

## POSTPRINT

Nadine Proske

# Digitale Infrastrukturen in der Gesprächsforschung

Auch Linguist\*innen, die gesprochene Sprache untersuchen, kommen schon seit längerem nicht mehr ohne digitale Infrastrukturen aus. Seit Beginn der Gesprochene-Sprache-Forschung werden Gespräche aufgezeichnet und anschließend transkribiert, da die flüchtigen, innerhalb von Bruchteilen von Sekunden stattfindenden Feinheiten des Gesprochenen paradoxerweise nur durch Verschriftung im Detail untersucht werden können. Diese Detailuntersuchungen beschränkten sich im vergangenen Jahrhundert meist auf wenige Einzelbelege für ein untersuchtes Phänomen. Das heißt, die Forschenden hatten den unmittelbaren Überblick über ihre Datenkollektionen und benötigten keine elaborierten digitalen Methoden zu deren Aufbereitung, Annotation und Analyse. Dies hat sich in den letzten beiden Jahrzehnten stark geändert: Es wurden vermehrt gezielt große Datenmengen gesammelt, in Datenbanken organisiert und der Forschungsgemeinschaft zur Nutzung zur Verfügung gestellt. An erster Stelle muss hier das Forschungs- und Lehrkorpus gesprochenes Deutsch (FOLK) genannt werden (vgl. Schmidt 2014). Dieses wird seit 2008 am Leibniz-Institut für Deutsche Sprache (IDS) aufgebaut und ist heute das größte Referenzkorpus für das gesprochene Deutsch.

FOLK enthält derzeit Audio- und Videoaufnahmen mit einer Gesamtdauer von 314 Stunden (Version 2.16). Es sind formelle und informelle Interaktionen unterschiedlichen Typs enthalten – vom privaten Tischgespräch über berufliche Meetings und Unterrichtsstunden bis hin zu öffentlichen Podiumsdiskussionen. Die Transkripte umfassen rund drei Millionen laufende Wörter – das ist zwar sehr wenig im Vergleich zur üblichen Größe von Schriftkorpora (DeReKo hat z. B. 50 Milliarden Wörter), aber sehr groß für die Verhältnisse in der Gesprächsforschung. Die Verschriftlichung richtet sich nach der Minimalversion des gesprächsanalytischen Transkriptionssystems (GAT, vgl. Selting et al. 2009). Mit einem Klick auf die Transkripte kann die Audio- oder Videoaufnahme abgespielt werden, zudem können Datenausschnitte zur weiteren Aufbereitung heruntergeladen werden. Die Suchfunktionen greifen u. a. auf eine Lemmatisierung und ein Part-of-Speech-Tagging (vgl. Westpfahl/Schmidt 2016) zu. Außerdem kann neuerdings mit verschiedenen, teilweise auch didaktisch ausgerichteten Tools aus dem Projekt ZuMult an den Daten gearbeitet werden (vgl. Fandrych et al. 2022).

Anhand von FOLK lassen sich die unterschiedlichsten Fragestellungen untersuchen. Durch die Nutzung von Metadaten zu den Gesprächen können z. B. Aussagen über die regionale Variation von Wörtern (*Appel* vs. *Apfel*), über die altersabhängige Verwendung von Ausdrücken (z. B. *mega* und *krass*) und über

stilistische Unterschiede grammatischer Varianten (*wegen dem* vs. *wegen des*) gemacht werden (vgl. Kaiser/Schmidt 2016). Meine eigene Forschung beschäftigt sich mit grammatischen Strukturen und Wörtern, die für die Mündlichkeit spezifisch sind, die wir also schriftlich nicht oder deutlich seltener verwenden. Das sind z. B. Imperative, die keine Aufforderungen mehr darstellen (*ach komm*, vgl. Proske 2014), und syntaktische Konstruktionen, die oft nicht mehr wörtlich zu verstehen sind, sondern sprecherseitige Bewertungen enthalten (*sich hinstellen und sagen*, vgl. Proske 2019). Aktuell untersuche ich Redebeiträge, die mit einer Konjunktion enden, z. B. mit *aber*. Beispiel (1) zeigt die Verwendung des ›finalen *aber*‹.

