

ADVERBS AND ADVERBIALS AS INDICATORS OF OBJECTIVITY IN GERMAN AUDIO DESCRIPTION SCRIPTS

EVA SCHAEFFER-LACROIX

SORBONNE UNIVERSITÉ | RESEARCH GROUP STIH (SENS TEXTE INFORMATIQUE HISTOIRE)



Sens Texte
Informatique
Histoire

International Conference *Discussing the Limits of Objective Audio Description*
5-7 October 2022
Institut für Übersetzungswissenschaft und Fachkommunikation
Universität Hildesheim (Germany)

OUTLINE

- Definition of "adverb" and "adverbial"
- Role of adverbs and adverbials for audio description (AD)
- Research question, hypotheses, and methods
- AD script corpus Buettenwarder
- TreeTagger annotation of adverbs and adverbials
- Adverbs versus adverbially used adjectives
- Data analysis
 - Quantification
 - Semantic categorisation
- Results
- Conclusion

DEFINITION OF "ADVERB" AND "ADVERBIAL"

- Adverb

- Closed class of endingless words: *da* [there], *heute* [today], *vielleicht* [maybe]

- Adverbial

- Words or phrases that serve as adverbs: *langsam* [slow], *grinsend* [grinning], *nach oben* [upwards]

(Drosdowski & Eisenberg, 1995, p. 335)

ROLE OF ADVERBS AND ADVERBIALS FOR AUDIO DESCRIPTION

- Adverbs support the task of condensing information and thus play a space-saving role for audio description (AD).
- Adverbs can provide a bird's eye view of a scene:
"Un flottement au-dessus de la scène [est] provoqué par la présence d'un adverbe." (Larrory 2022)
- Adverbs are likely to contribute to present AD events in an objective or subjective manner.

RESEARCH QUESTIONS (Q), HYPOTHESES (H) AND METHODS (M)

Q1: To which extent do adverbs and adverbials reflect objectivity in German AD scripts?

Q2: To which extent is emotion audiodescribed in a TV show not explicitly aimed at children?

- H1: The frequency of adverbs and adverbials in the recorded audio description part is an indicator for objectivity.
- H2: The meaning of the adverbs and adverbials is an indicator for objectivity.
- M1: Automatically annotate adverbs and adverbials in an AD script corpus with the help of the TreeTagger
- M2: Quantify the results obtained with the help of TXM
- M3: Classify the retrieved adverbs and adverbials into semantic categories

TV SHOW *NEUES AUS BÜTTENWARDER* [NEWS FROM BUETTENWARDER]

(Eberlein 2017)

Current episodes



Episode 12: Goethe year



Episode 16: Missing



Episode 72: Oh!

AD script corpus: episodes 20, 45, 52-73

Size: 71.850 words

Structures: POS (parts of speech), XML (extensible markup language)

AD SCRIPT (EPISODE 60)

10:23:29

"Ja, herzlichen Glückwunsch."
ss Adsche gibt Brakelmann
die Hand.

10:23:44

"Oh, eine Rede. Nee, dann
lieber singen."
Genervt setzt Schönbiehl
sich wieder.

10:23:29

"Well, congratulations!"
ss Adsche shakes
Brakelmann's hand.

10:23:44

"Oh, a talk. Nope, I would
rather sing a song."
Irritated, Schönbiehl takes
his seat back.



AD SCRIPT SECTIONS

Section	XML tag	Samples from episode 12
Speaker (text to be recorded)	<sp>	Brakelmann fährt durchs Dorf. [Brakelmann drives through the village.]
Stage (how to record the text)	<stage>	(Rest übersprechen) [Override the rest]
Stage type (speech rate)	<stage type="delivery">	s, ss, n [schnell, sehr schnell, normal] [quickly, very quickly, normal]
Prompt (words of the film dialogues that precede ADs)	<prompt>	"Goethe gehört mir!" [Goethe belongs to me!"]

