



LANGUAGE
AND LAW

LINGUAGEM
E DIREITO

VOLUME 10.1

ISSN 2183-3745

Language and Law

Linguagem e Direito

ISSN: 2183-3745 (online)

Volume 10, Issue 1, 2023

Editors / Diretores

Malcolm Coulthard & Rui Sousa-Silva

Aston University, UK & Universidade do Porto, Portugal

Book Reviews Editors / Editores de Recensões

David Wright (English), Rita Faria (Portugal) & Luciane Fröhlich (Brasil)

Nottingham Trent University, UK, Universidade do Porto, Portugal &

Universidade Federal de Santa Catarina, Brasil

Cover / Capa

Rui Effe

Publisher / Editora

Faculdade de Letras da Universidade do Porto

This publication was supported by project UIDB/00022/2020, FCT-Fundação para a Ciência e Tecnologia, Portugal, through national funds.

Esta publicação é financiada por fundos nacionais através da FCT – Fundação para a Ciência e a Tecnologia, I.P., no âmbito do projeto UIDB/00022/2020.



International Editorial Board / Conselho Editorial Internacional

Janet Ainsworth, *University of Washington, USA*

Carmen Rosa Caldas-Coulthard, *University of Birmingham, UK*

Le Cheng, *Zhejiang University, China*

Virginia Colares, *Universidade Católica de Pernambuco, Brasil*

Diana Eades, *University of New England, Australia*

Debora Figueiredo, *Universidade Federal de Santa Catarina, Brasil*

Maribel del Pozo Triviño, *Universidad de Vigo, Spain*

Ed Finegan, *University of Southern California, USA*

Tammy Gales, *Hofstra University, USA*

Núria Gavaldà, *Universitat Autònoma de Barcelona, Spain*

Maria Lúcia Gomes, *Universidade Tecnológica Federal do Paraná, Brasil*

Tim Grant, *Aston University, UK*

Patrick Juola, *Duquesne University, USA and Juola Associates*

Krzysztof Kredens, *Aston University, UK*

Iman Laversuch, *University of Cologne, Germany*

Janny Leung, *University of Hong Kong, Hong Kong*

Belinda Maia, *Universidade do Porto, Portugal*

Fernando Martins, *Universidade de Lisboa, Portugal*

Karen McAuliffe, *University of Birmingham, UK*

Frances Rock, *Cardiff University, UK*

Paolo Rosso, *Polytechnic University of Valencia, Spain*

Susan Sarcevic, *University of Rijeka, Croatia*

Roger Shuy, *Georgetown University Washington, USA*

Larry Solan, *Brooklyn Law School, USA*

Editorial Assistants / Assistentes Editoriais

Ana Sofia Meneses Silva, *Universidade do Porto, Portugal*

Viviane Maia, *Universidade do Porto, Portugal*

Copyright / Direitos de autor

The articles published in this volume are covered by the Creative Commons “Attribution-NonCommercial” (CC-BY-NC) license (see <http://creativecommons.org>). They may be reproduced in its entirety as long as Language and Law / Linguagem e Direito is credited, a link to the journal’s web page is provided, and no charge is imposed. The articles may not be reproduced in part or altered in form, or if a fee is charged, without the journal’s permission. Copyright remains solely with individual authors. The authors should let the journal Language and Law / Linguagem e Direito know if they wish to republish.

Os artigos publicados neste volume estão cobertos pela licença Creative Commons “Attribution-NonCommercial” (CC-BY-NC) (consultar <http://creativecommons.org>) e podem ser reproduzidos na íntegra desde que seja feita a devida atribuição à Language and Law / Linguagem e Direito, com indicação do link para a página da revista e desde que não sejam cobradas quaisquer taxas. Os artigos não podem ser parcialmente reproduzidos, o seu formato não pode ser alterado, e não podem ser cobradas taxas sem a autorização da revista. Os direitos de autor dos trabalhos publicados nesta revista pertencem exclusivamente aos seus respetivos autores. Os autores devem informar a revista Language and Law / Linguagem e Direito se pretenderem submeter o artigo noutra a outra publicação.

Language and Law / Linguagem e Direito

Language and Law / Linguagem e Direito is a free, exclusively online peer-reviewed journal published twice a year. It is available on the website of the Faculty of Arts of the University of Porto, at <http://ojs.letras.up.pt/index.php/LLLD/>.

All articles should be submitted via the system. See the guidelines for submission at the journal website.

Requests for book reviews should be sent to rssilva@letras.up.pt.

Language and Law / Linguagem e Direito é uma revista gratuita publicada exclusivamente online, sujeita a revisão por pares, publicada semestralmente e disponível no website da Faculdade de Letras da Universidade do Porto, em <http://ojs.letras.up.pt/index.php/LLLD/>.

Os materiais para publicação deverão ser enviados através do sistema, e devem seguir as instruções disponíveis na página da revista.

As propostas de revisão de livros devem ser enviadas para rssilva@letras.up.pt.

Indexing and abstracting / Indexação e bases de dados bibliográficas

Language and Law / Linguagem e Direito is covered by the following abstracting and indexing services:

A Language and Law / Linguagem e Direito encontra-se indexada e catalogada nas seguintes bases de dados:

CrossRef

EBSCO Legal Source

ERIH PLUS: European Reference Index for the Humanities and the Social Sciences

Google Scholar

Journals for Free

JURN

Portal RCAAP

QUALIS Periódicos (A2)

Sherpa Romeo

ProQuest One Academic

Latin America & Iberia Database

Publicly Available Content Database

Linguistics Collection

Social Science Premium Collection

PUBLISHED BIANNUALLY ONLINE / PUBLICAÇÃO SEMESTRAL ONLINE

ISSN: 2183-3745

D.O.I.: <https://doi.org/10.21747/21833745/lanlaw>

D.O.I. Volume 10(1): https://doi.org/10.21747/21833745/lanlaw/10_1

THE ARTICLES ARE THE SOLE RESPONSIBILITY OF THEIR AUTHORS.

THE ARTICLES WERE PEER REVIEWED.

OS ARTIGOS SÃO DA EXCLUSIVA RESPONSABILIDADE DOS SEUS AUTORES.

OS ARTIGOS FORAM SUBMETIDOS A ARBITRAGEM CIENTÍFICA.

