Gerhard (2018) sind dies bei den öffentlich-rechtlichen Rundfunkanbietern die *Tagesschau* und *heute* und bei den Privatsendern *RTL aktuell* und die *SAT.1 Nachrichten*. Die Auswahl von je zwei Nachrichtensendungen pro System soll eine Gegenüberstellung der Rundfunk-Nachrichtenangebote dieser beiden Systeme ermöglichen, was ein verbreitetes Vorgehen in der Nachrichtenforschung darstellt (vgl. z. B. Bruns & Marcinkowski, 1997; Krüger, 1998; Lange et al., 2019). Zudem sind vier einzelne Sender miteinander vergleichbar.

Für das Zusammenstellen des Materials der Online-Nachrichten wurde folgendermassen vorgegangen: Über einen Zeitraum von zwei Wochen wurde zunächst eine künstliche Nachrichtenwoche der o.g. Nachrichtensendungen konstruiert. So wurden in der ersten Erhebungswoche Daten der Tage Montag, Mittwoch, Freitag und Sonntag und in der zweiten Woche Daten der Tage Dienstag, Donnerstag und Samstag erhoben. Im Anschluss wurden alle Beiträge einer Nachrichtensendung

hinsichtlich der präsentierten Themen analysiert. Im nächsten Schritt wurde überprüft, ob zum jeweiligen Themenbeitrag der TV-Nachrichtensendung eines jeden Senders ebenfalls ein Online-Beitrag auf der Webseite des entsprechenden Nachrichtensenders verfügbar war. Das Ziel war es, drei Themen pro Tag ausfindig zu machen, die sowohl in den Hauptnachrichtensendungen im Fernsehen ausgestrahlt als auch am selben Tag auf der Webseite veröffentlicht wurden. Insgesamt ergab sich anhand der beschriebenen Vorgehensweise eine Gesamtstichprobe von  $N=84$  Online-Texten mit einem Umfang von insgesamt über 7000 Wörtern. Für jeden der vier Sender wurden im Zuge dieser Arbeit somit insgesamt  $n=21$  Nachrichtenbeiträge untersucht, wobei die Länge der Texte zwischen 100 und 1122 Wörtern variierte.

Da wir die Werte für die Online-Texte jenen für klassische TV-Nachrichten gegenüberstellen wollten, analysierten wir neben den Online-Texten zusätzlich die Original-Transkriptionen der TV-Nachrichten von Lange, Kouros und Schwab (2019). Hier lagen  $N=28$  Transkriptionen vor; das heisst, pro Sender wurden  $n=7$  Beiträge untersucht. Die Länge der Transkriptionen variierte zwischen 1411 und 1706 Wörtern (vgl. Lange et al., 2019). Insgesamt analysierten wir Textmaterial (Online und TV) im Gesamtumfang von 112 Beiträgen bzw. 83.281 Wörtern.

#### 3.2 Instrumente und Operationalisierung

Unsere methodische Vorgehensweise orientiert sich weitgehend an der von Lange et al. (2019), die mithilfe des computergestützten quantitativen Textanalyseprogramms Linguistic Inquiry and Word Count (LIWC; Wolf et al., 2008; s. auch Pennebaker, Francis & Booth, 2001) TV-Nachrichten hinsichtlich ihres Sprachniveaus untersucht haben (s.o.). Wir haben uns zudem an Krüger (2009) orientiert, der den sog. Flesch-Index verwendet hat.

Bei LIWC handelt es sich um ein diktionsbasiertes Computer-Programm, das die automatische quantitative Analyse von natürlicher Sprache ermöglicht,

die bereits geschrieben vorliegt oder aber mündlich vorlag und dann transkribiert wurde. Das Instrument erlaubt eine hohe Abdeckung der Alltagssprache (Wolf et al., 2008). Die Funktionalität von LIWC zielt auf die Ein-Wort-Analyse; es verwendet einen Wortzählalgorithmus, der die gezählten Wörter – die Auftretenshäufigkeit wird i. d. R. prozentual in Relation zur Textlänge dargestellt – bestimmten Kategorien zuordnet (Wolf et al., 2008).

LIWC ist bereits vielfach in der psychologischen, aber auch kommunikationswissenschaftlichen und linguistischen Forschung eingesetzt worden, z. B. zur Analyse von Literatur oder politischer Reden (Überblick bei Wolf et al., 2008). Auch in der aktuellen kommunikationswissenschaftlichen Forschung wird LIWC immer wieder eingesetzt (z. B. Gil-Lopez, Shen, Benefield, Palomares, Kosinski, & Stillwell, 2018; Ho, Hancock, & Miner, 2018). Im Besonderen ist der (potenziell) breite Einsatz bei der Analyse von Mediensprache hervorzuheben (Wolf et al., 2008).

Lange et al. (2019) legen umfassend theoretisch wie empirisch dar, dass LIWC mehrere zentrale Aspekte des Sprachniveaus computergestützt-automatisiert quantitativ valide analysieren kann. Darauf aufbauend präsentierten sie eine umfassende Analyse des Sprachniveaus deutscher TV-Nachrichten mit überwiegend klaren Ergebnissen. Dies war einer der Gründe für uns, das Prozedere dieser Studie weitgehend auf unsere Analyse des Sprachniveaus von Online-Nachrichten zu übertragen. Dies hat den weiteren Vorteil, dass die Sprachniveau-Ergebnisse für TV-Nachrichten (Lange et al., 2019) unseren Sprachniveau-Ergebnissen für Online-Nachrichten direkt gegenübergestellt werden können.

Allerdings wollten wir das Vorgehen von Lange et al. (2019) erweitern. Dies betraf erstens den Aspekt der Umgangssprache, da Lange et al. (2019) diesen Aspekt selbst diskutieren, ohne ihn allerdings empirisch zu erfassen. Auch andere Arbeiten haben den Einsatz von Umgangssprache in den Nachrichten thematisiert (z. B. Arnold, 2016; Wittwen, 1995). Zweitens wollten wir mit dem oben bereits genannten

Flesch-Index einen weiteren Marker der Leichtigkeit bzw. Schwierigkeit der Sprache mitberücksichtigen, da auch dieser in der empirischen Forschung zur Nachrichtensprache bereits zum Einsatz kam (Krüger, 2009). Im Folgenden soll nochmals kurz elaboriert werden, wie Sprachniveau definiert ist und mittels der oben bereits kurz genannten linguistischen Marker operationalisiert werden kann.

Sprachniveau (Synonym: sprachliche Gewandtheit; Lange, 2012; Lange et al., 2016, 2017, 2019) ist grob u. a. mittels Wortschatzdiversität (Anteil unterschiedlicher Wörter), Wortkomplexität (Länge der Wörter) und Satzkomplexität (Länge der Sätze) definier- und operationalisierbar. Mehrere Rezeptionsstudien (Lange, 2012; Lange, Zaretsky, Schwarz & Euler, 2014; Lange et al., 2016) bestätigen die Validität dieses Ansatzes: Unterschiedlich sprachlich gewandte Stimuli (Text, Audio, Video), die auf Basis dieser Definition erstellt wurden, wurden von Probanden mit hoher Effektstärke als unterschiedlich sprachlich gewandt bewertet. Die Probanden zeigten dabei gute Inter-Rater-Reliabilitäten, waren sich in ihren Bewertungen also weitgehend einig. Die besagten Bewertungen korrelierten zudem hoch mit den Werten für die einzelnen sprachlichen Marker der verschiedenen Dimensionen des Konstrukts. Damit kann aufgrund guter Reliabilität und gegebener Validität von der Güte dieses so operationalisierten Konstrukts ausgegangen werden, auf dessen einzelne Dimensionen im Folgenden näher eingegangen wird.

Wortschatzdiversität ist über das sog. Type-Token-Verhältnis (TTR; LIWC-Variable «Unique») bestimmbar, d. h. über die Anzahl unterschiedlicher Wörter im Verhältnis zur Gesamtwortzahl (Wimmer, 2005). TTR ist ein in der Linguistik (s. z. B. Jarvis, 2013; Wimmer, 2005), aber auch in der Psychologie (s. z. B. Kemper & Sumner, 2001) etabliertes lexikalisches Diversitätsmass (Überblick bei Lange, 2012). Pennebaker, Chung, Ireland, Gonzales und Booth (2007) konnten einen starken negativen Zusammenhang zwischen dem «TTR» und der Textlänge (Gesamtwortzahl, LIWC-Va-

riable «WC») feststellen ( $r = -.8$ ): Je länger ein Text ist, desto weniger unterschiedliche Wörter lassen sich im Text finden. Dieser Umstand sollte daher – so könnte jedenfalls argumentiert werden – entweder bei der statistischen Auswertung oder – falls dies dort nicht geschieht – bei der Interpretation der Ergebnisse berücksichtigt werden (s. u.).<sup>2</sup>

Wortkomplexität lässt sich mittels LIWC insofern erfassen, als dass das Programm den Anteil der Wörter, die länger als sechs Buchstaben sind, ausgibt (LIWC-Variable «Sixltr»). Dieses Mass ist fast identisch mit dem der mittleren Wortlänge (Lange et al., 2019; s. auch Lange, 2012).

Zur Operationalisierung der Satzkomplexität kann das Mass der Satzlänge herangezogen werden. Diese wird von LIWC in Form der mittleren Wortzahl pro Satz ausgegeben (LIWC-Variable «WPS»). Dieses Mass ist annähernd identisch mit dem sonst in der Psycholinguistik geläufigen Mass «mean length of utterance» (MLU), bei dem die Äusserungslänge nicht in Wörtern, sondern in Morphemen angegeben wird (Aitchison, 2008; Überblick bei Lange, 2012). MLU kommt vor allem in der Diagnostik des kindlichen Sprachstandes zum Einsatz, kann aber auch für die Analyse von Erwachsenensprache verwendet werden (Nippold, Ward-Lonergan, & Fanning, 2005).

In Anlehnung an Lange (2012) und Lange et al. (2019) kann damit das Sprachniveau in der vorliegenden Arbeit über Wortschatzdiversität, Wortkomplexität und Satzkomplexität bzw. über die LIWC-Variable operationalisiert werden (Wolf et al., 2008). Auf weitere potenzielle Charakteristika sprachlicher Gewandtheit, wie etwa ein hoher Anteil an hypotaktischen Satzstrukturen (s. dazu Lange et al., 2019)

oder an Passivkonstruktionen, wurde verzichtet, und zwar aus im Wesentlichen zwei Gründen: Erstens liegt empirische Evidenz für die Annahme vor, dass mit der o. g. Operationalisierung sprachliche Gewandtheit valide erfasst wird (s. z. B. Lange et al., 2016). Zweitens kann LIWC weder hypotaktische noch Passivstrukturen detektieren. Wir werden diesen Aspekt in der Diskussion wieder aufgreifen.

Allerdings versuchten wir, wie weiter oben bereits angesprochen, LIWC sinnvoll zu ergänzen, und zwar erstens durch die Zusatzkategorie der Umgangssprache und zweitens durch die Bestimmung des Flesch-Indexes. Auf diese beiden Marker wird im Folgenden detailliert eingegangen.