KIRSTEN'S IDENTIFICATION OF ADVERBS AND ADVERBIALS

Les adverbos

SCr [sure]

?

ne, ja, nachdenklich
-adv
particules

Hörfilm
Neues aus Büttenwarder
60. Einladung
Text:
Martin Ovelgönne
November 2014

Sprecherangaben:

s
ss
→n
#

(noue mots elliptes?)

schnell sprechen
sehr, schnell sprechen
wieder in normalem Tempo sprechen
Szenenwechsel innerhalb des Textes,
weiterlesen, durch Betonung deutlich machen

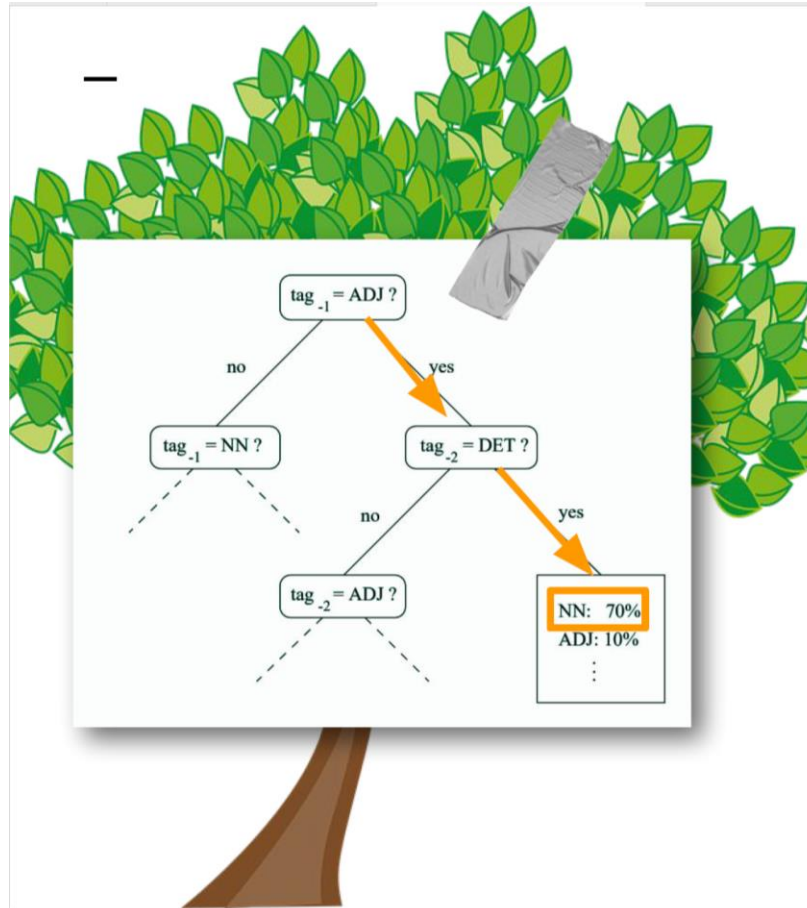
[speak fast]
[speak very fast]

(Geräusch)

Geräusch, freilassen

10:00:00
(KURZ MUSIK)

TREETAGGER



Decision Trees

Gesucht: $P(\text{NN}|\text{DET}, \text{ADJ})$

= 0,7

- Sobald zweimal ja gesagt wird, wird der Wahrscheinlichkeitsvektor ausgegeben
- Beispiel STTS: Maximal 106 Tests für 54 Tags
- Entscheidungsbäume werden durch einen Algorithmus berechnet, der die Tests so wählt, dass möglichst wenig Schritte gemacht werden müssen

Source:
[Deborah Watty](#)

PART-OF-SPEECH TAGS IN THE BUETTENWARDER CORPUS

Tag	Definition	Example	Translated example
ADV	Adverb	schon, bald	already, soon
ADJD	Adverbially used adjective	[sie fahren] schnell	[they drive] fast
PAV	Pronominal adverb	dafür	for this
ADVAD	AD-specific adverb	s, ss, n	f, ff, n [fast, very fast, normally]

[Stuttgart-Tübingen-TagSet \(STTS\)](#)