Contents / Índice

ARTICLES / ARTIGOS

- Interpretação de normas constitucionais na interface entre semântica e pragmática: estudo de caso de uma decisão do Supremo Tribunal Federal** 1
João Pedro Pádua
- Joke or threat? Competing genre uptakes in a Danish court case** 30
Marie Bojsen-Møller
- Disguise and imitation of language style in WhatsApp messages** 89
Meike M. de Boer, Susan A. Blackwell & Fleur van der Houwen
- Developing a Resource Model of Power and Authority in Anonymous Online Criminal Interactions** 110
Helen Newsome-Chandler & Tim Grant
- Jack Unterweger: An authorship analysis of the notorious killer's autobiography "Fegefeuer oder die Reise ins Zuchthaus"** 131
Karoline Marko & Alesia Locker

Interpretação de normas constitucionais na interface entre semântica e pragmática: estudo de caso de uma decisão do Supremo Tribunal Federal

João Pedro Pádua ¹

¹Universidade Federal Fluminense, Brasil, joapadua@id.uff.br

Resumo

Interpretação de textos normativos é a atividade básica, e talvez mais importante, do/a jurista. Por isso, trabalhos dogmáticos sobre a assim-chamada teoria da interpretação são abundantes na literatura jurídica, especialmente um subcampo dessa teoria, que estuda as particularidades atribuídas à interpretação constitucional. Em busca de conceitos analiticamente mais poderosos, frequentemente a teoria do direito recorre aos estudos da linguagem, nos âmbitos da linguística e da filosofia da linguagem. Porém os usos que se fazem desses conceitos são problemáticos em dois níveis: (i) transformam tais conceitos, originalmente descritivos, isto é, sobre como o/as usuário/as da língua realmente produzem sentido dos discursos, em conceitos normativos, isto é, como esses sentidos deveriam ser produzidos na interpretação de normas jurídicas. Além disso, (ii) os usos de conceitos dos estudos da linguagem são feitos de forma descuidada, com pouca conexão com como são compreendidos e utilizados em seus campos de origem. Nesse artigo, procuro apresentar uma revisão de literatura de sobre como a semântica e a pragmática podem ser utilizadas coerente e descritivamente para analisar empiricamente os processos reais de produção de sentido na interpretação de normas jurídicas (especialmente constitucionais). Depois de criar um modelo baseado nessa revisão, testo esse modelo em um estudo de caso sobre a Ação Direta de Inconstitucionalidade 4.277, julgada em 2011, pelo Supremo Tribunal Federal, e que reconheceu a constitucionalidade da união estável homoafetiva. Seções de discussão e conclusão, com propostas da relevância deste estudo para a teoria da interpretação jurídica, para a separação de poderes e para os estudos sobre a linguagem da interpretação de normas jurídicas encerram o artigo.

Palavras-chave: *Interpretação constitucional, Jurisdição Constitucional, União homoafetiva, Semântica, Pragmática, Filosofia da linguagem, Linguagem e direito, Linguística computacional.*

Abstract

The interpretation of normative texts is the basic, and perhaps the most important, activity of the jurist. Therefore, doctrinal works on the so-called theory of interpretation are abundant in the legal literature, especially a subfield thereof, which studies the particularities attributed to constitutional interpretation. In search of analytically more powerful concepts, legal doctrines often resort to language studies, in the fields of linguistics and philosophy of language. However, the uses that are made of these concepts are problematic on two levels: (i) they turn such concepts, originally descriptive, that is, about how the language user really makes sense of discourses, into normative concepts, that is, how these meanings should be produced in the interpretation of legal norms. Furthermore, (ii) uses of concepts from language studies are done carelessly, with little connection to how they are understood and used in their fields of origin. In this article, I try to present a literature review of how semantics and pragmatics can be used coherently and descriptively to empirically analyze the real processes of sense making in the interpretation of legal (especially constitutional) rules. After creating a model based on this review, I test this model in a case study on Direct Constitutional Complaint 4,277, a case judged in 2011 by the Brazilian Federal Supreme Court, which recognized the constitutionality of the same-sex civil union. Discussion and conclusion sections, with proposals on the relevance of this study for the theory of legal interpretation, for the separation of powers, and for the study of the language of legal rules construction close the article.

Keywords: *Constitutional interpretation, Constitutional adjudication, Same-sex union, Semantics, Pragmatics, Philosophy of language, Language and law, Computational linguistics.*

1. Introdução

Interpretação de textos normativos é a atividade básica, e talvez mais importante, do/a jurista. Grande parte do treinamento na faculdade de direito envolve compreender, se acostumar com e ver como outro/as juristas compreendem e interpretam textos de lei, cláusulas contratuais e, mais importante, normas da constituição. A própria existência do direito, como fenômeno, depende da evidência e da mediação simbólica do discurso, especialmente escrito, que ao mesmo tempo registra e constitui o fenômeno jurídico. Como Peter Tiersma colocou, mais sucintamente, “nosso direito é um direito de palavras” (Tiersma, 2000, p. 1).

Não é surpreendente, portanto, que as técnicas e conceitos de interpretação jurídica sejam o capítulo inicial de praticamente todo manual ou livro-texto em qualquer ramo do direito. No direito constitucional, a atenção à interpretação das normas constitucionais é objeto de especial cuidado, já que a teoria do direito parte da assunção de que normas constitucionais são ontologicamente distintas de, e mais importantes que, normas infraconstitucionais¹. De tão comuns, referências a como as normas con-

¹“Norma infraconstitucional” é um termo de arte no campo do direito. Refere-se a qualquer norma jurídica que não esteja na Constituição. São, pois, normas infraconstitucionais, leis ordinárias, como o

stitucionais são a “base axiológica”, ou o “fundamento de validade”, ou o parâmetro para a “filtragem” de todas as demais normas jurídicas são até supérfluas (cf., p. ex., J. Afonso da Silva, 2001, pp. 95-96; Britto, 2006, p. 3; Schier, 2005).

Desde pelo menos o entorno da promulgação da Constituição de 1988 vem-se aceitando como um axioma no campo jurídico brasileiro que a interpretação de normas constitucionais reclama um capítulo conceitual à parte, em adição à interpretação de normas em geral. “A ideia de uma nova interpretação constitucional liga-se ao desenvolvimento de algumas fórmulas originais de realização da vontade da Constituição”, dizem Luiz Roberto Barroso e Ana Paula de Barcellos. “[...] [A]s cláusulas constitucionais, por seu conteúdo aberto, principiológico e extremamente dependente da realidade subjacente, não se prestam ao sentido unívoco e objetivo que uma certa tradição exegética lhes pretende dar” (Barroso & Barcellos, 2006, pp. 331-332, grifo no original).