(1) Telefongespräch (FOLK\_00407, T\_01, c749)

HR: SIEHT ma den stellplatz wenn ma vorbeifährt?  
weil ich KUCK nämlich als **aber**-  
(0.85)

JM: ja;  
man SIEHT den.

Zunächst ein Wort zur Darstellung: Das Beispiel ist wie für Publikationen üblich nach der Basiskonvention von GAT transkribiert, das heißt, die Minimalversion aus dem Korpus ist um prosodische Informationen wie Fokusakzente (Großbuchstaben) und finale Intonationskonturen (repräsentiert durch Satzzeichen) ausgebaut worden.

In Beispiel (1) ergänzt die Sprecherin HR ihre Frage *Sieht man den Stellplatz, wenn man vorbeifährt?* durch eine Begründung dafür, warum sie sich die Frage nicht selbst beantworten kann: *Ich gucke nämlich als, aber* (*als* bedeutet hier ›immer‹). Die Begründung endet mit *aber* und bleibt formal unvollständig. Es erscheint – sogar auch für uns als Außenstehende, die den Gesprächskontext nicht kennen, – relativ klar, was nach *aber* ergänzt werden müsste, nämlich in etwa: *aber ich habe ihn noch nicht entdecken können*. Die Gesprächspartnerin JM beantwortet die Frage nach einer Pause von einer knappen Sekunde mit *Ja, man sieht den*. Sie wartet also nicht auf die Vervollständigung des *aber*-Satzes. Ihre Antwort zeigt, dass sie HRs Frage und Begründung auch in unvollständiger Form verstanden hat.

Solche Sprecherwechsel nach *aber* kommen häufig vor. Sind dies nun ungeplante Abbrüche, mit denen gut umgegangen werden kann, wenn der Kontext klar ist, oder sind es eher schon gezielt eingesetzte Routinen? Wenn Fälle wie Beispiel (1) überwiegen, in denen es nicht zu Rederechtskonkurrenz, Nachfragen oder Verständnisproblemen kommt, dann liegt eine Interpretation als Routine nahe. Eben dies ist schon für die Äquivalente von *aber* im Englischen und Finnischen beschrieben (Mulder/Thompson 2008; Koivisto 2015). Die Studie zum Deutschen soll zeigen, welche *verschiedenen* Routinen durch das finale *aber* realisiert werden. Zu einem kleinen Ausschnitt der Ergebnisse später mehr.

Wie findet man nun zunächst Belege für das finale *aber* in FOLK? Die Tokensuche nach der normalisierten Form *aber* ergibt 10.000 zufällig ausgewählte Treffer – das Korpus enthält mehr als doppelt so viele Vorkommen des Wortes, aber eine KWIC-Liste ist auf diese maximale Anzahl begrenzt (zu den Einzelheiten der Korpusrecherche vgl. Kaiser 2016). Die Trefferliste enthält Belege für *aber* in allen möglichen Positionen: zwischen zwei Hauptsätzen wie in (2), zu Beginn eines reaktiven Redebeitrags wie in (3), im Mittelfeld des Satzes als Modalpartikel wie in (4) und schließlich am Ende eines inhaltlich unvollständigen Redebeitrags wie in (1).

(2) Mädelsabend (FOLK\_00287, T\_01, c331)

IR: ich würd\_s wahrscheinlich nie im leben beNUTzen;  
**aber**\_s sieht hammerGEIL aus.

(3) Backen (FOLK\_00331, T\_03, c496)

RA: ich RIECH nix.  
CA: **aber** ICH riech grade karaMELL.

(4) Fahrschulstunde (FOLK\_00415, T\_01, c135)

HM: das sind **aber** keine fünfundvierzig GRAD;  
STOPP martin.