ADVERBS VERSUS ADVERBIALY USED ADJECTIVES

adverbial use (uninflected form)	predicative use (uninflected form)	attributive use (inflected form)	example	translated example
ADV			• lediglich	only
ADV	ADV		• (ist) vergebens	(is) futile
ADV meaning 1		ADJA meaning 2	• nämlich • nämliche...	1. namely 2. the appointed xxx
ADV meaning 1	ADJD meaning 2	ADJA meaning 2	• eben • (ist) eben • ebene...	1. just, a second ago 2. (is) flat 2. the flat xxx
ADV meaning 1	ADJD meaning 2	ADJA meaning 2	• wahrscheinlich • (ist) wahrscheinlich • wahrscheinliche...	1. probably 2. (is) probable 2. the probable xxx
	ADJD	ADJA	• (ist) schuldig • schuldige...	• (is) guilty • the guilty xxx
		ADJA	• obere...	the upper xxx

(Building on Zinsmeister & Heid, 2003)

TREETAGGER ANNOTATION TEST

" \$" "
Oh ITJ oh
, \$, ,
eine ART eine
Rede NN Rede
. \$. .
Nee PTKANT Nee
, \$, ,
dann ADV dann
lieber ADJD lieb ADV lieber
singen VVINF singen
. \$. .
" \$" "
Genervt VVPP nerven genervt ADJD
setzt VVFIN setzen
Schönbiehl NE Schönbiehl
sich PRF sich
wieder ADV wieder
. \$. .

" \$" "
Oh ITJ oh
, \$, ,
a ART a
talk NN talk
. \$. .
Nope PTKANT nope
, \$, ,
then ADV then
nicer ADJD nice ADV rather
sing VVINF sing
. \$. .
" \$" "
Irritated VVPP irritate irritated ADJD
sits (down) VVFIN sit (down)
Schönbiehl NE Schönbiehl
himself PRF himself
again ADV again
. \$. .

TXM (OPEN-SOURCE TEXTUAL ANALYSIS SOFTWARE)

Selected features

- Import, annotate, and analyse your own corpus
- Divide the corpus into several parts
- Display the size of the sub-corpora
- Calculate the frequency of a form or pattern
- Calculate the co-occurrence between two forms or patterns
- Calculate the specificity of the vocabulary of a text
- Calculate the lexical distance between four texts or more
- Localise one or more forms or patterns in the corpus
- Display the text or video segment corresponding to a word or pattern



(Heiden et al. 2010)

TXM INTERFACE

Er umklammert die Decke.

10: 14: 12 " Und du? ... Ja und nicht so, wie sonst immer. " (Poltern frei) 10: 14: 24

Alle drehen sich um. Der Blick fällt auf eine große Kiste **draußen** vor dem Fenster. Ein weißer Transporter fährt weg. Vor der Kneipe stehen die Männer um die hüfthohe, würfelförmige Holzkiste.

default [Navigation icons] 5 / 9

BUETTENWARDERWITHSPV23/@depos ≤71 850 /... BUETTENWARDERWITHSPV23/<[depos="ADV"]> ☒

Requête ✎ [depos="ADV"]

text_id, sp_n	Contexte gauche	Pivot	Contexte droit
20, null	" Und du ?... Ja und nicht	so	, wie sonst immer. " (Poltern frei) 10 :
20, null	?... Ja und nicht so, wie	sonst	immer. " (Poltern frei) 10 : 14 : 24
20, null	... Ja und nicht so, wie sonst	immer	. " (Poltern frei) 10 : 14 : 24 Alle
20, 59	. Der Blick fällt auf eine große Kiste	draußen	vor dem Fenster. Ein weißer Transporter fährt weg. Vor der

[The gaze falls on a large box **outside** the window. A white van drives away.]

QUANTIFICATION: ABSOLUTE AND WEIGHTED* FREQUENCIES

	ADV	Wt.	ADJD	Wt.	ADVAD	Wt.	PAV	Wt.
prompt	1752	1000	524	11,3	1	-107,4	60	2,1
speaker	436	-185,3	725	0,8	9	-202,6	103	4
stage	136	-3	181	9,7	779	1000	1	-4,5
Total	2324		1430		789		164	

*The weighting is calculated with the TXM function "spécificités" (abbreviated as "wt."). It results from the comparison of the absolute frequency found for each section with its frequency in the corpus as a whole.