Lange et al. (2019) stellten die Vermutung auf, dass die PS, insbesondere Sat.1, mehr Umgangssprache verwendeten, was mit einer höheren Wortschatzdiversität einhergehen könnte. Sie lieferten zudem einige exemplarische Belege für einen höheren Gehalt an Umgangssprache in den Nachrichten von Sat.1, quantifizierten den Anteil an Umgangssprache allerdings nicht; LIWC bietet hierfür auch keine eigene Kategorie an. Daher erweiterten wir das deutsche LIWC-Diktionär mit seinen 68 Kategorien um die Kategorie Umgangssprache (US). Dies geschah einerseits induktiv, d. h. entlang des Untersuchungsmaterials, welches sowohl aus den Online-Texten, als auch aus den Original-Transkriptionen der TV-Nachrichten von Lange et al. (2019) bestand, andererseits deduktiv auf Basis von Arbeiten zur Umgangssprache (Bernhard, 2012; Wittwen, 1995).

Die neue Kategorie der Umgangssprache («US») umfasste Wörter, die nicht einer formal gehobenen Sprache entsprechen, sondern sich in salopp und/oder alltagssprachlich formulierten Wörtern (Begriffen, Bezeichnungen) und Sätzen (Aussagen) äussern (vgl. Bernhard, 2012; Wittwen, 1995), z. B. «Ausruhter», «Bauchplatscher», «cool», «Happening» und «Kerl» (Beispiele aus unserem Textkorpus der Online-Texte) sowie «Boom», «happy», «Job», «Promi» und «tja» (Beispiele aus dem Textkorpus der TV-Transkripte

2 Die starke Abhängigkeit der Variable «Unique» von der Textlänge kann zu Fehlgebrauch und -interpretationen der Ergebnisse führen; diese LIWC-Funktion ist schlicht sehr sensibel gegenüber der Wortzahl. Dies ist einer der Gründe, warum «Unique» aus der neueren Version von LIWC (vgl. z. B. LIWC 2015 Version 1.3.1. von 2016 mit Version LIWC 2001 Version 1.82 von 2001) entfernt wurde (Wolf, 2016, persönliche Mitteilung; vgl. Wolf et al., 2008).



onen; Lange et al., 2019). Diese neue Kategorie entstand konkret folgendermassen: Die Zweitautorin erstellte für das Material der Online-Nachrichten eine Liste an möglichen umgangssprachlichen Wörtern. Was als umgangssprachlich aufgefasst wurde und was nicht, folgte dabei den Kriterien, die in der entsprechenden Literatur genannt werden (Bernhard, 2012; Wittwen, 1995; s. o.). Es wurden nur solche Wörter aufgenommen, die im tatsächlichen Material vorkamen. Der Erstautor erstellte analog dazu für das Material der TV-Nachrichten eine ähnliche Liste. Beide Listen zusammen sollten umgangssprachliche Wörter des gesamten analysierten Materials (Online und TV) enthalten. Erst- und Senior-Autor bestimmten dann gemeinsam, welche Wörter Teil der finalen Kategorie sein sollten. Dabei wurden nur solche Wörter aufgenommen, über die die beiden letztgenannten Autoren Konsens erzielten. Es wurden zudem nur diejenigen Wörter für die Konstruktion der Kategorie berücksichtigt, die insgesamt mehr als einmal in den Texten vorkamen. Insgesamt konnten durch diese (qualitative) Zuordnung unter Berücksichtigung der Kategoriendefinition (vgl. Bernhard, 2012; Wittwen, 1995) 155 umgangssprachliche Wörter aus der Worthäufigkeitsliste für die Erweiterung des Diktionärs gewonnen werden.

Auf Basis des LIWC-Diktionärs von Wolf et al. (2008) erstellten wir somit ein erweitertes Diktionär. Davon wurden die folgenden fünf Kategorien verwendet: «WC» (Wortzahl), «Unique» (Type-Token-Ratio; misst Wortschatzdiversität), «Sixltr» (Anteil an Wörtern, die mehr als sechs Zeichen lang sind; misst Wortkomplexität) und «WPS» (Wörter pro Satz; misst Satzkomplexität); ausserdem die eigens erstellte Kategorie «US» (Anteil an Umgangssprache).<sup>3</sup>

Schliesslich wurde der Flesch-Index für alle Beiträge auf Basis der vorliegenden Texte bestimmt. Dieser Index wird auch «Lesbarkeitsindex» oder «Reading-Ease-Index» genannt und besteht aus den quantitativen Massen der durchschnittli-

chen Wortlänge in Silben und der durchschnittlichen Satzlänge in Wörtern (s. z. B. Ballstaedt, 2019; Immel, 2014). Er scheint damit zwei wesentliche Dimensionen des o. g. Konstrukts des Sprachniveaus ebenfalls zu erfassen, nämlich Wortkomplexität (vgl. «Sixltr») und Satzkomplexität (vgl. «WPS»).

Für die Bestimmung des Flesch-Indexes kam die für deutsches Sprachmaterial entwickelte Formel zum Einsatz:  $180 - ASL - (58,5 * ASW)$ . «ASL» steht für «Average Sentence Length», womit die durchschnittliche Satzlänge in Wörtern gemeint ist. «ASW» steht für «Average Number of Syllables per Word», d. h. für die durchschnittliche Wortlänge in Silben (s. z. B. Ballstaedt, 2019; Immel, 2014). Für die inhaltliche Interpretation der so errechneten Werte griffen wir auf die bestehende Literatur zum Flesch-Index zurück (Aziz, Fook & Alsree, 2010; Ballstaedt, 2019; Flesch, 1948; Immel, 2014; Krüger, 2009; Merten, 1995, S. 179). Die entsprechend der genannten Formel errechneten Werte drücken demnach je unterschiedlich gute Verständlichkeits- bzw. Schwierigkeitsgrade aus, nämlich (Personen, für die ein entsprechender Text verständlich ist, in Klammern): 0 bis 30: sehr schwer verständlich bzw. sehr schwierig (Akademiker / Hochschulabsolventen), 30 bis 50: schwer verständlich bzw. schwierig (Studierende), 50 bis 60: schwer bis mittelschwer verständlich bzw. ziemlich schwierig (10. bis 12. Schulklasse), 60 bis 70: mittelschwer verständlich bzw. normal (8. und 9. Schulklasse / Alter: 13–15 Jahre), 70 bis 80: mittelschwer bis leicht verständlich bzw. ziemlich leicht (7. Schulklasse), 80 bis 90: leicht verständlich (6. Schulklasse), 90 bis 100: sehr leicht verständlich (5. Schulklasse / Alter: 11 Jahre) (vgl. Aziz et al., 2010; Immel, 2014; Krüger, 2009; Merten, 1995, S. 179). Dies stellt das Standardvorgehen für den Einsatz des Flesch-Indexes dar (Aziz et al., 2010; Ballstaedt, 2019; Flesch, 1948; Immel, 2014; Krüger, 2009; Merten, 1995, S. 179). Uns war und ist allerdings die Problematik bewusst, dass an den in der besagten Literatur zum Flesch-Index genannten Kategoriengrenzen (z. B. 0–30 vs.

3 Das erweiterte Diktionär ist (im Format «\*.dic») als e-appendix verfügbar.

30–50) in Form doppelter Vergaben ein Zuordnungsproblem entstehen kann, nämlich dann, wenn ein empirisch ermittelter Flesch-Index-Wert genau auf der Grenze liegt (z. B. 30). Aus unserer Sicht wäre ein solcher Flesch-Index-Wert notgedrungen als zwischen zwei Kategorien liegend zu interpretieren (z. B. sehr schwierig bis schwierig).

Der in der Analyse von Nachrichtensprache bereits eingesetzte Flesch-Index demonstrierte eine schwere Verständlichkeit für die *Tagesschau* (Krüger, 2009), was als hohes Sprachniveau interpretierbar ist. Konkret wird für die *Tagesschau* der Wert 39 (schwer verständlich) berichtet, während die Kindernachrichtensendung *logo!* einen Wert von 65 (mittelschwer verständlich) aufweist (Krüger, 2009).

### 3.3 Statistische Analysen

Um Unterschiede hinsichtlich des Sprachniveaus zwischen Online- und TV-Berichterstattung und zwischen den beiden Rundfunksystemen statistisch zu untersuchen, wurde zunächst für verschiedene Untersuchungskriterien (konkret: «Sixltr», «WPS», «US» und Flesch-Index) und für jedes Medium (d. h. Online und TV) ein t-Test für unabhängige Stichproben berechnet. Um Aussagen über die Grösse der ermittelten Unterschiede treffen zu können, wurde mittels Mittelwerten und Standardabweichungen jeweils die Effektstärke Cohens  $d$  berechnet. Für «Unique» wurde, wegen der Wortzahlabhängigkeit dieses Masses (s. o.), eine Varianzanalyse mit Kovariate «WC» gerechnet, da t-Tests die Berücksichtigung von Kovariaten nicht ermöglichen. Da Varianzanalysen das partielle Eta-Quadrat ( $\eta^2$ ) als Effektstärke ausgeben, wurde diese in  $d$  umgerechnet. Nach Cohen (1988) sind  $\eta^2$ -Werte von .01 gering, solche von .06 moderat und solche von .14 hoch. Die übliche Klassifikation der  $d$ -Werte spricht bei einem Wert von 0.2 von einem geringen, bei einem Wert von 0.5 von einem mittleren Effekt und bei einem Wert von 0.8 oder höher von einem hohen Effekt (s. jedoch Gignac & Szodorai, 2016 für eine weniger konservative Klassifikation). Für alle  $d$ -Werte wird zusätzlich das 95-prozentige Konfidenzintervall an-

gegeben, d. h. der Wertebereich, in den der wahre  $d$ -Wert mit 95-prozentiger Wahrscheinlichkeit fällt.

Wir entschieden uns dafür, weniger über statistische Signifikanz als vielmehr über Effektstärken zu argumentieren, da erstere zu sehr von Stichprobengrössen abhängt, während zweite ein besseres Bild der tatsächlichen praktischen Relevanz zeichnen.

Für Textlänge wurden zwei (einmal für Online, einmal für TV) einfaktorielle Varianzanalysen mit Sender als Faktor und «WC» als Kriterium gerechnet. Um mögliche Unterschiede im Sprachniveau je nach Medium (Online/TV) nicht nur auf System-, sondern auch auf Senderebene feststellen zu können und um des Weiteren auch mögliche statistische Interaktionen zwischen Medium einerseits und System oder Sender andererseits prüfen zu können, wurden zusätzlich zweifaktorielle univariate Varianzanalysen durchgeführt. Die beiden Faktoren (d. h. Gruppenvariablen / unabhängige Variablen) waren (1) Medium (Online/TV) und (2) entweder Rundfunksystem oder Sender (in dieser Reihenfolge). Das Untersuchungskriterium (d. h. die abhängige Variable) war das jeweilige linguistische Merkmal (neben «WC» noch «Sixltr», «WPS», «US» und Flesch-Index). Bei den Analysen von «Unique» entschieden wir, Analysen durchzuführen, die «WC» als Kovariate berücksichtigen (s. Abschnitt 3.2 und Fussnote 2). Um die Ergebnisse aber mit der Vorgängerforschung vergleichbarer zu machen, präsentieren wir ausserdem «Unique»-Auswertungen ohne diese Kovariate. Zur Quantifizierung der Unterschiede diente als Effektstärke jeweils das besagte partielle Eta-Quadrat ( $\eta^2$ ). Um etwaige Interaktionen zwischen Medium und Sender besser zu elaborieren, wurden zusätzlich jeweils zwei (einmal für Online, einmal für TV) einfaktorielle Varianzanalysen mit Sender als Faktor und dem jeweiligen linguistischen Mass als Kriterium gerechnet.