Com toda a relevância que a interpretação jurídica assume para o campo jurídico, é notável o quanto as assunções sobre o funcionamento da linguagem e do discurso embutidas nas teorias e conceitos interpretativos ficam sem problematização. De modo geral, a teoria da interpretação jurídica que parece predominar no ensino e na prática do direito brasileiro é uma soma da teoria dos elementos da interpretação de Savigny, publicada no Século XIX, com novos conceitos e técnicas oriundos de transplantes jurídicos da teoria jurídica alemã, especialmente a desenvolvida pelo comentário das decisões do Tribunal Federal Constitucional Alemão a partir da edição da Lei Fundamental de Bonn (cf. Kommers, 1997, p. 2; Pádua, 2019a, com referências).

O presente artigo pretende contribuir para suprir essa falta de atenção aos conceitos linguísticos e discursivos subjacentes às teorias da interpretação jurídica predominante no Brasil, com foco na interpretação constitucional. Para isso, em primeiro lugar, farei uma breve revisão da literatura jurídica brasileira que procurou estudar conceitos linguísticos, especialmente relativos às áreas da semântica e da pragmática, por vezes sob o guarda-chuva conceitual da semiótica. Como veremos, essa literatura procura incorporar tais conceitos aos conceitos da teoria da interpretação jurídica, ou contrastá-los com esses conceitos. Em seguida, compararei a literatura do campo do direito com a literatura do campo da linguística, ambas focadas nas mesmas áreas da semântica e da pragmática, com ênfase na diferença entre o uso normativo e o uso descritivo dos conceitos. Esses dois momentos de revisão de literatura formarão a base para o estudo de caso, na seção seguinte.

O estudo de caso, a título ilustrativo será sobre o julgamento do Supremo Tribunal Federal da Ação Direita de Inconstitucionalidade (ADI) 4.277. Ao julgar essa ação, o tribunal, por unanimidade, deu interpretação conforme a constituição ao artigo 1.723 do Código Civil, para determinar que nele seja reconhecida também a proteção jurídica à união estável homoafetiva. Utilizando um método qualitativo de análise do discurso, com foco no voto do relator, procurarei demonstrar como a pragmática linguística explica e põe em evidência tensões normativas que ficam escamoteadas pelo recurso aos conceitos da teoria da interpretação constitucional. Especificamente, procurarei demon-

Código Civil (Lei 10.406 de 2001), leis complementares como a Lei de Responsabilidade Fiscal (Lei Complementar 101 de 2000) e decretos, resoluções, portarias e outras normas administrativas. Como a Constituição é vista como o ápice normativo da ordem jurídica, qualquer norma que não esteja nela está “abaixo” dela; daí a alegoria com o campo dêitico cima, baixo.

strar como o relator, falando em nome do tribunal², utiliza um mecanismo pragmático bem documentado na literatura linguística, porém não o apresenta como tal, utilizando, ao invés, conceitos jurídicos menos adequados a descrever a estratégia discursiva de fato utilizada para resolver o caso.

Uma seção de discussão se segue, com interpretações sobre os resultados da análise qualitativa e apresentação de resultados preliminares de uma análise quantitativa sobre os mesmos dados. A discussão se ocupará de como a análise evidencia uma postura criativa do STF no ato de interpretar normas constitucionais, e também de como o tribunal parece dar pouca importância à interpretação literal como potencial limite a essa atividade criativa.

O artigo se encerra com considerações sobre a relevância dos resultados dessa pesquisa para a interpretação jurídico-constitucional e a divisão de tarefas discursivas entre legislativo e judiciário na aplicação de normas a casos concretos; nomeadamente casos que envolvam questões morais sensíveis.

2. Semântica e pragmática na teoria do direito

Embora se apresente como um conjunto de técnicas para produzir sentido de textos (jurídicos), a teoria da interpretação jurídica utilizada no Brasil parece ter se desenvolvido (quase) sem recurso a teorias sobre a produção e negociação de sentido de textos desenvolvidas e testadas no âmbito dos estudos da linguagem.

Em resumo, a teoria da interpretação jurídica tem dois níveis: o da interpretação das normas jurídicas em geral e o da interpretação de normas constitucionais. O primeiro nível é composto pelo uso de quatro técnicas (assumidas como) adequadas para a fixação de sentido dos textos normativos: a interpretação literal ou gramatical, a interpretação histórica, a interpretação sistemática e a interpretação teleológica. Essas quatro técnicas, que seu autor, Friedrich von Savigny chamava de “elementos” (Savigny, 1878, pp. 149–151), podem ser usadas em combinações variadas, ora com ênfase em uma, ora em outra—embora, como veremos, a técnica ou elemento literal dificilmente predomine.

O segundo nível, exclusivo da interpretação constitucional, parte das quatro técnicas de Savigny, mas adiciona técnicas de outro conjunto, cujos elementos são mais numerosos e abertos. Esse segundo conjunto envolve técnicas que derivam de assim chamados “princípios”³, voltados à otimização dos valores constitucionais, tais como “princípio da razoabilidade ou da proporcionalidade”, “princípio da supremacia da constituição”, “princípio da unidade da constituição”, “princípio da máxima efetividade (de

²No direito brasileiro, as decisões de tribunais compostos por mais de um/a juiz/a, são registradas ou materializadas em documentos chamados de “acórdãos”. Esses acórdãos reúnem os votos de cada juiz/a. O/a relator/a é responsável por apresentar um resumo dos fatos do caso (o “relatório”) e um primeiro voto escrito, propondo como o tribunal deve decidir. Como o/a relator/a dá o primeiro voto e como qualquer juiz/a é livre para escrever votos individuais contendo suas razões para votar de uma ou outra maneira, se o voto do/a relator/a for seguido pela maioria do/as juize/as, esse voto é, ao mesmo tempo, o voto individual do/a relator/a e o voto oficial do tribunal como um todo. Uma maneira de organizar isso discursivamente é pensar que o/a relator/a, nesse caso, acumula os papéis de produção de “autor” e “porta-voz” do voto do tribunal, na tipologia que Levinson (2003) deriva da ideia original de Goffman (1979).

³Vide nota de rodapé 26, abaixo.

direitos fundamentais)”, entre outros (cf. V. Afonso da Silva 2005; Pádua 2019a, e referências nesses trabalhos)

Como se pode ver até pelo nome dessas técnicas, quase nenhuma delas busca maneiras de produzir ou negociar sentido de maneira semelhante à que os estudos da linguagem conceituam e estudam o fenômeno da interpretação de textos e discursos. A falta de comunicação entre teorias jurídicas e linguísticas sobre a produção de sentido de textos não é exclusividade do direito brasileiro ou da família romano-germânica (ou da “Civil Law”), senão que parece estar em toda parte (Coulthard, 1997; Felder & Friedemann, 2014; Pádua, 2020b; Shuy, 1993, pp. xvii–xviii).