SEMANTIC CATEGORISATION: FEATURES OF ADVERBS AND ADVERBIALS

Original citation

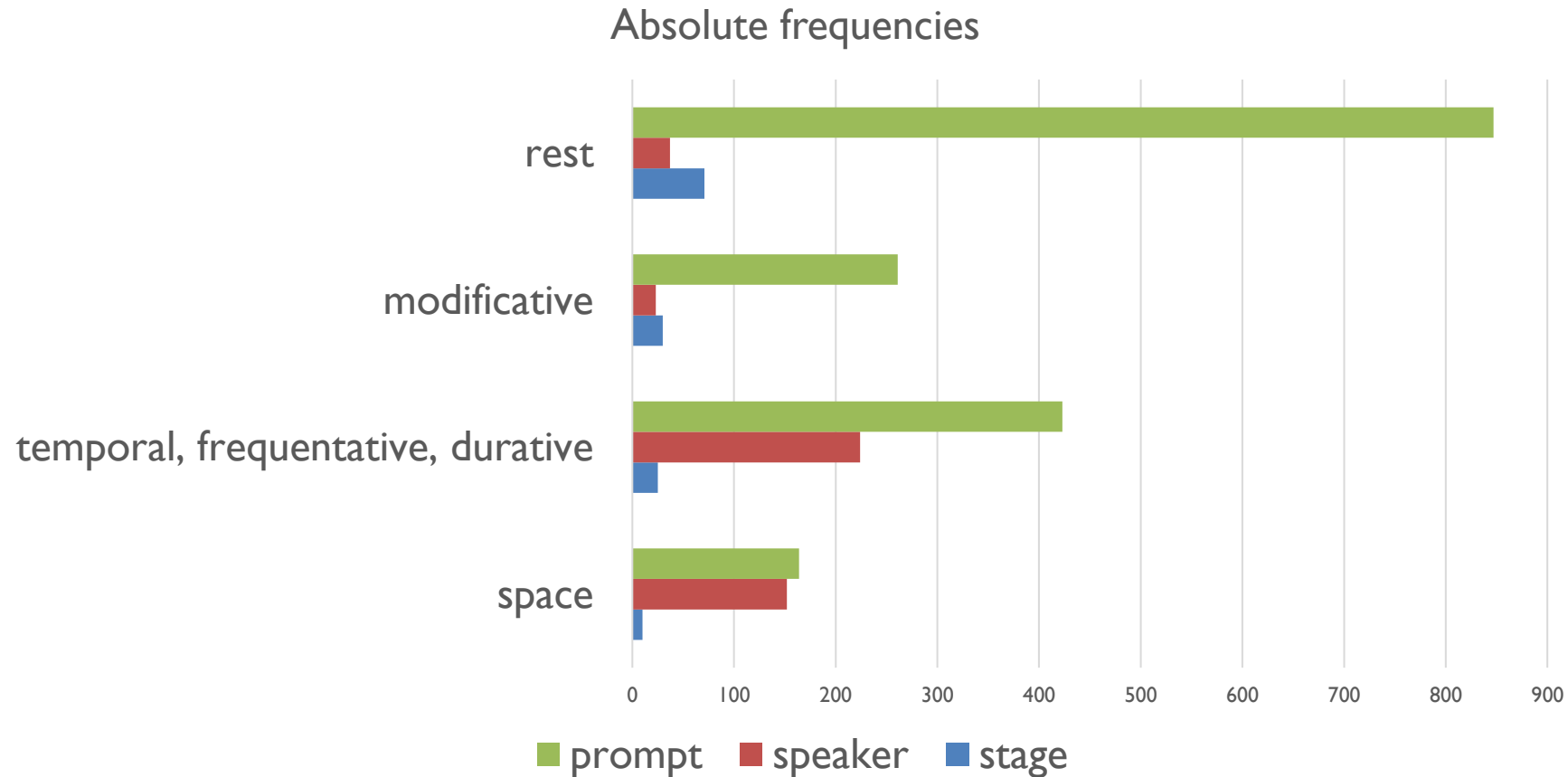
"Die prototypische Funktion von Adverbien besteht darin, ein Geschehen, einen Gegenstand oder einen Sachverhalt semantisch zu modifizieren. Die Art der Modifikation oder Spezifikation richtet sich nach den semantischen Eigenschaften des Adverbs: räumliche Spezifikation (oben liegen), zeitliche (die Aufführung morgen), konzessive (trotzdem nicht den Mut verlieren), **modale** (blindlings in sein Unglück rennen) usw."