Dieses Vorgehen führte zu einer relativ hohen Zahl an einzelnen statistischen Analysen. Dies erhöht die Gefahr, falsch-positiver Befunde (Alpha-Fehler).

In Ermangelung an bisheriger Forschung zu unserer Thematik erschien uns allerdings die Gefahr falsch-negativer Befunde (Beta-Fehler) schwerwiegender. Um keine bedeutsamen Befunde zu übersehen, wurde daher auf Alpha-Fehler-Adjustierung verzichtet (s. dazu Rothman, 1990; Perneger, 1998). Aufgrund dieser und anderer Probleme, die mit statistischer Signifikanz verbunden sind (s. o.), fokussieren wir auf das Berichten von Effektstärken; *p*-Werte zur statistischen Signifikanz (Alpha-Fehler-Niveau .05) werden nur zwecks Vollständigkeit mit berichtet (s. o.).

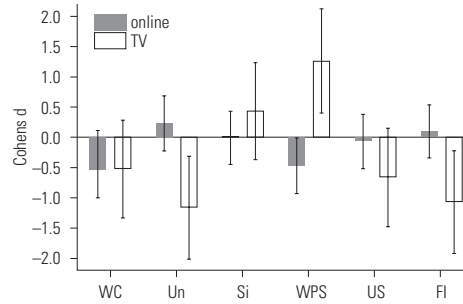
#### 4 Ergebnisse

Die Darstellung der Ergebnisse folgt der oben bereits angesprochenen Struktur. Entsprechend orientiert sie sich am jeweiligen linguistischen Merkmal (beginnend mit Textlänge und schliessend mit dem Fleisch-Index). Für jedes Merkmal wird nach Medium (Online vs. TV) und System (ÖRS vs. PS; für eine visuelle Übersicht über die wesentlichen Ergebnisse s. Abb. 1) sowie nach Sender (ARD, ZDF, SAT.1, RTL) analysiert. Statistische Masszahlen der inferenzstatistischen Auswertung, wie *p*-, *t*- und *d*-Werte, finden sich im Fliesstext; deskriptive Ergebnisse bzw. Rohwerte – konkret: (nicht adjustierte) Mittelwerte und Standardabweichungen – finden sich in Tabelle 1 als Online Supplement.

##### 4.1 Textlänge

Zunächst wurde untersucht, ob sich die Texte in ihrer Länge («WC») unterscheiden. Bei der Textlänge handelt es sich zwar nicht um ein prototypisches Mass für Sprachniveau; dennoch war diese Analyse bedeutsam, da das TTR (s. Abschnitt 4.2 und 4.3) textlängenabhängig ist. Es zeigte sich, dass Online-Nachrichtenartikel der PS signifikant und mit moderater Effektstärke länger waren als die der ÖRS,  $t(82) = -2.45$ ,  $p = .009$ ,  $d = -0.54$  (95%-KI[-0.98; -0.10]). Bei den TV-Nachrichten lagen die PS ähnlich deutlich über den ÖRS (s. Tab. 1),  $t(26) = -1.38$ ,  $p = .181$ ,  $d = -0.52$  (95%-KI[-1.31; 0.27]) (s. Abb. 1).

Abb. 1: Unterschiede (nach Effektstärke Cohens *d*) zwischen ÖRS und PS nach linguistischem Merkmal sowie nach Medium: TV (Lange et al., 2019) vs. Online



Anmerkungen: Positive *d*-Werte stehen jeweils für höhere Werte der ÖRS. Die linguistischen Merkmale sind: «WC» = Wortzahl; «Un» = Type-Token-Verhältnis (LIWC-Variante «Unique»); «Si» = Anteil der Wörter, die länger als sechs Buchstaben sind (LIWC-Variante «Sixltr»); «WPS» = Wörter pro Satz; «US» = Anteil an Umgangssprache; «FI» = Fleisch-Index. Beim Mass «Unique» der Online-Texte sowie der Transkriptionen von Lange et al. (2019) wurde statistisch für «WC» kontrolliert. Fehlerbalken zeigen das 95%-Konfidenzintervall des jeweiligen *d*-Wertes.

Auf Senderebene waren die Online-Nachrichtenbeiträge von *RTL aktuell* durchschnittlich am längsten gefolgt von der *Tagesschau*, den *SAT.1 Nachrichten* und *heute* (s. Tab. 1). Die Berechnung der einfaktorischen Varianzanalyse zeigte einen hoch signifikanten und hoch effektstarken Unterschied zwischen den vier Sendern hinsichtlich der Textlänge der Online-Berichte,  $F(3, 80) = 12.77$ ,  $p < .001$ ,  $\eta^2 = .324$ . Die sehr grossen Unterschiede lassen sich vor allem auf das Angebot von *heute* zurückführen (s. Tab. 1). Auch die TV-Transkriptionen waren auf Senderebene effektstark verschieden hinsichtlich der Wortzahl, obwohl die Signifikanz verfehlt wurde (s. Lange et al., 2019),  $F(3, 24) = 1.59$ ,  $p = .219$ ,  $\eta^2 = .165$ . Die Reihenfolge war: RTL, ZDF, Sat.1, ARD (s. Tab. 1). In der Gesamtschau lag damit ein Grund dafür vor, bei den Analysen für Wortschatzdiversität (s. Abschnitt 4.2) zusätzlich für «WC» statistisch zu kontrollieren.

Ein weiterer Grund, bei den Analysen der Wortschatzdiversität für «WC» zu kontrollieren, liegt in der hohen Korreliertheit von «WC» mit «Unique» in unserem Material: In den Transkriptionen der TV-Nachrichten (Lange et al., 2019) waren «Unique» und «WC» hoch negativ miteinander korreliert ( $r = -.59$ ,  $p = .001$ ); auch in der Stichprobe aus Online-Nachrichten war dies der Fall, sogar in stärkerem Masse ( $r = -.91$ ,  $p < .001$ ). Damit lagen mehrere Gründe dafür vor, bei der Wortschatzdiversität die Wortzahl statistisch zu kontrollieren.<sup>4</sup>

#### 4.2 Wortschatzdiversität (mit Kovariate)

Als nächstes Mass wurde das der Wortschatzdiversität, und zwar zunächst mit der Kovariate Textlänge, analysiert – erneut beginnend mit der Systemebene und gefolgt von der Senderebene.

Zunächst wurden die Werte für Wortschatzdiversität (Type-Token-Verhältnis; LIWC-Variable «Unique»; mit Kovariate «WC») für jedes Medium getrennt für ÖRS und PS miteinander verglichen. Für Online war das Ergebnis nicht signifikant und wies nur eine geringe Effektstärke auf,  $F(1, 81) = 1.09$ ,  $p = .300$ ,  $\eta^2 = .013$ ,  $d = 0.23$  (95%-KI[-0.21; 0.67]). Die ÖRS ( $M_{\text{adj}} = 59.45$ ,  $SE = 0.43$ ) lagen hier nur leicht über den PS ( $M_{\text{adj}} = 58.81$ ,  $SE = 0.43$ ). Für TV hingegen ergab sich ein signifikanter und hoch effektstarker Unterschied zwischen den beiden Systemen,  $F(1, 25) = 8.54$ ,  $p = .007$ ,  $\eta^2 = .255$ ,  $d = -1.17$  (95%-KI[-2.01; -0.33]). Die PS ( $M_{\text{adj}} = 49.68$ ,  $SE = 0.52$ ) lagen hier deutlich über den ÖRS ( $M_{\text{adj}} = 47.51$ ,  $SE = 0.52$ ) (s. Abb. 1).

Die zweifaktorielle Varianzanalyse ergab einen hoch signifikanten und hoch effektstarken Effekt für Medium,  $F(1, 107) = 138.78$ ,  $p < .001$ ,  $\eta^2 = .565$ . TV ( $M_{\text{adj}} = 69.79$ ,  $SE = 1.15$ ) lag hier deutlich

über Online ( $M_{\text{adj}} = 52.07$ ,  $SE = 0.45$ ). Der Haupteffekt für System wurde nicht signifikant,  $F(1, 107) = 1.68$ ,  $p = .197$ ,  $\eta^2 = .015$ , allerdings der Interaktionseffekt zwischen Medium und System,  $F(1, 107) = 6.01$ ,  $p = .016$ ,  $\eta^2 = .053$ . Dieser kam dadurch zustande, dass bei TV die PS über den ÖRS lagen, während sich bei Online das Muster umdrehte (s. o.; vgl. Abb. 1).

Die Varianzanalyse mit Sender statt System als Faktor ergab wieder einen hoch signifikanten und hoch effektstarken Effekt für Medium,  $F(1, 103) = 94.21$ ,  $p < .001$ ,  $\eta^2 = .478$ . Der Haupteffekt für Sender wurde nicht signifikant,  $F(3, 103) = 0.50$ ,  $p = .681$ ,  $\eta^2 = .014$ . Der Interaktionseffekt zwischen Medium und Sender wurde knapp nicht signifikant,  $F(3, 103) = 2.45$ ,  $p = .068$ ,  $\eta^2 = .014$ . Bei TV hatten die *SAT.1 Nachrichten* ( $M_{\text{adj}} = 49.83$ ,  $SE = 0.74$ ) den höchsten Wert gefolgt von *RTL aktuell* ( $M_{\text{adj}} = 49.49$ ,  $SE = 0.79$ ), der *Tagesschau* ( $M_{\text{adj}} = 47.82$ ,  $SE = 0.76$ ) und *heute* ( $M_{\text{adj}} = 47.24$ ,  $SE = 0.74$ ). Bei Online drehte sich, ähnlich wie bei Betrachtung auf Systemebene, das Muster (überwiegend) um: Den höchsten Wert hatte *heute* ( $M_{\text{adj}} = 60.04$ ,  $SE = 0.68$ ) gefolgt von *RTL aktuell* ( $M_{\text{adj}} = 58.98$ ,  $SE = 0.63$ ), der *Tagesschau* ( $M_{\text{adj}} = 58.95$ ,  $SE = 0.61$ ) und den *SAT.1 Nachrichten* ( $M_{\text{adj}} = 58.56$ ,  $SE = 0.60$ ) (s. Abb. 2, oben).

Der besagte Interaktionseffekt ist in Abbildung 2 (oben) visuell erkennbar. Er kommt, wie zwei weitere je einfaktorische Varianzanalysen zeigten, dadurch zustande, dass bei TV ein hoch effektstarker Unterschied zwischen den Sendern besteht,  $F(3, 23) = 2.79$ ,  $p = .063$ ,  $\eta^2 = .267$ , während der Unterschied zwischen den Sendern Online nur sehr gering war,  $F(3, 79) = 0.90$ ,  $p = .445$ ,  $\eta^2 = .033$ . Die Forschungsfragen lassen sich für Wortschatzdiversität somit wie folgt beantworten: TV-Nachrichten sind zwischen den Sendern verschieden, Online-Nachrichten hingegen sind einander eher ähnlich.

#### 4.3 Wortschatzdiversität (ohne Kovariate)

Unter anderem, um zu verdeutlichen, wie anfällig TTR für Unterschiede in der Wortzahl ist, führten wir sämtliche Analysen für Wortschatzdiversität erneut durch, al-

4 Lange et al. (2019) hatten in keiner ihrer Analysen für «WC» kontrolliert. Dort wurde mit Transkriptionen von stets ca. 10 Min. an TV-Nachrichten-Material gearbeitet, was ähnliche «WC»-Werte nahelegte. Die Ergebnisse unserer Re-Analysen des Materials und der Ergebnisse dieser früheren Studie fallen daher teilweise anders aus. Wir greifen diesen Aspekt in der Diskussion auf.

lerdings ohne für «WC» statistisch zu kontrollieren.