No Brasil, a primeira monografia clássica sobre interpretação jurídica no Século XX, de Carlos Maximiliano, editada pela primeira vez em 1924, não cita qualquer autor de qualquer das áreas dos estudos da linguagem. Ao tratar do que chamava de “processo [...] gramatical, ou melhor filológico”, Maximiliano cita apenas obras de juristas e usa conceitos intuitivos discutíveis sobre o funcionamento da linguagem para propor que, nesse “processo filológico”, o jurista, “graças ao manejo relativamente perfeito e ao conhecimento integral das leis e usos da linguagem, procura descobrir qual deve ou pode ser o sentido de uma frase, dispositivo ou norma” (Maximiliano, 2002, pp. 87–88)⁴.

De todo modo, esparsamente, a teoria do direito pareceu tentar dialogar com a teoria linguística e a filosofia da linguagem, especialmente a partir do meio do Século XX (Shecaira & Struchiner, 2015). Herbert Hart, por exemplo, em “O conceito de direito”, cuja primeira edição data de 1961, avançou uma concepção sobre interpretação jurídica baseada na “textura aberta da linguagem”, expressamente inspirada em estudos de filosofia da linguagem da época, e que até hoje tem importância nos estudos sobre semântica de textos normativos e interpretação jurídica em geral (Hart, 1961, pp. 120–132; Solan, 1998, pp. 78–83; Struchiner, 2002, 2011; Struchiner & Hannikainen, 2020⁵).

No Brasil, Tércio Sampaio Ferraz Júnior, ainda no final do Séc. XX, reconheceu essa falta de comunicação entre teoria do direito e estudos sobre a linguagem:

Os autores jurídicos, em sua maioria, têm uma visão conservadora da teoria da língua, sustentando, em geral, no que se refere aos objetos jurídicos, a possibilidade de definições reais, isto é, a ideia de que a definição de um termo deve refletir, por palavras, a coisa referida. (Ferraz Jr., 1994, p. 35, grifo no original)

Contra essa concepção que chama de “essencialista”, Ferraz Jr., citando especificamente a filosofia analítica da linguagem, propõe uma concepção que chama de “convencionalista”, na qual “o que deve ser levado em conta é o uso (social ou técnico dos conceitos)” (Ferraz Jr., 1994, p. 35, grifo no original). Para isso, uma divisão de trabalho conceitual entre sintaxe, semântica e pragmática deveria ser seguida:

[...] a análise é sintática [quando] estamos preocupados em definir o uso do termo tendo em vista a relação formal dele com outros vocábulos [...]. [...] [A] análise é semântica, isto é, queremos definir o uso do termo tendo em vista a

⁴Mesmo em 1924, quando da primeira edição da obra de Maximiliano, essa noção de “conhecimento perfeito” ou “integral” das “leis e usos da linguagem” era, na melhor das hipóteses, excessivamente ambicioso, à luz do que já se conhecia e propunha sobre o funcionamento complexo da linguagem.

⁵Um desvio nesse caminho das relações entre direito e estudos da linguagem é o inaugurado pela lógica deôntica, que insere estudos linguísticos formais na análise de estruturas lógicas de proposições jurídicas, que desembocam nas teorias sobre argumentação jurídica (cf. Atienza, 2003; Von Wright, 1999).

relação entre ele o objeto que se comunica [...]. [...] [Quando] definimos o uso do termo tendo em vista a relação do termo por quem e para quem o usa, [...], a análise é pragmática [...]. (Ferraz Jr., 1994, p. 37, grifos no original)

No excerto acima, o autor invoca a clássica distinção de Charles Morris entre sintaxe, semântica e pragmática (cf., p. ex., Levinson, 2007, pp. 1-2). Ele retoma concepções filosófico-linguísticas ao tratar da interpretação jurídica mais adiante, na mesma obra (Ferraz Jr., 1994, pp. 257–260). Trata-se de um esforço vanguardista na teoria jurídica brasileira para incorporar tais concepções na teoria do direito—isto é, concepções relativas às tarefas semânticas e pragmáticas na produção de sentido de textos.

Esse esforço parece ter reverberado em literatura mais recente, que procurou avançar em conceitos específicos da pragmática linguística para explicar eventuais questões e problemas na interpretação de normas jurídicas. Por exemplo, Sgarbi, retomando a distinção de Morris, adiciona conceitos oriundos da teoria dos atos de fala e outros esquemas conceituais pertinentes à linguística e à filosofia da linguagem (Sgarbi, 2007, p. 1)⁶.

Ao tratar especificamente da interpretação de textos normativos jurídicos, Sgarbi, depois de uma ampla revisão de literatura jurídica sobre o tema, recorre aos conceitos de um filósofo-chave para a pragmática linguística contemporânea, Herbert Paul Grice (Sgarbi, 2007, pp. 493–495). A maneira como Sgarbi tenta incorporar a teoria griceana a fenômenos jurídicos, todavia, demonstra a dificuldade de interlocução entre o campo dos estudos da linguagem e o campo jurídico, já que assume, incorretamente, que as máximas de Grice são equivalentes a normas jurídicas—isto é que as máximas de Grice são prescritivas:

[...] note-se que aspectos das máximas de Grice podem ser encontrados delimitando tanto o trabalho do legislador quanto o campo de disputa do significado em juízo. Pense-se aqui na disputa judicial e em aproximação estreita com a máxima [da qualidade] (“não diga o que crê ser falso”), a sanção prevista no art. 342 do Código Penal [...]. (Sgarbi, 2007, p. 495)⁷

Seguindo na mesma linha de Ferraz Jr. e Sgarbi, Bittar, depois de ampla revisão de literatura sobre o que chamou de semiótica jurídica, também lança mão da distinção de Morris sobre as subáreas da análise linguística — e também o faz com alguma dificuldade de interpretação do alcance dos conceitos originais:

Se à Semiótica Jurídica cabe estudar a significação, então, é claro, desse seu espectro de investigação não cabe alhear o triângulo semiótico [...], pois o triân-

⁶Não deixa de ser curioso como Sgarbi, mesmo recenseando literatura sobre diversas teorias e concepções de filosofia analítica da linguagem, tenha ao mesmo tempo afirmado que “o que importa é o entendimento de ser a linguagem tão-somente uma ferramenta de comunicação” (Sgarbi, 2007, p. 24). Essa visão, associada ao modelo da linguagem como conduto, embora ainda útil em alguns modelos, está severamente mitigada na linguística e na filosofia da linguagem, dada sua limitação na descrição das ações sociais associadas ao, e constituídas pelo, uso do discurso em contexto (cf. Martins, 2002). Isso sugere que, de fato, há ainda grande hiato entre as pesquisas linguísticas e seus resultados e as concepções jurídicas sobre linguagem, mesmo quando juristas se esforçam para conhecer os estudos linguísticos.