Translated citation

"The prototypical function of adverbs is to semantically modify an event, an object, or a state of affairs. The type of modification or specification depends on the semantic properties of the adverb: spatial specification (lie on top), temporal (the performance tomorrow), concessive (do not lose heart despite everything), **modal** (run blindly into one's misfortune), etc."

Source: Adverb. In: Leibniz-Institut für Deutsche Sprache: "Systematische Grammatik". Grammatisches Informationssystem grammis.

SEMANTIC CLASSIFICATION OF ADV FORMS



SPATIAL AND TEMPORAL EVENTS

ADV in the speaker section	Form	Absolute frequency	Spatial events	Temporal events, duration, iteration
wieder	<i>again</i>	49		49
Dann dann	<i>Then then</i>	31		31
Draußen draußen	<i>Outside outside</i>	27	27	
Heute	<i>Today</i>	23		23
oft	<i>often</i>	22		22
weiter	<i>further</i>	19	2	17
stets	<i>always</i>	16		16
Drinnen	<i>Inside</i>	13	13	
hinten	<i>in the back</i>	13	13	
Abends	<i>in the evening</i>	12		12
noch	<i>still</i>	11		11
Tags	<i>during the day</i>	11		11
vorbei	<i>past</i>	9	9	
vorne	<i>in front</i>	9	9	
nebeneinander	<i>side by side</i>	7	7	
Dort	<i>There</i>	6	6	
oben	<i>on the top</i>	6	6	


MODIFICATIVE EVENTS

Form	Translation	Degree	Judgment	Evaluation	Absolute frequency
Nur nur	<i>only</i>		x		6
etwa	<i>about</i>			x	5
fast	<i>almost</i>	x			3
Etwas etwas	<i>a little, some</i>	x			4
Ganz	<i>Quite</i>	x			2
Kaum	<i>Barely</i>			x	1
sehr	<i>very</i>	x			1
so	<i>so</i>			x	1

ADV AND ADJD FORMS IN THE SPEAKER SECTION

wasserfallartig	<i>waterfall-like</i>
ruckartig	<i>jerkily</i>
großspurig	<i>cockily</i>
griesgrämig	<i>grouchily</i>
beidhändig	<i>with both hands</i>

again 49
 briefly 45
 Today 23
 often 22
 Outside 20
 further 19
 Then 17
 elegantly 16
 always 16
 then 14
 high 14
 perplexed 14
 inside 13
 behind 13
 smiling 13
 in the evening 12
 sceptical 12
 slowly 11
 still 11
 during the day 11
 Smiling 10
 satisfied 10

Requête  [depos="ADV|ADJD"]

word	Fréquence
wieder	49
kurz	45
Heute	23
oft	22
Draußen	20
weiter	19
Dann	17
elegant	16
stets	16
dann	14
hoch	14
verdutzt	14
Drinnen	13
hinten	13
lächelnd	13
Abends	12
skeptisch	12
langsam	11
noch	11
Tags	11
Lächelnd	10
zufrieden	10

ADV AND ADJD FORMS IN THE STAGE SECTION

frei 73
bisschen 25
kurz 23
Ja 19
KURZ 17
FREI 13
langsam 11
so 7
mal 6
ja 5
doch 4
gut 4
So 4
Also 3
auch 3
ganz 3
GUT 3
hier 3
immer 3
noch 3