Alle 10.000 Belege durchzusehen und zu bestimmen, bei welchen davon *aber* final steht, wäre sehr aufwändig. Es kommt hinzu, dass man am Ende nicht alle im Korpus vorhandenen Belege fürs finale *aber* gefunden hätte, weil man die anderen 10.000 *aber*-Vorkommen gar nicht angezeigt bekommen hat. Deshalb bietet sich ein anderes Vorgehen an: Man gibt bei der Tokensuche eine kontextuelle Bedingung an – *aber* soll am Ende stehen. Aber am Ende wovon? Das Korpus enthält keine Interpunktion zur Anzeige von Satzgrenzen, weil diese bei mündlichen Daten schwer zu bestimmen sind (siehe aber Westpfahl/Gorisch 2018). Es enthält aber Kategorien zur Bestimmung der Position innerhalb eines Redebeitrags, unabhängig davon, ob der Redebeitrag aus einem oder mehreren Sätzen besteht. Für die Suche nach dem finalen *aber* eignet sich die Kategorie »höchstens N Wörter vor Ende eines Beitrags« (mit N=1). Das Suchergebnis beschränkt sich nun auf gut 2.000 Belege, unter denen aber immer noch viele händisch auszusortieren sind. Das liegt an der Definition von »Beitrag«, die für das Korpus genutzt wird: Beiträge werden abgegrenzt durch Sprecherwechsel oder durch Pausen ab 0,2 Sekunden Länge. Das heißt, auch Belege, bei denen ein\*e Sprecher\*in nach *aber* eine kurze Pause macht, dann aber weiterspricht, sind in den Suchergebnissen enthalten. Dies ist auch gewünscht, da sie dann zu den finalen

Fällen gehören, wenn die Fortsetzung nach der Pause syntaktisch inkohärent ist, also z. B. ein neuer Satz begonnen wird. Es sind aber auch viele Belege dabei, bei denen die Pause mitten in einem wohlgeformten Satz auftritt (z. B. *das wär aber (0.36) PECH für mich*). Diese müssen aussortiert werden, weil sie wie die Beispiele (2) bis (4) gar nicht die finale Position in einem Redebeitrag im gesprächsanalytischen Sinn einnehmen. Die Bestimmung des Umfangs eines solchen Redebeitrags ist weitaus komplexer, als es in der Aufbereitung eines Korpus berücksichtigt werden kann (vgl. Deppermann/Proske 2015). Dennoch bieten die Aufbereitung und Suchmöglichkeiten des Korpus einen Ausgangspunkt für eigene weitere Datenaufbereitungen, ohne den die Bearbeitung größerer Belegkollektionen zu einem Phänomen kaum handhabbar wäre, schon gar nicht für einzelne Forschende oder kleine Forschungsgruppen.

Für die noch laufende Studie wurden bisher rund 700 Belege durchgesehen, etwa ein Drittel davon (n=238) haben sich als Fälle von finalem *aber* herausgestellt, die dann näher untersucht wurden. Für die Durchsicht und Kodierung der Belege wurde die Trefferliste aus FOLK in Excel importiert. So kann für alle interessierenden Merkmale eines Belegs die entsprechende Ausprägung festgehalten werden. Anschließend kann über die Filter-Funktion die Häufigkeit einer Merkmalsausprägung und auch das gemeinsame Auftreten mit Ausprägungen weiterer Merkmale abgefragt werden. Unter den Merkmalen, die kodiert wurden, sind z. B. die lautliche und prosodische Realisierung von *aber* (lange oder kurze Vokale, gleichbleibender oder fallender Tonhöhenverlauf), das Vorhandensein von Pausen vor und nach *aber* und die Art der Reaktion auf den *aber*-Redebeitrag (äußert der\*die Gesprächspartner\*in nur eine Partikel wie *hm* oder *ja* oder formuliert er\*sie einen satzförmigen Redebeitrag?). Was sich nicht so eindeutig kodieren lässt, sind Funktionen, die das finale *aber* übernimmt, weil oft mehrere zugleich auftreten. Diese sind aber von zentralem Interesse sowohl für die qualitative Analyse als auch für den möglichen statistischen Zusammenhang mit klarer kodierbaren Merkmalen. Oft führt diese Schwierigkeit zu mehreren Kodierdurchläufen, bis man sich für eine Lösung entschieden hat, die am besten zur Fragestellung der Studie passt; dies kann z. B. eine feinkörnige, separate Kodierung aller Funktionen für jeden Beleg sein oder aber eine Beschränkung auf eine abstraktere Ebene, die mehre Funktionen umfasst, die nicht näher differenziert werden.