⁷Como veremos adiante, as máximas de Grice não são normativas no mesmo sentido que uma norma jurídica. A norma, em Grice, vem do campo semântico da normalidade, não da normatividade/obrigatoriedade. A violação de uma máxima conversacional aciona mecanismos não-literais de produção de sentido, nomeadamente o das implicaturas conversacionais. A violação de normas jurídicas aciona mecanismos sancionatórios.

gulo semiótico representa a totalidade do fenômeno do sentido, seja com relação ao sujeito (conceito, interpretante), seja com relação ao referente (psicológico, lógico, físico, metafísico), seja com relação ao significante (simbologia, linguagem, sistema de representação). Semântica, sintática [sic], e pragmática são os três grandes ramos da Semiótica Jurídica [...]. (Bittar, 2008, pp. 52–53)⁸

Finalmente, em texto mais recente, Ferraz Jr. retoma Grice e demonstra maior sutileza conceitual com a operacionalização dos conceitos linguísticos para fenômenos jurídicos—nomeadamente no âmbito da divisão de trabalho entre semântica e pragmática e suas consequências em termos de assunções para a análise dos sentidos produzidos em textos jurídicos:

[...] Grice procura identificar certas interferências que são justificadas a partir de certas máximas de interpretação que instituem o compartilhamento de certos padrões de racionalidade entre os comunicantes. Tais máximas aproximam-se [...] dos postulados de competência que guiam a construção interpretativa da conhecida figura do legislador racional. [...] Isso envolve uma conceptualização do emissor, a partir do compartilhamento de determinados padrões de racionalidade (além das regras semânticas convencionadas na comunidade linguística). [...]. Essa mudança de foco, passando do que foi dito na sentença para o que o agente quis dizer na comunicação é o ponto de partida para a análise pragmática de Grice sobre o que seria uma lógica da conversação. Para Grice, o acesso ao que se quis dizer, a partir do que se disse consiste em um processo de inferência, não dedutiva, que é chamada de implicatura. (Ferraz Jr., 2011, grifos omitidos)

Para Ferraz Jr., o conceito hipotético de legislador racional e as assunções que o intérprete de normas jurídicas faz sobre ele na produção de sentido de normas jurídicas concretas são as bases filosófico-linguísticas da interpretação jurídica. Segundo ele, ao contrário da “lógica interpretativa da conversação ordinária”, a “lógica da interpretação jurídica” não busca apenas identificar significados e inferências sobre materiais linguísticos, mas principalmente, “organizada em torno do valor justiça”, busca “expressar [no processo de interpretação] uma escolha capaz de separar o certo do errado, o justo do injusto [...]” (Ferraz Jr., 2011, p. 105, grifos omitidos).

Essa intuição de Ferraz Jr. sobre a ligação entre a atividade de produção de sentido de textos normativos e avaliações sobre o resultado normativo dessa atividade é refletida em parte da literatura jusfilosófica brasileira, que, para além da divisão de trabalho analítico entre semântica e pragmática, utiliza conceitos linguísticos para delimitar o que é interpretação jurídica do que é criação de normas com base em concepções morais. Essa questão diz com a relevância da interpretação “literal” como delimitação entre o que são argumentos referentes ao sentido específico das normas jurídicas—que

⁸Também em Bittar (2008) há grande desproporção entre, de um lado, a quantidade de informações e referências levantadas na revisão de literatura e na discussão de conceitos teóricos; e, de outro, a tentativa de sua operacionalização empírica, com exemplos ou dados concretos de discursos jurídicos. Bittar dedica 358 páginas (p. 3-361) à revisão de literatura teórica e apenas 16 (p. 367-383) à tentativa de aplicar esses conceitos a exemplos concretos. Mesmo nessa aplicação, ao invés de privilegiar dados de ocorrência real de discursos jurídicos—uma decisão judicial, artigos de normas jurídicas, petições, etc.—Bittar privilegia exemplos hipotéticos de excertos de textos jurídicos, como frases comumente enunciadas por juristas práticos ou textos de normas jurídicas hipotéticas (cf., p. ex., Bittar, 2008, pp. 370 e 373), seguidos de comentários breves e intuitivos que não se conectam com a miríade de conceitos teóricos expostos na parte maior da obra.

são argumentos de autoridade sobre o que o direito “diz” —, por um lado; e argumentos que buscam acesso a razões morais que justificariam a existência dessas normas—que são argumentos substantivos sobre o que o direito “deveria dizer” (Struchiner, 2011; Struchiner & Shecaira, 2012). No contexto dessa delimitação, teorias sobre a interpretação das normas que incorporam conceitos linguísticos não funcionariam como técnicas para descobrir o melhor sentido da norma, mas sim o real sentido da norma (Pádua, 2018, 2019a) — querendo dizer o sentido que os mecanismos linguísticos de produção de sentido utilizados empiricamente pelos usuários da língua preveem como sendo o mais provável, condicionado ao texto normativo em questão⁹.

De todo modo, na maior parte dessa literatura jurídica—assim como em outras obras que poderiam ser mencionadas (cf., p. ex., Carvalho, 2016; Streck, 2006) — o ponto em comum parece ser a utilização de conceitos linguísticos, seja da linguística propriamente dita, seja da filosofia da linguagem, no bojo da teoria da interpretação de normas jurídicas. Esses conceitos se incorporam à lógica normativa própria do campo dogmático-jurídico. Em outras palavras, esses conceitos são incorporados como parte de argumentos dogmáticos de *dever-ser*—argumentos “deônticos”, com o objetivo de propor qual a maneira correta, adequada, ou obrigatória de produzir o sentido de normas jurídicas e de aplicá-lo a casos concretos.

No entanto, no seu âmbito de origem, conceitos linguísticos e linguístico-filosóficos são parte de argumentos ônticos, isto é do campo do *ser*. Seu valor conceitual é predominantemente o de verdade, não o de correção normativa (Habermas, 1984, p. III). Ou seja, eles procuram ser descrições verdadeiras de como o/as usuário/as da língua de fato produzem sentido a partir dos materiais linguísticos com os quais se deparam como parte da sua vida social. Ou, o que é dizer o mesmo, procuram ser teorias descritivas sobre o funcionamento da linguagem.