×

free 73
a little 25
short, briefly 23
Yes 19
SHORT, BRIEFLY 17
FREE 13
slowly 11
like this 7
sometimes 6
yes 5
but 4
well 4
So 4
So 3
also 3
quite 3
GOOD 3
here 3
always 3
still 3

SEMANTIC CLASSIFICATION OF ADJD FORMS (ADVERBIALS) IN THE SPEAKER SECTION

word class	local, directional	temporal, frequentative, durative	modificative		total
			degree, evaluation, judgment	facial expression, emotion	
participle I	29	1	0	91	121
rest	118	99	83	305	605
total	147	100	83	396	726

EMOTION IN THE SPEAKER SECTION

Example from episode 53

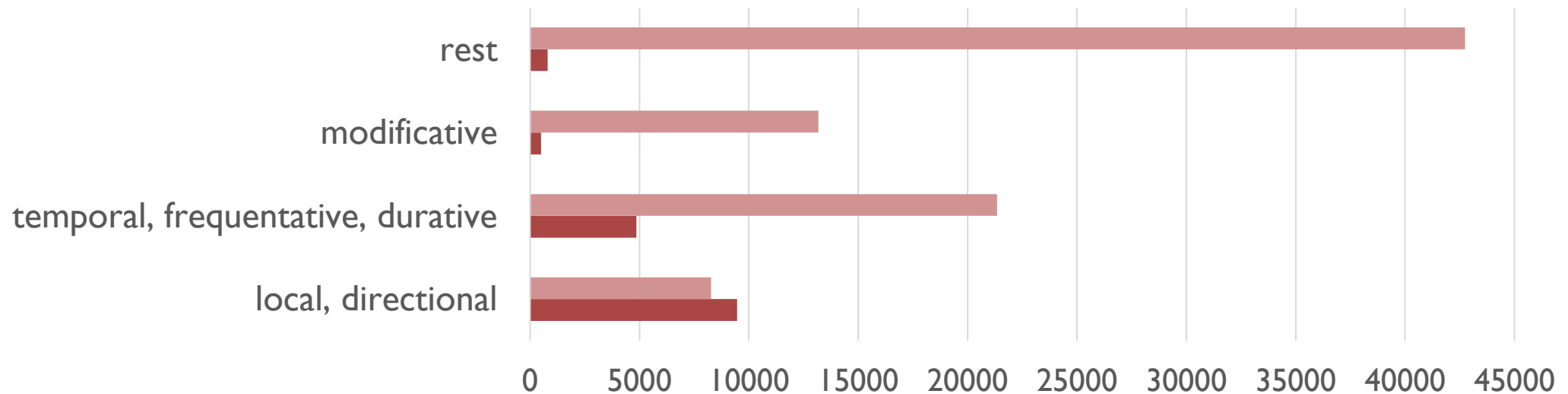
Energisch schneiden die beiden Zweigstücke aus den Flügeln raus. Adsche steht auf und geht einen Schritt zurück. Er beäugt sein Werk **skeptisch** und blickt **verunsichert** zu Krischan. Der schnippelt **konzentriert** weiter. Adsche hockt sich zurück an seinen Flügel und arbeitet weiter.

Translated example

Vigorously, the two cut pieces of twigs out of the wings. Adsche stands up and takes a step back. He eyes his work **sceptically** and looks **unsecurely** at Krischan who continues to snip with **concentration**. Adsche crouches back at his wing and continues working.

RELATIVE FREQUENCY OF ADV FORMS IN THE SECTIONS PROMPT AND SPEAKER

Relative frequencies (words per million)



	local, directional	temporal, frequentative, durative	modificative	rest
prompt	8276,5	21347,4	13171,8	42745,3
speaker	9457,7	4859	498,9	802,6

RESULTS

- Stage section
 - Contains few weakly varying adverbs and adverbials covering two semantic concepts (degree and frequency).
- Prompt section
 - Contains a high number of adverbs and adverbials covering a wide range of concepts.
 - Contains dialogue endlines, which causes a bias for section analysis.
 - Contains a high number of parts-of-speech which share linguistic features with adverbs and adverbials (conversational particles, modal particles, degree particles and interjections), which causes TreeTagger annotation errors.
- Speaker section
 - Contains a modest proportion of adverbs and adverbials covering a wide range of concepts (place, time, degree, frequency, concomitance, etc.).