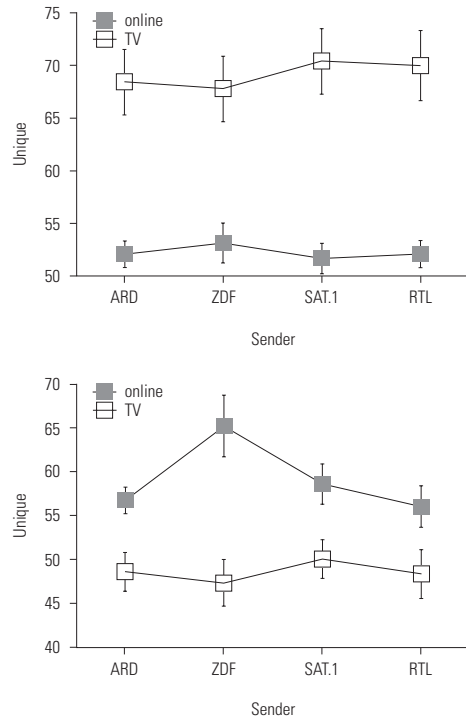
Es wurden also zunächst erneut die Werte für Wortschatzdiversität für jedes Medium getrennt für ÖRS und PS analysiert. Für Online lagen die ÖRS über den PS (s. Tab. 1),  $t(82)=2.68, p=.009, d=0.59$  (95%-KI[0.147; 1.033]). Für TV lagen die PS über den ÖRS (s. Tab. 1),  $t(26)=-1.26; d=-0.48$  (95%-KI[-1.268; 0.308]).

Die zweifaktorielle Varianzanalyse (mit Medium und System als Faktoren) ergab einen hoch signifikanten und hoch effektstarken Effekt für Medium,  $F(1, 108)=75.71, p<.001, \eta^2=.412$ . TV und Online unterschieden sich deutlich (s. Tab. 1),  $d=2.33$  (95%-KI[1.799; 2.861]). Der Haupteffekt für System wurde nicht signifikant,  $F(1, 107)<1, p=.325$ , allerdings der Interaktionseffekt zwischen Medium und System,  $F(1, 108)=4.02, p=.048, \eta^2=.036$ . Dieser kam dadurch zustande, dass bei TV die PS über den ÖRS lagen, während sich bei Online das Muster umdrehte (s. o.).

Die Varianzanalyse mit Sender statt System als Faktor ergab wieder einen hoch signifikanten und hoch effektstarken Effekt für Medium,  $F(1, 104)=97.22, p<.001, \eta^2=.483$ . Der Haupteffekt für Sender wurde ebenfalls signifikant,  $F(3, 104)=2.95, p=.036, \eta^2=.078$ . Auch der Interaktionseffekt zwischen Medium und Sender wurde signifikant,  $F(3, 104)=5.19, p=.002, \eta^2=.130$ . Bei TV waren die Werte untereinander recht ähnlich (s. Tab. 1). Bei Online fanden sich deutlichere Unterschiede: Den höchsten Wert hatte *heute* gefolgt von den *SAT.1 Nachrichten*, der *Tagesschau* und *RTL aktuell* (s. Tab. 1 und Abb. 2, unten).

Der Interaktionseffekt zwischen Sender und Medium ist in Abbildung 2 (unten) visuell gut erkennbar. Er kommt, wie zwei weitere je einfaktorielle Varianzanalysen zeigten, dadurch zustande, dass bei Online ein effektstarker Unterschied zwischen den Sendern bestand,  $F(3, 80)=12.46, p<.001, \eta^2=.318$ , während der Unterschied zwischen den Sendern bei TV nicht signifikant wurde,  $F(3, 24)=1.34, p=.286, \eta^2=.143$ . Die Forschungsfrage lässt sich für Wortschatzdiversität, wenn auf statistische Kontrolle von «WC» verzichtet wird, somit wie folgt beantworten: Online-Nachrichten

Abb. 2: Unterschiede (Rohwerte) zwischen den vier Sendern beim linguistischen Merkmal «Unique» (Type-Token-Verhältnis; TTR) als Mass für Wortschatzdiversität nach Medium: TV (Lange et al., 2019) vs. Online (oben: mit «WC» als Kontrollvariable, unten: ohne)



Anmerkungen: Fehlerbalken zeigen das 95%-Konfidenzintervall des jeweiligen Mittelwertes. Für Online und TV wurde einmal statistisch für «WC» kontrolliert (vgl. Lange et al., 2019; s. dazu Fussnote 2) und einmal nicht (vgl. Abb. oben mit Abb. unten).

sind zwischen den Sendern verschieden, TV-Nachrichten hingegen sind diesbezüglich einander eher ähnlich.

#### 4.4 Wortkomplexität

Bei Online wiesen beide Systeme fast identische Werte auf (s. Tab. 1),  $t(82)=-0.05, p=.963, d=-0.01$  (95%-KI[-0.44; 0.42]). Für TV war der Unterschied zwar ebenfalls nicht signifikant, aber fast moderat effektstark,  $t(26)=1.15, p=.260, d=0.44$

(95%-KI[-0.35; 1.23]). Die ÖRS lagen hier über den PS (s. Tab. 1 und Abb. 1).

In der zweifaktoriellen Varianzanalyse wurde nur der Haupteffekt für Medium signifikant,  $F(1, 108) = 178.43$ ,  $p < .001$ ,  $\eta^2 = .622$ : Online lag hier deutlich über TV (s. Tab. 1). Der Haupteffekt für System wurde nicht signifikant,  $F(1, 108) = 0.27$ ,  $p = .606$ ,  $\eta^2 < .002$ .

Bei Sender statt System als Faktor war der Befund sehr ähnlich: Medium wies einen sehr starken Effekt auf,  $F(1, 104) = 177.07$ ,  $p < .001$ ,  $\eta^2 = .630$ . Auch hier wurden also über 60 Prozent der Varianz im Mass für Wortkomplexität durch Medium (s. o.) erklärt.

Der Haupteffekt für Sender wurde nicht signifikant,  $F(3, 104) = .81$ ,  $p = .492$ ,  $\eta^2 = .021$ . Abbildung 3 fasst die Befunde grafisch zusammen. Hier deutet sich ein Interaktionseffekt zwischen Medium und Sender an, der durch ein positives Herausstechen der *Tagesschau* (bei TV) zustande zu kommen scheint. Der Interaktionseffekt wurde aber statistisch nicht signifikant und erklärte nur 2 Prozent Varianz,  $F(3, 104) = .73$ ,  $p = .538$ ,  $\eta^2 = .021$ . Je eine weitere einfaktorielle Varianzanalyse für beide Me-

dien konnte hier für mehr Klarheit sorgen. Für Online gab es keinen signifikanten Effekt,  $F(3, 80) = .20$ ,  $p = .894$ ,  $\eta^2 = .008$ , allerdings einen signifikanten und effektstarken Effekt für TV,  $F(3, 244) = 4.31$ ,  $p = .014$ ,  $\eta^2 = .350$ . Somit lässt sich bezüglich der Forschungsfrage Folgendes festhalten: eher Gleichheit bei Online-Nachrichten zwischen den Sendern, jedoch Unterschiede zwischen ihnen bei TV-Nachrichten.

#### 4.5 Satzkomplexität

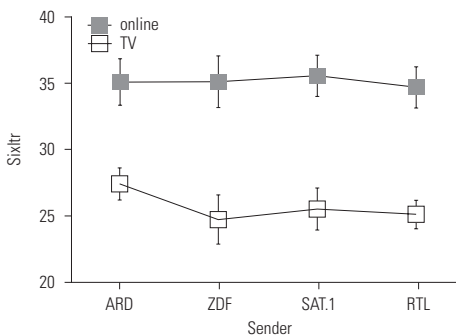
Die Online-Texte der PS enthielten mehr Wörter pro Satz als die der ÖRS (s. Tab. 1),  $t(82) = -2.22$ ,  $p = .029$ ,  $d = -0.48$  (95%-KI[-0.92; -0.04]). Bei TV-Nachrichten war das Muster umgekehrt: Die ÖRS wiesen hier substantiell höhere Werte auf als die PS (s. Tab. 1),  $t(26) = 2.31$ ,  $p = .029$ ,  $d = 1.27$  (95%-KI[0.42; 2.12]).

In der zweifaktoriellen Varianzanalyse wurde der Haupteffekt für Medium nicht signifikant ( $p = .948$ ,  $\eta^2 < .001$ ), allerdings der Interaktionseffekt zwischen Medium und System,  $F(1, 108) = 4.65$ ,  $p = .033$ ,  $\eta^2 = .041$ . Dieser kam dadurch zustande, dass Online die PS über den ÖRS lagen, während bei TV genau das umgekehrte Muster beobachtet werden konnte (s. o.).

Bei der gleichen Analyse mit Sender statt System fand sich ebenfalls kein Haupteffekt für Medium ( $p = .947$ ,  $\eta^2 < .001$ ); allerdings zeigte sich ebenfalls der besagte Interaktionseffekt, der allerdings knapp nicht mehr signifikant wurde ( $F(3, 104) = 2.49$ ,  $p = .064$ ,  $\eta^2 = .067$ ), bei visueller Inspektion von Abbildung 4 allerdings vergleichsweise klar zu erkennen und mit 7 Prozent erklärter Varianz auch moderat stark ist. Demnach sackt das ZDF bei Online ab, Sat.1 und etwas schwächer RTL dafür bei TV. Die ARD hat praktisch identische Werte bei Online und TV. Die Reihenfolge für Online war: RTL gefolgt von Sat.1, der ARD und schliesslich dem ZDF. Die Reihenfolge für TV war: ARD gefolgt von ZDF, RTL und schliesslich Sat.1 (s. Tab. 1 und Abb. 4).

In beiden Medien existierten nennenswerte Unterschiede zwischen den Sendern, d. h. sowohl bei Online,  $F(3, 80) = 4.63$ ,  $p = .005$ ,  $\eta^2 = .148$ , als auch bei TV,  $F(3, 244) = 2.12$ ,  $p = .124$ ,  $\eta^2 = .209$ . Anhand

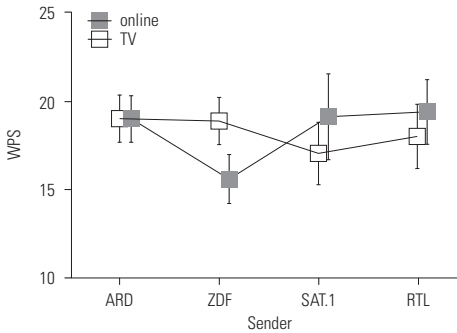
Abb. 3: Unterschiede (Rohwerte) zwischen den vier Sendern beim linguistischen Merkmal «Sixltr» (Anteil der Wörter, die länger als sechs Buchstaben sind) als Mass für Wortkomplexität nach Medium: TV (Lange et al., 2019) vs. Online



Anmerkung: Fehlerbalken zeigen das 95%-Konfidenzintervall des jeweiligen Mittelwertes.

der Effektstärken (15 % erklärte Varianz bei Online, 21 % erklärte Varianz bei TV) liesse sich mit Blick auf die Forschungsfrage von einer leicht stärkeren Senderähnlichkeit bei Online als bei TV sprechen.