Também por isso, tais conceitos estão em constante mutação, refletindo as modificações que a sua utilização em situações empíricas impõe às teorias que são testadas nessas situações. Um passo necessário, portanto, é comparar como os conceitos linguísticos e linguístico-filosóficos são utilizados na teoria jurídica — o que foi feito acima — com como eles foram concebidos e são utilizados nos seus contextos acadêmicos de origem—o que será feito a seguir.

3. Semântica, pragmática e a produção de sentido na teoria linguística

As origens das pesquisas sobre semântica e pragmática na linguística são as mesmas que parte da teoria do direito buscou. Nomes como o do já citado Charles Morris e de Charles Sanders Peirce são seminais no tipo de discussão ainda hoje vigente na teoria linguística sobre a divisão de trabalho entre semântica e pragmática, como dois subcampos que lidam com a atividade de produção e identificação de sentidos linguísticos (Levinson, 2007, p. 1; Oliveira & Basso, 2007). No entanto, a evolução da teorização e da aplicação empírica desses conceitos em linguística vem mostrando que o critério

⁹Na literatura internacional, há debate semelhante, incorporando igualmente conceitos linguísticos, lógicos e dogmático-jurídicos (p. ex., Eskridge & Levi, 1995; Schauer, 1998; Solan & Gales, 2016). Sobre a ideia de que o significado derivado de um texto é um cálculo de probabilidade condicional, vide nota de rodapé 11, abaixo.

derivado das relações entre signo, referente e usuário da língua, tão popular entre juristas, é reducionista, para não dizer equivocado (Oliveira & Basso, 2007, pp. 5–9).

Em primeiro lugar, parece estar claro que a atividade de produção e negociação de sentido em textos não é bem modelada por entidades estanques, que encapsulem aspectos separados dessa atividade. Qualquer modelo realista dessa atividade precisa incorporar relações entre o/as usuário/as da língua junto com níveis de significação linguística. Aliás, já desde o final do Século XIX, com a teoria do significado de Frege, havia na filosofia da linguagem a intuição de que a significação da língua não se esgotava numa relação entre “signo” e “referente”, mas num complexo sistema de referenciação que envolvia níveis de significação: um interno ao próprio sistema linguístico e um, que, partindo daí, referenciava o mundo “externo” à língua¹⁰.

Esse sistema de significação complexo é o domínio inicial da semântica, campo dos estudos da linguagem que se ocupa, predominantemente, do significado hoje chamado de literal, ou convencional ou de superfície (Oliveira & Basso, 2007; Pádua, 2018, 2020c; Slocum, 2017; Terkourafi, 2021). O que a pragmática demonstrou ao longo do Século XX, no entanto, é que o estudo desse significado de superfície, embora necessário, é insuficiente para dar conta do complexo fenômeno de produção e negociação de sentidos da linguagem humana. Em múltiplas situações comunicativas, os usuários da língua precisam adicionar elementos de sentido ao significado de superfície para que mesmo frases simples possam atingir objetivos interacionais e comunicativos (Geis, 1995; Huang, 2017; Pádua, 2020c; Recanati, 2001; Stein, 2017). Vejam-se os seguintes exemplos:

1. Voltei do mercado — pressupõe que eu fui ao mercado;
2. Vou ao jogo domingo — implica que é o próximo domingo, não outro qualquer

A constatação da insuficiência do significado de superfície levou Grice a propor que a significação linguística se dá em dois níveis distintos e hierarquicamente estruturados: o nível do “que é dito” (*what is said*) e o nível das implicaturas (*implicatures*):

No sentido em que estou usando a palavra dizer [*say*], pretendo que o que alguém disse seja relacionado intimamente com o sentido convencional das palavras (da frase) que ele/a enunciou. [...] Em alguns casos, o significado convencional das palavras usadas vai determinar o que é implicado, além de ajudar a determinar o que é dito. (Grice, 1975, p. 44, grifo no original)

Em ambos os níveis de análise, como corretamente intuído por Ferraz Jr., acima, objetivo é estabelecer a intenção comunicativa evidenciada pelo produtor do sentido, definida pelo conceito griceano de “significado_{nn}” [*meaning_{nn}*]:

[P]ara que x tenha significado_{nn} qualquer coisa, não apenas [essa coisa] deve apenas ter sido “enunciada” com a intenção de induzir uma certa crença, mas também o enunciador precisa ter tido a intenção de que uma “audiência” reconhecesse a intenção por trás do enunciado (Grice, 1957, p. 382)

¹⁰Frege desenvolveu uma interessante prova lógica sobre a existência de um significado que é interno à língua e outro, logicamente distinto, que liga o primeiro a entidades do mundo exterior. Essa prova ficou conhecida como a teoria do “senso” e da “referência”, tradução dos conceitos fregeanos no original em Alemão, “*Sinn*” e “*Bedeutung*” (Frege, 2007).

Parte desse desdobramento em aspectos internos e externos do significado é também capturada pela distinção na filosofia analítica da linguagem entre “intensão” e “extensão” do significado (Auroux, 2018, p. III).

No modelo de produção de sentido griceano, portanto, temos três variáveis que geram os resultados de sentido: o significado convencional das palavras contidas no texto, o significado implicado pelo uso dessas palavras nesse texto e a intenção comunicativa do/a enunciador, que é (re)construível pela audiência como o resultado de uma função envolvendo esses dois tipos de significado (Geis, 1995; Goodman & Frank, 2016, Huang, 2017, pp. 47–49; Terkourafi, 2021, pp. 77–79).

Com a evolução do modelo griceano, ocasionada por tensionamentos teóricos e empíricos às proposições originais do modelo, perspectivas neogriceanas foram surgindo e aperfeiçoando a linguística teórica e a linguística aplicada, especialmente no aspecto pragmático (Huang 2017). Um dos modelos neogriceanos que parece funcionar melhor com discursos jurídicos é o proposto por François Recanati (Recanati, 2001, 2002; Slocum, 2017; Stein, 2017).

No modelo de Recanati, há três níveis de significado: o nível da superfície, o nível do que é dito e o nível do que é implicado, como pode ser graficamente esquematizado a seguir:

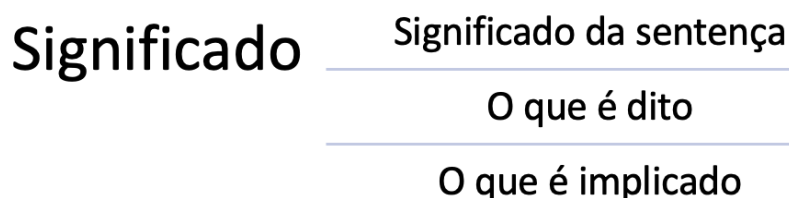


Figure 1. Modelo de níveis de significado (Adaptado de Pádua (2020c, p. 70))

Para Recanati, o critério que faz a análise (ou compreensão) passar de um nível para o outro é o custo cognitivo de processamento do significado. No nível da superfície, o significado é imediatamente conectado à forma; deriva diretamente do sentido literal das palavras e expressões (ou sintagmas) e sua organização sintática. No nível do que é dito, processos de expansão, redução ou alteração do significado literal são acionados, mas ainda num nível subconsciente e com custo cognitivo baixo—como nos exemplos (1) e (2), acima. Já no nível do que é implicado, há uma maior transformação do sentido do texto, acionado pelas palavras e sintagmas da superfície, mas apenas indiretamente relacionados com elas, como no caso dos atos de fala indiretos por exemplo¹¹.