Die Funktionen des finalen *aber* lassen sich auf abstrakter Ebene als Relevanzrückstufung zusammenfassen. Das heißt, der unmittelbar der Konjunktion vorangehende Redebeitragsteil wird in den Hintergrund gerückt und die nicht verbalisierte, zu erschließende Aussage, auf die *aber* verweist, wird in den Vordergrund gerückt. Darüber hinaus hat das »Weglassen« der eigentlich zentralen Aussage weitere kommunikative Effekte, die recht unterschiedlich sein können: Wenn *aber* auf eine vorerwähnte Äußerung verweist, wie in (5), dann hat es eine Abschlussfunktion. Es wird signalisiert, dass *wie, das weiß man nicht* nicht der letzte, zentrale Gedanke sein soll, sondern dass auf etwas bereits Gesagtes (of-

*fensichtlich bleiben diese Instinkte doch irgendwie erhalten*) reagiert werden soll, dessen Wiederholung aber unnötig wäre. Zugleich entsteht so trotz der Konzession, die *aber* vorausgeht, eine Bekräftigung der Sicherheit über die zentrale Aussage.

(5) WG-Gespräch (FOLK\_00055, T\_09, c111)

NH: offensichtlich bleiben diese instinkte DOCH irgendwie (0.86)  
erHALten.

WIE das WEISS man nich **aber-**  
(1.66)

AM: ich glaub AUCH we\_man ne-

In Fällen ohne Verweis auf vorher Gesagtes ist das Spektrum der Effekte noch größer. Hier nur ein Beispiel: In (6) folgt *aber* auf eine negative Antwort auf eine Frage. Die finale Konjunktion lässt Raum für unterschiedliche nicht ablehnende Ergänzungen (z.B. *ich würde trotzdem mitkommen*) und schwächt damit die Ablehnung etwas ab, ohne aber durch Explikation einer positiveren Aussage diese als relevant für den Fortgang des Gesprächs zu setzen.

(6) Tischgespräch (FOLK\_E\_00052, T\_01, c947)

AM: habt ihr lust in ne JAZZkneipe zu gehen,  
mit LIVEmusik?  
(1.11)

LP: hm.  
(1.26) ich BIN nich so der jazzfan **aber-**  
(0.65)

AM: prag is dafür berühmt lenali

Bei solchen Redebeiträgen, auf deren ersten und einzigen Teil bereits *aber* folgt, reagieren Gesprächspartner\*innen entweder auf den zurückgestuften Teil oder es folgt gar keine Reaktion. Dies ist also anders als in den Fällen wie (5), in denen *aber* einen Teil eines längeren Redebeitrags in den Hintergrund rückt, auf den üblicherweise nicht reagiert wird.

Wenn die qualitative Analyse und Kodierung aller Belege abgeschlossen ist, wird sich zeigen, ob sich auch sinnvolle quantitative Aussagen machen lassen. Einige Aspekte sind leichter »zählbar«, z.B. das Vorhandensein eines Antezedens, d.h. eines Bezugs auf Vorerwähntes. Auch wird sich gut zeigen lassen, wie durchgehend die dem Eindruck nach häufigste prosodische Realisierung von *aber* ist; diese zeichnet sich durch kurze Vokale, gleichbleibende Tonhöhe (in den Transkripten als *aber-* wiedergegeben) sowie Integration in die Intonationskontur der vorausgehenden Äußerung aus. Offen bleiben muss dagegen z.B. noch, ob es einen Zusammenhang zwischen unterschiedlichen prosodischen Realisierungen und bestimmten Funktionen oder Kontexten gibt oder nicht, weil bei einer

diesbezüglichen Quantifizierung zahlreiche Merkmale im Spiel sind, die sich der Intuition entziehen.

Das Fallbeispiel veranschaulicht, dass in der Gesprächsforschung eine gründliche qualitative Arbeit die Voraussetzung ist, um große Datenmengen überhaupt sinnvoll quantifizieren zu können, weil oft keine binären Kategorien vorliegen bzw. solche erst erarbeitet werden müssen. Die digitalen Infrastrukturen ermöglichen die Arbeit mit großen Datenmengen und damit ganz neue Arten von Erkenntnissen überhaupt erst. Die Gesprächsforschung wird dadurch komplexer, aber...