Abb. 4: Unterschiede (Rohwerte) zwischen den vier Sendern beim linguistischen Merkmal «WPS» (Wörter pro Satz) als Mass für Satzkomplexität nach Medium: TV (Lange et al., 2019) vs. Online



Anmerkungen: Fehlerbalken zeigen das 95%-Konfidenzintervall des jeweiligen Mittelwertes. Die Datenpunkte sind horizontal verschoben, um Überlappungen zu verringern.

#### 4.6 Umgangssprache

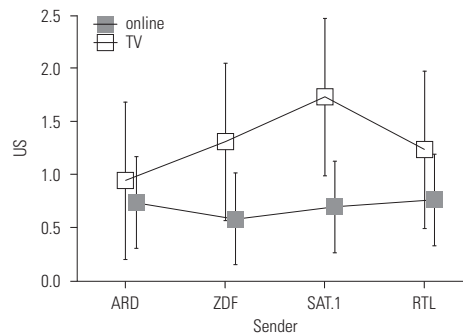
Die Online-Texte der PS enthielten nicht nennenswert mehr Umgangssprache als die der ÖRS. Der Unterschied war nicht nur nicht signifikant, sondern auch hinsichtlich der Effektstärke nahe Null,  $t(82) = -0.32$ ,  $p = .754$ ,  $d = -0.07$  (95%-KI[-0.50; 0.36]). Bei TV-Nachrichten deutete sich hingegen ein Systemunterschied an,  $t(26) = -1.78$ ,  $p = .087$ ,  $d = -0.66$  (95%-KI[-1.46; 0.14]) (s. Tab. 1 und Abb. 1).

In der zweifaktoriellen Varianzanalyse wurde der Haupteffekt für Medium signifikant,  $F(1, 108) = 8.39$ ,  $p = .005$ ,  $\eta^2 = .072$ : TV war durch mehr Umgangssprache charakterisiert als Online (s. Tab. 1). Weder der Haupteffekt für System noch der Interaktionseffekt zwischen Medium und System wurden signifikant ( $p > .3$ ,  $\eta^2 < .01$ ).

Bei Sender statt System als Faktor war das Ergebnis ähnlich: Medium wurde signifikant,  $F(3, 104) = 8.21$ ,  $p = .005$ ,  $\eta^2 = .073$ , jedoch nicht der Haupteffekt für Sender

( $p > .6$ ,  $\eta^2 < .02$ ). Abbildung 5 legt nahe, dass die Sender in ihren Online-Nachrichten eher ähnlich wenig Umgangssprache verwenden, während bei TV Sat.1 nach oben heraussticht, ZDF und RTL in etwa ähnlich hohe Werte haben und die ARD relativ wenig Umgangssprache verwendet (s. Tab. 1). Dieser sich visuell in der Abbildung andeutende Interaktionseffekt wurde allerdings nicht signifikant und war mit 2 % erklärter Varianz gering,  $F(3, 104) = 0.69$ ,  $p = .563$ ,  $\eta^2 = .019$ .

Abb. 5: Unterschiede (Rohwerte) zwischen den vier Sendern beim linguistischen Merkmal «US» (Umgangssprache) nach Medium: TV (Lange et al., 2019) vs. Online



Anmerkungen: Fehlerbalken zeigen das 95%-Konfidenzintervall des jeweiligen Mittelwertes. Die Datenpunkte sind horizontal verschoben, um Überlappungen zu verringern.

Mit erneut zwei separaten je einfaktoriellen Varianzanalysen liess sich der Forschungsfrage näher auf den Grund gehen: So waren die Senderunterschiede bei TV effektstark verschieden,  $F(3, 24) = 3.01$ ,  $p = .050$ ,  $\eta^2 = .273$ , bei Online hingegen nahe Null,  $F(3, 80) = 0.12$ ,  $p = .950$ ,  $\eta^2 = .004$ . Das bedeutet erneut: deutliche Unterschiedlichkeit bei TV, hingegen eher Gleichheit bei Online.

Lange et al. (2019) diskutierten ihren Befund, dass die PS in ihren TV-Nachrichten eine höhere Wortschatzdiversität aufwiesen, indem sie einen Zusammenhang zwischen Wortschatzdiversität und Um-

gangssprache vermuteten. Auf Basis der nun vorliegenden Daten lässt sich diese Vermutung empirisch überprüfen: Für TV-Nachrichten war die Korrelation zwischen den beiden Massen nahe Null und deutlich nicht signifikant ( $r=.07$ ,  $p=.74$ ); für Online-Nachrichten war sie vergleichsweise klein und knapp nicht signifikant ( $r=-.20$ ,  $p=.06$ ), allerdings negativ, was der besagten Vermutung von Lange et al. (2019) widerspricht.

Abschliessend wurde für die neue Kategorie der Umgangssprache ermittelt, wie hoch sie in unserer Gesamtstichprobe mit den anderen Markern korreliert. Diese Analyse wurde durchgeführt, um erstens einen besseren Eindruck davon zu ermitteln, was sich hinter Umgangssprache verbirgt, d. h. für welche anderen sprachlichen Muster sie indikativ ist, und zweitens, da diese Kategorie in der Vorgängerpublikation (Lange et al., 2019) fehlte. Demnach korrelierte «Umgangssprache» signifikant mit ( $r$ -/ $p$ -Werte in Klammern): «WC» (.308 / .001), «Unique» (-.307 / .001), «Sixltr» (-.358 / <.001) und «Flesch» (.312 / .001). Die Korrelation mit «WPS» war positiv, aber nicht signifikant (.126 / .186).

#### 4.7 Flesch-Index

Die Online-Texte beider Systeme wiesen ähnliche Flesch-Werte auf,  $t(82)=0.46$ ,  $p=.643$ ,  $d=0.07$  (95%-KI[-0.36; 0.50]). Für TV-Nachrichten hingegen fand sich in Form eines höheren Flesch-Index eine deutlich leichtere Verständlichkeit der PS,  $t(26)=-2.75$ ,  $p=.011$ ,  $d=-1.07$  (95%-KI[-1.90; -0.24]) (s. Tab. 1 und Abb. 1).

Die zweifaktorielle Varianzanalyse mit System als einem der Faktoren bestätigte dies, da der Haupteffekt für Medium signifikant wurde,  $F(1, 108)=44.18$ ,  $p<.001$ ,  $\eta^2=.290$ : TV wies einen höheren Flesch-Index auf als Online (s. Tab. 1), war also leichter verständlich. Der Haupteffekt für System wurde nicht signifikant,  $F(1, 108)=1.12$ ,  $p=.293$ ,  $\eta^2=.010$ . Auch der Interaktionseffekt zwischen Medium und System verfehlte in der Varianzanalyse die Signifikanz,  $F(1, 108)=2.47$ ,  $p=.119$ ,  $\eta^2=.022$ .

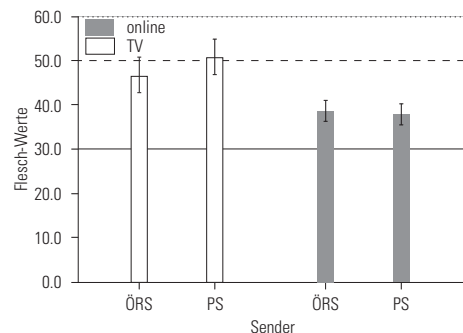
In der zweifaktoriellen Varianzanalyse mit Sender (statt System) war das Ergebnis ähnlich. So wurde wieder der Haupteffekt

für Medium signifikant,  $F(1, 104)=43.69$ ,  $p<.001$ ,  $\eta^2=.296$ . Der Haupteffekt für Sender wurde nicht signifikant,  $F(3, 104)=1.21$ ,  $p=.310$ ,  $\eta^2=.034$ . Auch der Interaktionseffekt zwischen Medium und Sender verfehlte in der Varianzanalyse die Signifikanz,  $F(3, 104)=1.25$ ,  $p=.295$ ,  $\eta^2=.035$ . In Tabelle 1 finden sich alle Flesch-Werte nach Sender und Medium (Online vs. TV).

Abschliessend wurden erneut zwei separate je einfaktorielle Varianzanalysen gerechnet: Demnach waren die Senderunterschiede bei TV hoch effektstark verschieden,  $F(3, 24)=7.01$ ,  $p=.002$ ,  $\eta^2=.467$ , bei Online hingegen nahe Null,  $F(3, 80)=0.17$ ,  $p=.917$ ,  $\eta^2=.006$  (s. Tab. 1). Demnach ergibt sich auch für die Analyse des Flesch-Indexes eine deutliche Unterschiedlichkeit bei TV, hingegen eher Gleichheit bei Online.

Konkret waren unsere Flesch-Indexwerte nach System für TV/Online diese: 46.71 / 38.71 (ÖRS) und 50.86 / 37.90 (PS) und nach Sender für TV/Online diese: 43.71 / 38.29 (ARD), 49.71 / 39.14 (ZDF), 50.43 / 37.38 (SAT.1) und 51.29 / 38.43 (RTL). Abbildungen 6 (Systemvergleich) und 7 (Sendervergleich) geben einen grafischen Überblick.

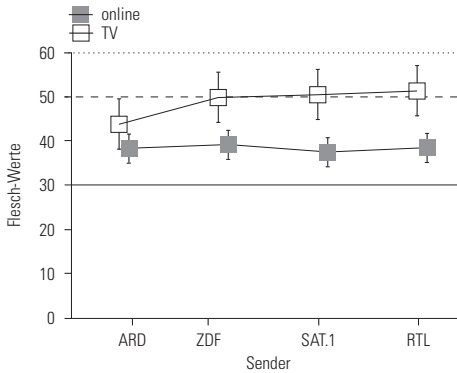
Abb. 6: Unterschiede (Rohwerte) zwischen den Systemen beim Flesch-Index nach Medium: TV (Lange et al., 2019) vs. Online



Anmerkungen: Fehlerbalken zeigen das 95%-Konfidenzintervall des jeweiligen Mittelwertes. Unterschiedliche horizontale Linien markieren unterschiedliche Verständlichkeitsgrade: 0–30 = sehr schwer, 30–50 = schwer, 50–60 = schwer bis mittelschwer. ÖRS = Öffentlich-rechtliche Sender; PS = Privatsender.



Abb. 7: Unterschiede (Rohwerte) zwischen den vier Sendern beim Flesch-Index nach Medium: TV (Lange et al., 2019) vs. Online



Anmerkungen: Fehlerbalken zeigen das 95%-Konfidenzintervall des jeweiligen Mittelwertes. Unterschiedliche horizontale Linien markieren unterschiedliche Verständlichkeitsgrade: 0–30=sehr schwer, 30–50=schwer, 50–60=schwer bis mittelschwer.

Hinsichtlich der Forschungsfrage kann auch beim Flesch-Index von einer leicht stärkeren Senderähnlichkeit bei Online als bei TV gesprochen werden (vgl. Abb. 7).

Abschliessend wurde auch für den Flesch-Index ermittelt, wie hoch er in unserer Gesamtstichprobe mit den anderen Markern korreliert (vgl. Abschnitt 4.6). Demnach korrelierte der Flesch-Index signifikant mit ( $r$ - /  $p$ -Werte in Klammern): «WC» (.501 / <.001), «Unique» (-.371 / <.001), «Sixltr» (-.776 / <.001), «WPS» (-.264 / .005) und Umgangssprache (.312 / .001).