Embora o modelo de Recanati assuma fatos empíricos sobre a cognição que não parecem ser corretos em todos os casos (Rühlemann, 2019, pp. 25-31), e não consiga incorporar o aspecto intersubjetivo da comunicação humana (Martins, 2002; Terkourafi, 2021), a intuição básica do modelo tem se demonstrado correta. Isto é, há uma divisão de trabalho entre diferentes níveis de significado, ela é fluida nos limites entre os níveis e sua função comunicativa é possibilitar, ao mesmo tempo, a criatividade necessária à produção de sentido e uma certeza mínima sobre o significado que permita uma comunicação estável, a partir da atribuição de alguma intenção comunicativa a/o falante.

Essa perspectiva neogriceana, com mais ou menos modificações e maior ou menor formalização, é a principal perspectiva teórica que orienta os estudos contemporâneos

¹¹Exemplos clássicos de atos de fala indiretos em Português incluem sentenças do tipo “Está quente aqui!”, que, embora na superfície pareça uma afirmação sobre o clima, indiretamente implica um pedido, que uma janela seja aberta ou um ar-condicionado ligado.

em linguística teórica e aplicada sobre como os usuários da língua produzem, negociam e atribuem significado a materiais linguísticos (Goodman & Frank, 2016; Huang, 2017; Terkourafi, 2021, entre outros).

Assim como qualquer modelo ou teoria linguística, a perspectiva neogriceana não prescreve como o/a usuário/a da língua deve produzir, negociar e identificar significado, mas como ele/a de fato o faz. Trata-se, portanto, de um conjunto de modelos descritivos, cuja validade empírica depende da demonstração da sua aplicabilidade para explicar ou prever fenômenos reais (Epstein & King, 2002; Pádua, 2020a).

Se um modelo neogriceano do tipo apresentado acima estiver correto, fenômenos jurídicos empíricos que dependem de produção de sentido poderão ser explicados por uma aplicação desse modelo. O estudo de caso a seguir procura fazer um teste dessa aplicabilidade.

4. A ADI 4.277 como caso de estudo

4.1. Dados e métodos

Usarei como fontes primárias para o estudo de caso apenas a íntegra do acórdão do Supremo Tribunal Federal na Ação Direita de Inconstitucionalidade (ADI) 4.277, o texto da Constituição Federal brasileira de 1988 e o texto do artigo 1.723 do Código Civil, que é a norma impugnada nessa ADI.

No Direito Constitucional brasileiro, a ação direta da inconstitucionalidade é um dos instrumentos que levam ao chamado “controle concentrado de constitucionalidade”. Esse controle é exercido direta e originariamente pelo Supremo Tribunal Federal, no seu papel de tribunal constitucional, isto é, de avaliador da constitucionalidade das leis e demais normas infraconstitucionais editadas pelo Poder Público. Um conjunto limitado de atores, listados no artigo 103 da Constituição, pode propor uma ADI perante o STF, a fim de iniciar um procedimento de controle de constitucionalidade. Dois desses atores propuseram essa ação, neste caso: o Governador do Estado do Rio de Janeiro e o Procurador-Geral da República.

A ADI 4.277 foi escolhida como estudo de caso intencionalmente, por se mostrar um bom exemplo de como mecanismos discursivos de produção de sentido explicáveis e condizentes com o modelo linguístico neogriceano operam na tomada de decisões judiciais em casos complexos—mesmo quando esses mesmos mecanismos não são conscientemente percebidos pelo tribunal (vide Discussão abaixo).

O inteiro teor do acórdão foi coletado do site do tribunal¹² como um arquivo em formato pdf de 270 páginas. Convertido este arquivo para formato em texto bruto (.txt), e feita uma limpeza parcial, para retirar cabeçalhos, referências ao número do processo e número de páginas, além de marcas de assinatura digital, pontuação e outras não-palavras, o resultado é um texto de quase 75 mil palavras¹³.

Sobre esses dados, foram feitas duas análises. A principal delas, de natureza qualitativa, consistiu na análise do discurso do relatório e voto do Ministro Relator, Carlos

¹²<https://jurisprudencia.stf.jus.br/pages/search/sjur200017/false>, acesso em 13 fev. 2024.

¹³74.957 palavras, para ser mais exato. Os dados e documentos brutos, bem como o código utilizado para as análises feitas para este artigo, em linguagem de programação Python, estão publicamente disponíveis em https://github.com/joaoppadua/adi_4277.

Britto, com a finalidade de identificar mecanismos e padrões linguísticos de produção e negociação de sentido das normas jurídicas em questão (cf., para métodos em análise qualitativa do discurso, Arribas-Ayllon, Sarangi, & Clarke, 2011; Pádua, 2019b; Roberts & Sarangi, 2005; Sarangi, 2005). Uma leitura mais superficial e sem as mesmas pretensões analíticas também foi feita sobre os demais votos.

Após essa análise do discurso, cujas conclusões serão apresentadas na seção seguinte, foi feita uma nova análise complementar. Essa análise foi quantitativa e assistida por computador, usando principalmente técnicas do campo da linguística de *corpus* e do processamento de linguagem natural (Egbert, Larsson, & Biber, 2020; Freitas, Martins, & Biar, 2022; Hovy, 2022; Jurafsky, 2015; Rühlemann, 2019; Sardinha, 2000). Nessa segunda análise, motivada pelos resultados da primeira, o objetivo era testar o quanto o STF, nesta decisão, levou em consideração conceitos tipicamente levantados pela doutrina jurídica e pela própria jurisprudência do STF como limites à atividade criativa na interpretação constitucional e, notadamente, na técnica de interpretação conforme a constituição—tais como a “literalidade” ou os “limites semânticos” dos textos normativos em consideração (cf. voto do Min. Gilmar Mendes, p. 753-68).

Essa segunda análise será explorada na seção de discussão deste artigo, já que tem uma natureza principalmente auxiliar e exploratória, em relação à análise qualitativa e seus resultados.