CONCLUSIONS

- Sound method to analyse the degree of objective and subjective discourse in audio description scripts
- A challenging task: share with blind or visually impaired people non-verbal messages needing interpretation

ACKNOWLEDGEMENTS

My warmest thanks go to :

- the NDR and to Ursula Heerdegen-Wessel, who helped me obtain the Büttenwarder data;
- Kirsten Berland for her contribution to the annotation of the data;
- Céline Largier Vié, who supported me in my first language analysis steps of this data set.
- You: thank you for listening!

REFERENCES (I)

- Abraham, W./Leiss, E. (Hrsg.) (2009): Modalität: Epistemik und Evidentialität bei Modalverb, Adverb, Modalpartikel und Modus. Tübingen: Stauffenburg. (= Studien zur deutschen Grammatik Bd. 77).
- Arndt, W. (1960): “Modal Particles” in Russian and German. In: *WORD* 16, S. 323–336.
- Doval, I. (2017): POS-tagging a bilingual parallel corpus: methods and challenges. In: *Research in Corpus Linguistics*, S. 35–46.
- Drosdowski, G./Eisenberg, P. (Hrsg.) (1995): Duden, Grammatik der deutschen Gegenwartssprache. 5., völlig neu bearbeitete und erw. Aufl. Mannheim: Dudenverlag. (= Der Duden in 12 Bänden Bd. 4).
- Eberlein, N. (1997): Neues aus Büttenwarder. https://www.ndr.de/mediathek/mediatheksuche105_broadcast-129.html.
- Geuder, W. (2019): Eine Art Wortart: Das Adverb im Deutschen. In: *Zeitschrift für Sprachwissenschaft* 38, S. 191–241.
- Heiden, S. (2010): The TXM Platform: Building Open-Source Textual Analysis Software Compatible with the TEI Encoding Scheme. In: *Proceedings of the 24th Pacific Asia Conference on Language, Information and Computation*. Tohoku University, Sendai, Japan: Institute of Digital Enhancement of Cognitive Processing, Waseda University. S. 389–398. <https://www.aclweb.org/anthology/Y10-1044> (Abruf 20.11.2019). Hirvonen, M. (2013): Perspektivierungsstrategien und -mittel kontrastiv - Die Verbalisierung der Figurenperspektive in der deutschen und finnischen Audiodeskription. In: *trans-kom - Journal of Translation and Technical Communication Research* 6, S. 8–38.
- Lang, E./Zifonun, G. (Hrsg.) (1996): Partikelverben im deutsch-englischen Vergleich. In: *Deutsch - Typologisch*. De Gruyter, S. 261–288.
- Leibniz-Institut für Deutsche Sprache (2018a): Adverb. In: *Grammatisches Informationssystem grammis*. <https://grammis.ids-mannheim.de/progr@mm/5293> (Abruf 6.8.2021).

REFERENCES (2)