## 5 Diskussion und Fazit

Zunächst fassen wir die wesentlichen Befunde zusammen und interpretieren sie vor dem Hintergrund unserer Forschungsfrage und unserer Hypothese. Wir schließen mit Limitationen unserer Studie und einem Ausblick auf potenziell zukünftige Forschung.

### 5.1 Zusammenfassung und Interpretation der Befunde

Ziel der vorliegenden Arbeit war die Analyse des Sprachniveaus der Online-Nachrichten nach Rundfunksystem bzw. Sender. Weiterhin sollten diese Ergebnisse jenen für klassische TV-Nachrichten gegenübergestellt werden. Die hierbei zu beantwortende Forschungsfrage war, ob Unterschiede im Sprachniveau von Online-Nachrichten zwischen den beiden Systemen sowie auf Senderebene existieren und ob sich hierbei ähnliche Muster wie in TV-Nachrichten (Lange et al., 2019) zeigen. Es fand sich, dass Online eher Gleichheit über Systeme und Sender hinweg herrschte, bei TV-Nachrichten hingegen eher Unterschiedlichkeit.

Basierend auf der etablierten Unterscheidung zwischen Mündlichkeit und Schriftlichkeit (Koch & Oesterreicher, 1985) wurde weiterhin die Hypothese aufgestellt, dass Online-Nachrichten (da stärker der Schriftlichkeit zuzuordnen) ein höheres Sprachniveau aufweisen sollten als klassische TV-Nachrichten. Für drei («Sixltr», «US» und Flesch-Index) der fünf Masse konnte die entsprechende Hypothese bestätigt werden: Online-Nachrichten sind wortkomplexer, enthalten weniger Umgangssprache und weisen einen niedrigeren Flesch-Index auf als TV-Nachrichten. Dies könnte tatsächlich daran liegen, dass erstere eher der Schriftlichkeit (Koch & Oesterreicher, 1985) entsprechen als letztere. Für diese drei Masse existieren zudem bei TV-Nachrichten teils sehr deutliche Unterschiede zwischen ÖRS und PS und zwischen den einzelnen Sendern, während Online-Nachrichten bei diesen Massen für beide Systeme und die einzelnen Sender recht ähnlich waren.

Hinsichtlich der Gründe für die relative Gleichheit bei Online- und die relative Unterschiedlichkeit bei TV-Nachrichten (deren Validität vorausgesetzt, s.u.) kann Folgendes vermutet werden: Die TV-Nachrichten der ARD haben ihren Ursprung in den frühen 1950er Jahren; die des ZDF folgten Mitte der 1960er Jahre, die von Sat.1 und RTL schliesslich Mitte bzw. Ende der 1980er Jahre. Damit entstanden die unterschiedlichen Sendungen zu deut-

lich unterschiedlichen Zeiten, und es ist naheliegend, dass diesen Unterschieden auch unterschiedliche Philosophien entsprechen (vgl. Bartel, 1997; Matzen, 2009; Wahdat, 2009), die sich dann auch sprachlich niederschlagen, und zwar – sozusagen der jeweiligen Tradition folgend – bis in die Gegenwart hinein. Die jeweiligen Online-Nachrichten hingegen entstanden allesamt etwa in den 1990er Jahren, d. h. unter ähnlichen Bedingungen und – so darf vermutet werden – unter etwa gleicher Berücksichtigung der «Regeln» der (für alle) neuen Online-Welt. Für Online-Nachrichten könnte also von Konvergenz, für TV-Nachrichten von Konkurrenz der Systeme und Sender die Rede sein (vgl. Krüger, 1998; Weiss, 2003).

Einige der Befunde weisen interessanterweise auf die Sonderstellung von *heute* hin: So kann *heute* bei TV weder bei Wort- oder Satzkomplexität noch bei Umgangssprache mit dem hohen Niveau der *Tagesschau* mithalten. Bei Online fällt *heute* insbesondere bei Satzkomplexität wiederum gegenüber der *Tagesschau* stark ab. Die Sonderstellung von *heute* lässt sich gut mit der Zielsetzung dieser Sendung erklären, einen Kompromiss zwischen Seriosität einerseits und Unterhaltungsorientierung andererseits darzustellen (Bartel, 1997; vgl. dazu Lange et al., 2019). Hinsichtlich der nach Rundfunksystem unterschiedlichen Unterhaltungsorientierung, wie einleitend dargelegt, lässt sich schlussfolgern: Für TV-Nachrichten existieren klare Sprachniveau-Korrelate dieser Orientierung, die grob betrachtet tatsächlich einen Unterschied zwischen ÖRS und PS aufzeigen, feiner betrachtet aber vor allem die Ausnahmestellung der *Tagesschau* gegenüber anderen Sendungen dokumentieren. Für Online-Nachrichten sind die entsprechenden Befunde, vor allem durch die relative Gleichheit der Systeme und Sender hinsichtlich des Sprachniveaus, weniger klar.

Gegenüber Lange et al. (2019) berücksichtigte unsere Studie auch Umgangssprache, die ebenfalls als Korrelat der Unterhaltungsorientierung angesehen werden kann (z. B. Arnold, 2016; Wittwen, 1995). Im Einklang mit den sonstigen Befunden wiesen die TV-Nachrichten der

PS (aber auch des ZDF) verglichen mit den ÖRS (bzw. der ARD) vergleichsweise hohe Werte bei diesem Mass auf. Für Online-Nachrichten konnten, erneut wegen starker Ähnlichkeit, auch hier keine klaren Unterschiede gefunden werden. Da Online-Nachrichten weniger Umgangssprache enthielten als TV-Nachrichten, könnten letztere als unterhaltungsorientierter eingestuft werden. Inwiefern diese stärkere Unterhaltungsorientierung von TV-Nachrichten tatsächlich der Fall ist und sich auch unabhängig vom Phänomen des Sprachniveaus zeigt, sollte weiter erforscht werden.

Wiederum über die Untersuchung von Lange et al. (2019) hinausgehend ermittelten wir für das gesamte Material den Flesch-Index. Dabei fand sich vor allem, dass Online-Nachrichten als schwerer verständlich einzustufen sind als TV-Nachrichten, was mit dem Rezeptionsmodus erklärt werden kann (Lesen im erstgenannten, Zuhören im letztgenannten Fall). Erstaunlich war, dass selbst die höchsten von uns gefundenen Flesch-Werte (also die mit der höchsten Verständlichkeit) immer noch im Bereich schwerer bis mittelschwerer Verständlichkeit lagen. Dies steht im Widerspruch zu dem Ziel, möglichst breite Teile der Bevölkerung mit Nachrichtenangeboten zu erreichen.

Bezüglich des Flesch-Indexes ist noch zu betonen, dass, wie in Abschnitt 3.2 erläutert, ein potenzielles Problem an den Kategoriengrenzen besteht, nämlich für den Fall, dass ein empirischer Flesch-Index-Wert genau auf der Grenze zwischen zwei Kategorien liegt. Da die Literatur zum Flesch-Index (z. B. Aziz et al., 2010) in diesem Punkt leider eine Schwachstelle aufweist, wir aber keine andere Möglichkeit hatten, als uns auf die Literatur zu berufen, die es zu dem Thema gibt, und diese die Kategorien eben genauso definiert, mussten wir notgedrungen diese potenzielle Schwachstelle mit in Kauf nehmen. In unserem Fall entstand jedoch konkret kein solches Problem, da keiner unserer Flesch-Index-Werte (s. Abschnitt 4.7) genau auf einer der Grenzen lag. Damit sind sämtliche von uns empirisch ermittelten

Flesch-Werte eindeutig einer Kategorie zuzuordnen.

Ein Befund, der weder Teil der Forschungsfrage noch der Hypothese war und sich somit zusätzlich ergeben hat, soll abschliessend kurz Erwähnung finden, da er interessant erscheint: Die Sprachniveau-Werte der verschiedenen Online-Nachrichten scheinen heterogener zu sein als die entsprechenden Werte für TV-Nachrichten. So finden sich in Tabelle 1 (siehe Online Supplement) deskriptiv-nummerisch durchweg höhere Standardabweichungen für Online- verglichen mit TV-Nachrichten. Als potenzielle Erklärung kommen aus unserer Sicht Unterschiede in der Rezeption der verschiedenen Nachrichten in Frage: Nachrichten im (linearen) TV erfordern eine eher passive Rezeption; der Rezipient kann praktisch keinen aktiven Einfluss nehmen. Bei Online-Nachrichten hingegen ist stärkere Selektion seitens des Rezipienten möglich: Was dem einen nicht gefällt oder aber zu schwer verständlich ist, kann dieser wegklicken und stattdessen ein anderes Angebot aufrufen; dem anderen ist ein hohes Sprachniveau möglicherweise gerade recht.

## 5.2 Limitationen der Studie

Eine Einschränkung dieser Studie ist, dass wir die Themen der Online- und der TV-Nachrichten nicht aneinander angehängen haben; es wurde zu unterschiedlichen Zeitpunkten über jeweils künstliche Wochen Nachrichtenmaterial gesammelt. Dieses bestand aus 21 Beiträgen aus TV-Nachrichten (Lange et al., 2019) und aus – zu einem späteren Zeitpunkt gesammelten – 84 Beiträgen aus Online-Nachrichten. Denkbar wäre, dass thematische Unterschiede zwischen den Beiträgen der beiden Medienbereiche mit Unterschieden im Sprachniveau konfundiert waren. Allerdings liesse sich dem entgegenhalten, dass erstens die Hauptnachrichten ganz grundsätzlich betrachtet immer wieder ähnliche Themen behandeln (z. B. wichtige Neuigkeiten aus den Bereichen Politik und Wirtschaft) und dass zweitens einzelne thematisch bedingte sprachliche Abweichungen in irgendeiner

Richtung (sofern sie existieren) durch die vergleichsweise hohe Zahl an Beiträgen nicht übermässig ins Gewicht fallen. Dabei kann allerdings kritisch gesehen werden, dass insgesamt nur 105 Texte in die Analyse eingingen, wobei die Zahl der Texte, die das TV-Material (Lange et al., 2019) betrafen, besonders niedrig war; entsprechend gross waren die Konfidenzintervalle. Nachfolgende Forschung sollte diesen Kritikpunkt von vornherein adressieren, indem erstens eine grössere Stichprobe an Untersuchungsmaterial gewählt und zweitens ein Matching der Inhalte durchgeführt wird.