## Literaturverzeichnis

- Deppermann, Arnulf/Proske, Nadine (2015): Grundeinheiten der Sprache und des Sprechens. In: Dürscheid, Christa/Schneider, Jan-Georg (Hrsg.): Handbuch Satz, Äußerung, Schema. Berlin/Boston: de Gruyter. S. 17–47.
- Fandrych, Christian/Frick, Elena/Kaiser, Julia/Meißner, Cordula/Portmann, Annette/Schmidt, Thomas/Schwendemann, Matthias/Wallner, Franziska/Wörner, Kai (2022): ZuMult: Neue Zugangswege zu Korpora gesprochener Sprache. In: Kämper, Heidrun/Plewnia, Albrecht (Hrsg.): Sprache in Politik und Gesellschaft. Perspektiven und Zugänge. Jahrbuch des Instituts für Deutsche Sprache 2021. Berlin: de Gruyter. S. 305–312.
- Kaiser, Julia (2016): Reformulierungsindikatoren im gesprochenen Deutsch. Die Benutzung der Ressourcen DGD und FOLK für gesprächsanalytische Zwecke. In: Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion Nr. 17. S. 196–230.
- Kaiser, Julia/Schmidt, Thomas (2016): Einführung in die Benutzung der Ressourcen DGD und FOLK für gesprächsanalytische Zwecke. Handreichung: Einfache Recherche-Anfragen als Übungsbeispiele. URN: <https://nbn-resolving.org/urn:nbn:de:bsz:mh39-55360>.
- Koivisto, Aino (2015): Taking an interactional perspective on final particles. The case of Finnish *mutta* (>but<). In: Hancil, Sylvie/Haselow, Alexander/Post, Margje (Hrsg.): Final particles. Berlin: de Gruyter. S. 55–76.
- Mulder, Jean/Thompson, Sandra A. (2008): The grammaticization of *but* as a final particle in English conversation. In: Laury, Ritva (Hrsg.): Crosslinguistic studies of clause combining. The multifunctionality of conjunctions. Amsterdam: Benjamins. S. 179–204.
- Proske, Nadine (2014): *°h ach KOMM; hör AUF mit dem kleInkram*. Die Partikel *komm* zwischen Interjektion und Diskursmarker. In: Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion Nr. 15. S. 121–160.
- Proske, Nadine (2019): Emergent pseudo-coordination in spoken German. A corpus-based exploration. In: Yearbook of the German Cognitive Linguistics Association (GCLA) Nr. 7. S. 115–136.
- Schmidt, Thomas (2014): Gesprächskorpora und Gesprächsdatenbanken am Beispiel von FOLK und DGD. In: Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion Nr. 15. S. 196–233.

- Selting, Margret/Auer, Peter/Barth-Weingarten, Dagmar/Bergmann, Jörg/Bergmann, Pia/Birkner, Karin/Couper-Kuhlen, Elizabeth/Deppermann, Arnulf/Gilles, Peter/Günthner, Susanne/Hartung, Martin/Kern, Friederike/Mertzluft, Christine/Meyer, Christian/Morek, Miriam/Oberzaucher, Frank/Peters, Jörg/Quasthoff, Uta/Schütte, Wilfried/Stukenbrock, Anja/Uhmann, Susanne (2009): Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). In: Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion Nr. 10. S. 353–402.
- Westpfahl, Swantje/Gorisch, Jan (2018): A syntax-based scheme for the annotation and segmentation of German spoken language interactions. In: Proceedings of the Joint Workshop on Linguistic Annotation, Multiword Expressions and Constructions (LAW-MWE-CxG-2018). S. 109–120. URL: <https://aclanthology.org/W18-4913.pdf> (18.03.2022).
- Westpfahl, Swantje/Schmidt, Thomas (2016): FOLK-Gold – A GOLD standard for part-of-speech-tagging of spoken German. In: Proceedings of the Tenth Conference on International Language Resources and Evaluation (LREC'16), Portorož, Slovenia. Paris: European Language Resources Association (ELRA). S. 1493–1499.

Dr. Nadine Proske, Leibniz-Institut für Deutsche Sprache (IDS), Mannheim,  
[proske@ids-mannheim.de](mailto:proske@ids-mannheim.de)