- Leibniz-Institut für Deutsche Sprache (2018b): Semantisch begründete Adverb-Subklassen. In: „Wissenschaftliche Terminologie“. Grammatisches Informationssystem grammis. <https://grammis.ids-mannheim.de/progr@mm/5278> (Abruf 6.8.2021).
- Lüdeling, A. (2001): On particle verbs and similar constructions in German. Stanford, CA: CSLI Publications. (= Dissertations in linguistics).
- Norddeutscher Rundfunk (2019): Vorgaben für Audiodeskriptionen. In: NDR. https://www.ndr.de/fernsehen/barrierefreie_angebote/audiodeskription/Vorgaben-fuer-Audiodeskriptionen,audiodeskription140.html (Abruf 7.3.2020).
- Nuyts, J. (2001): Subjectivity as an evidential dimension in epistemic modal expressions. In: Journal of Pragmatics 33, S. 383–400.
- Palmer, A./Salway, A. (2015): Audio Description on the Thought-Action Continuum. In: Style 49, S. 126–148.
- Revers, N. (2018): Studying the language of Dutch audio description: An example of a corpus-based analysis. In: Translation and Translanguaging in Multilingual Contexts 4, S. 178–202.
- Revers, N./Remael, A./Daelemans, W./Jankowska, A./Szarkowska, A. (2015): The Language of Audio Description in Dutch: Results of a Corpus Study. In: New Points of View on Audiovisual Translation and Media Accessibility. Peter Lang, S. 167–190.
- Schaeffer-Lacroix, E./Berland, K. (2020): Buettenwarder corpus.
- Schiller, A./Teufel, S./Stöckert, C./Thielen, C. (1999): Guidelines für das Tagging deutscher Textcorpora mit STTS (Kleines und großes Tagset). <https://www.ims.uni-stuttgart.de/documents/ressourcen/lexika/tagsets/stts-1999.pdf> (Abruf 15.7.2021).

REFERENCES (3)

- Schmid, H. (1994): Probabilistic Part-of-Speech Tagging Using Decision Trees. In: Proceedings of International Conference on New Methods in Language Processing. Manchester, UK. <http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/data/tree-tagger1.pdf>.
- TEI Consortium (2020): TEI P5: Guidelines for Electronic Text Encoding and Interchange. Zenodo. <https://zenodo.org/record/3413524> (Abruf 16.1.2021).
- Troiano, E./Klinger, R./Padó, S. (2020): Lost in Back-Translation: Emotion Preservation in Neural Machine Translation. In: Proceedings of the 28th International Conference on Computational Linguistics. Barcelona, Spain (Online): International Committee on Computational Linguistics. S. 4340–4354. <https://www.aclweb.org/anthology/2020.coling-main.384> (Abruf 1.5.2021).
- Trotzke, A./Quaglia, S. (2016): Particle topicalization and German clause structure. In: The Journal of Comparative Germanic Linguistics 19, S. 109–141.
- Tyfour, M. (2021): Sprachmacht auf engstem Raum: die Inszenierung der Stadt in den Hörfilmen der Münchner Tatort-Filmserie: eine korpusgeleitete Studie zur Audiodeskription. Berlin: Frank & Timme, Verlag für wissenschaftliche Literatur. (= Easy-plain-accessible vol. 7).
- Volk, M./Graën, J. (2017): Multi-word Adverbs – How well are they handled in Parsing and Machine Translation?
- Westpfahl, S./Schmidt, T./Jonietz, J./Borlinghaus, A. (2017): STTS 2.0. Guidelines für die Annotation von POS -Tags für Transkripte gesprochener Sprache in Anlehnung an das Stuttgart Tübingen Tagset (STTS). IDS Leibniz-Institut für deutsche Sprache. https://ids-pub.bsz-bw.de/frontdoor/deliver/index/docId/6063/file/Westpfahl_Schmidt_Jonietz_Borlinghaus_STTS_2_0_2017.pdf.
- Zinsmeister, H./Heid, U. (2003): Identifying predicatively used adverbs by means of a statistical grammar model. In: Archer, D./Rayson, P./Wilson, A./McEnery, T. (Hrsg.): UCREL Technical Paper. Lancaster. S. 932–939. <http://ucrel.lancs.ac.uk/publications/CL2003/papers/zinsmeister.pdf>.

GRAMMAR REFERENCES

- Adverb. In: Leibniz-Institut für Deutsche Sprache: "Systematische Grammatik". Grammatisches Informationssystem grammis.
DOI: 10.14618/grammatiksystem
Permalink: <https://grammis.ids-mannheim.de/systematische-grammatik/278>
- Semantisch begründete Adverb-Subklassen. In: Leibniz-Institut für Deutsche Sprache: "Propädeutische Grammatik". Grammatisches Informationssystem grammis.
DOI: 10.14618/programm
Permalink: <https://grammis.ids-mannheim.de/progr@mm/5278>