Ebenfalls in Zukunft untersucht werden sollte die Interaktion zwischen dem Sprachniveau und der visuellen Gestaltung von TV- und Online-Nachrichten. Wie in der Einleitung erwähnt, findet die Übermittlung von Nachrichten auf sprachlicher und visueller Ebene statt; Nachrichten stellen multimodale Medienprodukte dar (Brosius, 1998; Holly, 2004; Lobinger, 2012). Ob bei klassischen audiovisuellen Nachrichten im TV, bei denen u. a. (Bewegt-)Bild und Ton (inklusive gesprochener Sprache) zusammenwirken, oder bei Online-Nachrichten, die neben geschriebenem Text oftmals aus zusätzlichen Grafiken und sonstigen Bildern bestehen – Nachrichten werden meist über mehr als einen Kanal kommuniziert, wobei die Kanäle komplementär zusammenspielen (Brosius, 1998; Burger, 2005; Holly, 2004; Lobinger, 2012). Die vorliegende Arbeit hat sich auf sprachliche Aspekte von Nachrichten beschränkt – ähnlich, wie frühere Forschung u. a. deren visuelle Gestaltung fokussiert hat (z. B. Winterhoff-Spurk, Unz & Schwab, 2005). Idealerweise werden alle Kanäle zusammen adressiert, was jedoch in einem Aufsatz wie dem vorliegenden nicht möglich ist. Zudem funktioniert der sprachliche Kanal von Nachrichten bereits für sich, bedarf also nicht zwingend immer (bewegt-)bildlicher Aspekte: Das Radio-Programm von *hr-info* etwa sendet jeden Abend einen Audio-Stream der 20-Uhr-Ausgabe der *Tageschau*. Und *Die Welt*, die Nachrichten klassisch als Print-Produkt, aber auch Online verbreitet, bietet Rezipienten die Möglich-

keit, sich die Nachrichten per App vorlesen zu lassen. Zudem dienen Text und Bild teils unterschiedlichen Zwecken (Brosius, 1998; Lobinger, 2012). Gleichwohl könnte die vorliegende Forschung in Zukunft um Aspekte der visuellen Gestaltung erweitert werden.

Die Operationalisierung sprachlicher Gewandtheit fand unter Rückgriff auf aktuelle sprachpsychologische Studien statt, die ihrerseits auf etablierte Masse aus der Psycholinguistik und besonders der Sprachdiagnostik fussten. Dennoch könnte kritisiert werden, dass Sprachniveau auch anders hätte operationalisiert werden können (z. B. u. a. über den Anteil von Passivsätzen). Dem lässt sich allerdings entgegenhalten, dass neben der Berücksichtigung der entsprechenden Fachliteratur bei der Operationalisierung des Konstrukts auch Rezeptionsstudien durchgeführt wurden, die die Validität dieser Operationalisierung bestätigten (Überblick bei Lange, 2012). Gleichwohl beinhaltet diese nicht alle denkbaren Facetten sprachlicher Gewandtheit, sondern vor allem solche, die computergestützt durch LIWC erfasst werden können (s. dazu die nachfolgenden Ausführungen).

Als allgemeine Kritik an der vorliegenden Studie und ihrer Verwendung von LIWC lässt sich u. a. nennen, dass wir «blind» auf die Kategorien von LIWC vertraut haben, weil vertrauen mussten. LIWC verwendet jedoch lediglich einen Ein-Wort-Zähl-Algorithmus und kann daher komplexe Sinnzusammenhänge nicht erfassen. Mit unserer selbst erstellten Kategorie der Umgangssprache haben wir versucht, die eigentlichen Möglichkeiten von LIWC zu erweitern; doch auch hier trifft prinzipiell die gleiche Kritik zu. Zudem mag unser Vorgehen bei der Aufnahme eines Wortes in die Kategorienbildung kritikwürdig erscheinen. Was umgangssprachlich ist (und was nicht), ist weniger trennscharf bestimmbar, als dies bei anderen sprachlichen Dimensionen (z. B. Wortkomplexität) der Fall war. Wir haben uns zwar an verfügbarer Literatur zur Thematik (Bernhard, 2012; Wittwen, 1995) orientiert; letztlich basieren die Zuordnungen von Wörtern zur entsprechenden Kategorie

«US» allerdings auf Plausibilität, waren also lediglich augenscheinvalid. Andere Autoren wären vermutlich mitunter zu einer leicht anderen Kategorienbildung gekommen. Dennoch sind die Ergebnisse in sich schlüssig und bestätigen die entsprechenden Annahmen. Gleichwohl ist bei der Interpretation insbesondere unserer Ergebnisse zu Unterschieden in der Umgangssprache Vorsicht geboten. Es handelt sich bei der Kategorie «Umgangssprache» um keine etablierte LIWC-Kategorie, sondern um eine, die wir selbst für unsere Textkorpora entwickelt haben und die auch nur für unsere Korpora an Texten passte und daher nicht generalisierbar ist.

Konkreter lässt sich an der Verwendung von LIWC weiterhin folgendes kritisieren: Bei der LIWC-Variable «Unique», als Mass für TTR (Wortschatzdiversität), ist deren Textlängenabhängigkeit als Problem zu nennen, das in der vorliegenden Arbeit allerdings durch statistisches Herauspartialisieren von «WC» kontrolliert wurde. Was passiert, wenn auf eine solche statistische Kontrolle verzichtet wird, wollten wir explizit zeigen, indem wir zusätzlich ohne statistische Kontrolle ausgewertet haben. Wie aus Abbildung 2 (vgl. oben mit unten) hervorgeht, sind die Ergebnisse praktisch gegenläufig – je nachdem, ob für «WC» statistisch kontrolliert wurde oder nicht. Das verdeutlicht nochmals die Problematik dieses Masses und unterstreicht die Notwendigkeit, hier besonders sorgfältig vorzugehen (vgl. Fussnote 2).

Bei der LIWC-Variable «WPS», als Mass für Satzkomplexität, wäre als Problem denkbar, dass Satzkomplexität neben der blossen Wortzahl auch von der Satzstruktur beeinflusst wird: Zwei Sätze gleichen Inhalts können aus der gleichen Anzahl an Wörtern bestehen, aber während der eine parataktisch schlicht aneinanderreihet, besteht der andere aus hypotaktischen Verschachtelungen. Etwas Ähnliches würde für zwei Sätze gleichen Inhalts mit etwa gleicher Anzahl an Wörtern gelten, wenn der eine im Aktiv, der andere aber im Passiv steht. LIWC, das in der linguistischen und kommunikationswissenschaftlichen Forschung vielfach eingesetzt wurde und wird (s. Einleitung), kann derlei Unterschiede

allerdings nicht erfassen. Gleichwohl kann «WPS», da es hoch mit dem etablierten Satzkomplexitätsmass «MLU» korreliert, als valides Mass für Satzkomplexität gelten (Lange et al., 2019). Dennoch könnten zusätzliche Aspekte sprachlicher Gewandtheit, etwa die Satzstruktur betreffend oder hinsichtlich des Sprachstils (vgl. Pennebaker & King, 1999), weitere Varianz z. B. in der Wahrnehmung von sprachlicher Gewandtheit durch Rezipienten erklären und in zukünftiger Forschung daher mitberücksichtigt werden.

Einen ersten Schritt weg davon, sich zu sehr auf LIWC zu verlassen, das neben den bereits genannten Problemen noch jenes hat, nicht 100 Prozent der Sprache abzudecken (vgl. Wolf et al., 2008), haben wir in der vorliegenden Arbeit bereits getan, indem wir nicht nur zusätzlich Umgangssprache untersuchten, sondern für das gesamte Material den etablierten Flesch-Index berechneten. Die Korrelation der Flesch-Werte mit den Werten der anderen Sprachniveau-Marker bestätigt dessen Brauchbarkeit.

### 5.3 Fazit und Ausblick

Hinsichtlich der sprachlichen Gestaltung von Online-Nachrichtenberichten der Rundfunkanstalten scheint eine Unterscheidung der Anbieter hinsichtlich des Sprachniveaus des Online-Angebots weniger eindeutig, als dies bei klassischen Fernsehnachrichten der Fall ist (Lange et al., 2019). Zwar zeigen sich einige Unterschiede im Sprachniveau der Online-Berichterstattung der ÖRS verglichen mit jener der PS, jedoch scheinen sich Nachrichten in der Online-Welt im System- und Sendervergleich eher zu ähneln (Konvergenz) als sich zu unterscheiden (Konkurrenz). Teils ergab sich ein höheres Sprachniveau der Online-Nachrichten im Vergleich zu ihren TV-Pendants. Dies war bei Wortkomplexität und bei Umgangssprache der Fall, was daran liegen könnte, dass Online-Nachrichten eher der Schriftlichkeit, TV-Nachrichten eher der Mündlichkeit (Koch & Oesterreicher, 1985) zuzuordnen sind. Aufgrund einiger Limitationen unserer Forschung (z. B. nur 84 Online-Texte bei nur 21 TV-Texten, deren Themen zum Teil

nicht übereinstimmen) sollte zukünftige Forschung evaluieren, inwiefern die hier präsentierten Befunde tatsächlich robust sind.

Als Konsequenz aus unseren Ergebnissen könnte in zukünftiger Forschung der Frage nachgegangen werden, wie Unterschiede im Sprachniveau klassischer TV-Nachrichten auf Rezipienten wirken, etwa bezüglich der Informationsvermittlung oder darauf, als wie angenehm die Nachrichtenrezeption empfunden wird (Lange et al., 2019). Unsere Befunde zu Online-Nachrichten legen also nahe, in einem nächsten Schritt die Wirkung von TV- aber auch von Online-Nachrichten in experimentellen Rezeptionsstudien in den Blick zu nehmen. Damit könnte u. a. die Frage nach der optimalen sprachlichen Gestaltung crossmedialer Informationsvermittlung beantwortet werden.

### Danksagung

Wir bedanken uns sehr herzlich für die sehr guten und hilfreichen Kommentare des Editors und der zwei anonymen Reviewer. An dieser Stelle sei auch Stefanie Kouros gedankt für ihre Arbeit im Kontext der Original-Transkriptionen der TV-Nachrichten aus Lange, Kouros und Schwab (2019).

### Supplementary Material

Zu diesem Beitrag ist ein von den AutorInnen erstelltes Online Supplement verfügbar.

### Literatur

- Aitchison, J. (2008). *The articulate mammal: An introduction to psycholinguistics* (5. Auflage). London: Routledge.
- Arnold, B. P. (2016). *Nachrichten: Schlüssel zu aller Information*. Baden-Baden: Nomos.
- Aziz, A., Fook, C. Y., & Alsree, Z. (2010). Computational text analysis: A more comprehensive approach to determine readability

- of reading materials. *Advances in Language and Literary Studies*, 1(2), 200–219.
- Ballstaedt, S.-P. (2019). *Sprachliche Kommunikation: Verstehen und Verständlichkeit*. Tübingen: Narr Francke Attempto.
- Bartel, R. (1997). *Fernsehnachrichten im Wettbewerb. Die Strategien der öffentlichen und privaten Anbieter*. Köln: Böhlau.
- Bernhard, U. (2012). *Infotainment in der Zeitung: Der Einfluss unterhaltungsorientierter Gestaltungsmittel auf die Wahrnehmung und Verarbeitung politischer Informationen*. Baden-Baden: Nomos.
- Brosius, H.-B. (1998). Visualisierung von Fernsehnachrichten. In K. Kamps & M. Meckel (Hrsg.), *Fernsehnachrichten. Prozesse, Strukturen, Funktionen* (S. 213–224). Opladen: Westdeutscher Verlag.
- Bruns, T. & Marcinkowski, F. (1997). *Politische Information im Fernsehen: Eine Längsschnittstudie zur Veränderung der Politikvermittlung in Nachrichten- und politischen Informationssendungen*. Opladen/Wiesbaden: Leske & Budrich.
- Burger, H. & Luginbühl, M. (2014). *Mediensprache. Eine Einführung in Sprache und Kommunikationsformen der Massenmedien* (4. Auflage). Berlin: de Gruyter.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. Auflage). Hillsdale: Erlbaum.
- Donsbach, W. & Büttner, K. (2005). Boulevardisierungstrend in deutschen Fernsehnachrichten: Darstellungsmerkmale der Politikberichterstattung vor den Bundestagswahlen 1983, 1990 und 1998. *Publizistik*, 50(1), 21–38.
- Dürscheid, C. (2003). Medienkommunikation im Kontinuum von Mündlichkeit und Schriftlichkeit. Theoretische und empirische Probleme. *Zeitschrift für Angewandte Linguistik*, 38, 37–56.
- Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32(3), 221–233.
- Gignac, G. E. & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, 102, 74–78.
- Gil-Lopez, T., Shen, C., Benefield, G. A., Palomares, N. A., Kosinski, M. & Stillwell, D. (2018). One size fits all: Context collapse, self-presentation strategies and language styles on Facebook. *Journal of Computer-Mediated Communication*, 23(3), 127–145.
- Half, G. (1998). Wa(h)re Bilder? Zur Glaubwürdigkeit von Fernsehnachrichten. In K. Kamps & M. Meckel (Hrsg.), *Fernsehen: Prozesse, Strukturen, Funktionen* (S. 127–134). Opladen/Wiesbaden: Westdeutscher Verlag.
- Ho, A., Hancock, J. & Miner, A. S. (2018). Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot. *Journal of Communication*, 68(4), 712–733.
- Holly, W. (2004). *Fernsehen*. Tübingen: Niemeyer.
- Hölig, S. & Hasebrink, U. (2018). *Reuters Institute Digital News Report 2018: Ergebnisse für Deutschland*. Hamburg: Verlag Hans-Bredow-Institut.
- Immel, K.-A. (2014). *Regionalmeldungen im Hörfunk. Verständlich schreiben für Radiohörer*. Wiesbaden: Springer VS.
- Jarvis, S. (2013). Capturing the diversity in lexical diversity. *Language Learning*, 63, 87–106.
- Kamps, K. (1998). «Zur Politik, nach Bonn ...»: Politische Kommunikation in Fernsehnachrichten. In K. Kamps & M. Meckel (Hrsg.), *Fernsehen: Prozesse, Strukturen, Funktionen* (S. 33–48). Opladen/Wiesbaden: Westdeutscher Verlag.
- Kemper, S. & Sumner, A. (2001). The structure of verbal abilities in young and older adults. *Psychology & Aging*, 16(2), 312–322.
- Koch, P. & Oesterreicher, W. (1985). Sprache der Nähe — Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. *Romanistisches Jahrbuch*, 36, 15–43.
- Krüger, U. M. (2009). Altersgerechte Nachrichten für Kinder. Programmanalyse der Kindernachrichtensendung «logo!». *Media Perspektiven*, 11, 577–592.
- Krüger, U. M. (1998). Zwischen Konkurrenz und Konvergenz: Fernsehnachrichten öffentlich-rechtlicher und privater Rundfunkanbieter. In K. Kamps & M. Meckel (Hrsg.), *Fernsehen: Prozesse, Strukturen, Funktionen* (S. 65–84). Opladen/Wiesbaden: Westdeutscher Verlag.
- Krüger, U. M., Zampf-Schramm, T. & Jung, M. (2018). Sendungsformen, Themen und

- Akteure im Nonfictionangebot von Das Erste, ZDF, RTL und Sat.1. *Media Perspektiven*, 5, 242–262.
- Landmeier, C. & Daschmann, G. (2011). Im Seichten kann man ertrinken? Boulevardisierung in der überregionalen deutschen Qualitätspresse. In R. Blum, H. Bonfadelli, K. Imhof & O. Jarren (Hrsg.), *Krise der Leuchttürme öffentlicher Kommunikation. Vergangenheit und Zukunft der Qualitätsmedien* (S. 177–191). Wiesbaden: VS.
- Lange, B. P. (2012). *Verbal proficiency as fitness indicator: Experimental and comparative research on the evolutionary psychology of language and verbal displays*. Saarbrücken: Südwestdeutscher Verlag für Hochschulschriften.
- Lange, B. P. (2016). Befunde zur Rolle sprachlicher Gewandtheit bei der Partnerwahl. In C. Hennighausen, B. P. Lange & F. Schwab (Hrsg.), *Evolution des Sozialen* (S. 128–143). Lengerich: Pabst Science Publishers.
- Lange, B. P., Bögemann, H. & Zaretsky, E. (2017). Ästhetische Dimensionen von Sprache, Sprechen, Stimme. In C. Schwender, B. P. Lange, & S. Schwarz (Hrsg.), *Evolutionäre Ästhetik* (S. 225–246). Lengerich: Pabst Science Publishers.
- Lange, B. P., Hennighausen, C., Brill, M. & Schwab, F. (2016). Only cheap talk after all? New experimental psychological findings on the role of verbal proficiency in mate choice. *Psychology of Language and Communication*, 20(1), 1–22.
- Lange, B. P., Kouros, S. & Schwab, F. (2019). Schön gesagt! Aspekte der Gewandtheit der Nachrichtensprache: Ein empirischer Vergleich der sprachlichen Gewandtheit der Nachrichtensendungen von ARD, ZDF, SAT.1 und RTL. *Medien und Kommunikationswissenschaft*, 67(1), 45–62.
- Lange, B. P., Zaretsky, E., Schwarz, S. & Euler, H. A. (2014). Words won't fail: Experimental evidence on the role of verbal proficiency in mate choice. *Journal of Language and Social Psychology*, 33(5), 482–499.
- Lobinger, K. (2012). *Visuelle Kommunikationsforschung. Medienbilder als Herausforderung für die Kommunikations- und Medienwissenschaft*. Wiesbaden: Springer VS.
- Mangold, R. (2004). Infotainment und Edu-tainment. In R. Mangold, P. Vorderer & G. Bente (Hrsg.), *Lehrbuch der Medienpsychologie* (S. 527–542). Göttingen: Hogrefe.
- McQuail, D. (2001). Television news research: Retrospect and prospect. In K. Renckstorff, D. McQuail & N. Jankowski (Hrsg.), *Television news research: Recent European approaches and findings* (S. 393–404). Berlin: Quintessenz-Verlag.
- Meckel, M. & Kamps, K. (1998). Fernsehnachrichten: Entwicklung in Forschung und Praxis. In K. Kamps & M. Meckel (Hrsg.), *Fernsehen: Prozesse, Strukturen, Funktionen* (S. 11–29). Opladen/Wiesbaden: Westdeutscher Verlag.
- Merten, K. (1995). *Inhaltsanalyse: Einführung in Theorie, Methode und Praxis* (2., verbesserte Auflage). Wiesbaden: Springer Fachmedien.
- Neuberger, C. (2018). Journalismus in der Netzwerköffentlichkeit: Zum Verhältnis zwischen Profession, Partizipation und Technik. In C. Nuernbergk & C. Neuberger (Hrsg.), *Journalismus im Internet: Profession – Partizipation – Technisierung* (S. 11–80). Wiesbaden: Springer.
- Newman, N., Fletcher, R., Kalogeropoulos, A., Levy, D. A. L. & Nielsen, R. K. (2018). *Reuters Institute Digital News Report 2018*. Reuters Institute for the Study of Journalism. Abgerufen von <https://www.digitalnewsreport.org/survey/2018/>.
- Nippold, M. A., Ward-Loneragan, J. M., & Fanning, J. L. (2005). Persuasive writing in children, adolescents, and adults: A study of syntactic, semantic, and pragmatic development. *Language, Speech, and Hearing Services in Schools*, 36(2), 125–138.
- Paetzold, U. (1973). Wie objektiv können Nachrichten sein. *Gewerkschaftliche Monatshefte*, 24(8), 488–495.
- Pennebaker, J. W., Francis, M. E. & Booth, R. J. (2001). *Linguistic Inquiry and Word Count: LIWC 2001*. Mahwah, NJ: Erlbaum Publishers.
- Pennebaker, J. W., Chung, C., Ireland, M., Gonzales, A. & Booth, R. J. (2007). *The development and psychometric properties of LIWC 2007*. Austin: University of Texas.
- Pennebaker, J. W. & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology*, 77(6), 1296–1312.

- Perneger, T. V. (1998). What's wrong with Bonferroni adjustments. *British Medical Journal*, 316(7139), 1236–1238.
- Rankl, D. (2014). *Ressorts im Zeitalter im Zeitalter von Medienwandel und Boulevardjournalismus: Themenwahl als Qualitätskriterium*. Wiesbaden: Springer Gabler.
- Rothman, K. J. (1990). No adjustments are needed for multiple comparisons. *Epidemiology*, 1(1), 43–46.
- Schwiesau, D. & Ohler, J. (2016). *Nachrichten – klassisch und multimedial*. Wiesbaden: Springer VS.
- Semetko, H. & Valkenburg, P. M. (2000). Framing European politics: A content analysis of press and television news. *Journal of Communication*, 50(2), 93–109.
- Stanyer, J. (2008). Web 2.0 and the transformation of News and Journalism: New possibilities and challenges in the Internet Age. In A. Chadwick and P. N. Howard (Hrsg.), *The Handbook of Internet Politics* (S. 201–214). New York: Routledge.
- Storrer, A. (2013). Sprachstil und Sprachvariation in sozialen Netzwerken. In B. Frank-Job (Hrsg.), *Die Dynamik sozialer und sprachlicher Netzwerke. Konzepte, Methoden und empirische Untersuchungen an Beispielen des WWW* (S. 331–366). Wiesbaden: Springer VS.
- Unz, D. & Schwab, F. (2004). Nachrichten. In R. Mangold, P. Vorderer & G. Bente (Hrsg.), *Lehrbuch der Medienpsychologie* (S. 493–525). Göttingen: Hogrefe.
- Uribe, R. & Gunter, B. (2007). Are »sensational» news stories more likely to trigger viewers' emotions than non-sensational news stories? A content analysis of British TV news. *European Journal of Communication*, 22(2), 207–228.
- Wahdat, A. (2009). Die Welt im Zeitraffer. Die Sprache und das Sprechen. In N. Matzen & C. Radler (Hrsg.), *Die Tagesschau. Zur Geschichte einer Nachrichtensendung* (S. 149–167). Konstanz: UVK.
- Weiss, J. (2003). *Das Internet und die klassischen Medien: Konvergenz – Konkurrenz oder Komplementierung? Eine medienpolitische Betrachtung*. Frankfurt am Main: Peter Lang.
- White, D. M. (1950). The »gate keeper»: A case study in the selection of news. *Journalism Quarterly*, 27(4), 383–390.
- Wimmer, G. (2005). The type-token relation. In R. Köhler, G. Altmann & R. G. Piotrowski (Hrsg.), *Quantitative linguistics. An international Handbook* (S. 361–368). Berlin: de Gruyter.
- Winterhoff-Spurk, P., Unz, D. C. & Schwab, F. (2005). Häufiger, schneller, variabler. Ergebnisse einer Längsschnittuntersuchung über Gewalt in TV-Nachrichten. *Publizistik*, 50(2), 225–237.
- Wittwen, A. (1995). *Infotainment: Fernsehnachrichten zwischen Information und Unterhaltung*. Bern: Peter Lang.
- Wolf, M., Horn, A. B., Mehl, M. R. & Haug, S. (2008). Computergestützte quantitative Textanalyse: Äquivalenz und Robustheit der deutschen Version des Linguistic Inquiry and Word Count. *Diagnostica*, 54(2), 85–98.
- Zubayr, C. & Gerhard, H. (2018). Tendenzen im Zuschauerverhalten. *Media Perspektiven*, 3, 102–117.