4.2. Contexto e questão jurídica na ADI 4.277

A ADI 4.277 foi ajuizada pelo Procurador-Geral da República. Ela foi julgada em conjunto com a Arguição de Descumprimento de Preceito Fundamental (ADPF) 132¹⁴, ajuizada pelo Governador do Estado do Rio de Janeiro.

A ADPF 132 foi, na verdade, proposta antes, ainda no ano de 2008, no contexto de iniciativa política do então Governador para promoção dos direitos da comunidade LGBT+. Originalmente, o pedido nesta ação era o reconhecimento das uniões homoafetivas para efeito de proteção de direitos previdenciários de servidores públicos do Estado do Rio de Janeiro. Dois anos depois, em 2010, a Procuradoria-Geral da República ajuizou a ADI 4277, com escopo mais amplo.

No que tinham de comum, ambas as ações pediam que fosse declarada a inconstitucionalidade parcial com interpretação conforme a constituição do artigo 1.723 do Código Civil, que trata da união estável, a fim de interpretar que o mesmo dispositivo também abarca as uniões estáveis homoafetivas.

Ambas as ações também fundamentavam seus pedidos em argumentos calcados no direito constitucional à igualdade e à dignidade da pessoa humana. Em conjunto com outros princípios constitucionais, tais como a razoabilidade e proporcionalidade e

¹⁴No direito constitucional brasileiro, adotou-se uma técnica de controle de constitucionalidade das leis chamada “interpretação conforme a constituição”. A técnica é oriunda do direito alemão e utilizada com crescente frequência pelo Tribunal Federal Constitucional Alemão (Afonso Da Silva, 2006; Vosskuhle, 2000), com o objetivo de, sem declarar que uma lei é totalmente inconstitucional, dizer que, para ser constitucional, essa mesma lei deve ser interpretada de uma determinada maneira. No caso em estudo, a maneira como a norma do Código Civil deveria ser interpretada para ser constitucional seria a de que essa norma abarcaria, também, a união estável homoafetiva. Como explicado no texto, a dificuldade de ordem linguística com a técnica da interpretação conforme à constituição, nesse caso, decorre do fato de que a própria Constituição Federal tem uma norma com texto idêntico, no essencial, ao do Código Civil.

à vedação a qualquer forma de discriminação, tais parâmetros normativos resultariam em que a leitura do artigo 1723 do Código Civil que limitasse a união estável a casais heterossexuais seria inconstitucional¹⁵.

A votação foi unânime pela procedência do pedido, porém todo/as o/as ministro/as em atuação na época fizeram declarações de voto separadas. O julgamento parece ter sido bem recebido na comunidade jurídica, tanto que foi logo usado como precedente pelo Conselho Nacional de Justiça para estender a possibilidade de união estável para casamentos homoafetivos (v. Resolução nº 175/2013).

Uma peculiaridade importante na causa de pedir dessa ADI é que, embora os parâmetros de controle invocados pelo autor da ação tenham sido os princípios da igualdade, liberdade, dignidade da pessoa humana, segurança jurídica e razoabilidade ou proporcionalidade (voto do Relator, p. 618-619), há na Constituição de 1988 um texto superficialmente quase idêntico ao do artigo 1723 do Código Civil. O quadro a seguir coloca os textos lado a lado, com as palavras idênticas grifadas.

Artigo 226, § 3º da Constituição Federal	Artigo 1723 do Código Civil (Lei. 10.406/2002)
Para efeito da proteção do Estado, é reconhecida a união estável entre o homem e a mulher como entidade familiar , devendo a lei facilitar sua conversão em casamento	É reconhecida como entidade familiar a união estável entre o homem e a mulher , configurada na convivência pública, contínua e duradoura e estabelecida com o objetivo de constituição de família

Figure 2. Comparativo de textos entre Constituição e Código Civil. Fonte: Autor.

Essa circunstância, e a maneira como o STF—principalmente no voto-condutor do Relator—lidou com ela foi a principal unidade de análise deste estudo de caso. Afinal, como mencionado acima, os “limites semânticos” dos textos normativos são pacificamente reconhecidos como limitadores à interpretação conforme à constituição—e às demais técnicas de interpretação constitucional como um todo (cf., p. ex., Barroso & Barcellos, 2006, p. 361; Pádua, 2019a, com referências). Por isso, entender o mecanismo discursivo utilizado para negociar esses limites semânticos é importante para entender, empiricamente, quais são os limites a esses limites.

4.3. Modelo analítico

Os conceitos analíticos utilizados em análise do discurso dependem de que fenômenos se fazem relevantes a partir dos dados (Roberts & Sarangi, 2005). Neste caso, a principal questão que surgiu da análise qualitativa dos dados foi a estratégia discursiva utilizada pelo STF para justificar um controle de constitucionalidade do texto do Código Civil, quando a Constituição contém norma de redação quase igual à da norma impugnada. Como veremos, na seção seguinte, essa estratégia centra-se no cancelamento da

¹⁵Esse resumo do caso e dos argumentos dos autores é baseado no relatório da ADI 4277 (Voto do Relator, p. 616-624).

implicatura gerada pela forma como a norma do artigo 226, § 3º, da Constituição foi textualizada superficialmente.

Para ver isso, antes precisamos voltar à questão da semântica e da pragmática. Como vimos, um modelo neogriceano sobre o significado implica que a produção do sentido de um texto assume que há uma intenção comunicativa projetada pelo/a falante e reconstruída pelo/a ouvinte, que reconhece essa intenção¹⁶. Uma elaboração recente de um modelo desse tipo, a postula que o significado de um discurso, tal como entendido pelo ouvinte, varia em função das probabilidades relativas à intenção comunicativa atribuída ao falante, condicionada à superfície textual e ao contexto, este definido como o conhecimento de mundo compartilhado entre falante e ouvinte.

No modelo que estamos usando para esta análise, a superfície do texto, que se liga à dimensão semântica da produção de significado, é suficiente para derivar o nível do que é dito. No entanto, para dar conta do que é implicado, a superfície textual é insuficiente. Originalmente, para mapear de um para o outro nível, Grice sugeriu outra assunção fundamental da comunicação humana—além da intenção comunicativa. Essa segunda assunção, compartilhada por qualquer conjunto de dois ou mais interlocutores, poderia ser conceituada como um princípio da cooperação:

3. PRINCÍPIO DA COOPERAÇÃO DE GRICE¹⁷

Faça sua contribuição conversacional tal como ela for requerida, no estágio em que ela ocorrer e de acordo com o propósito ou direção da troca verbal na qual você está participando

Desse princípio, surgiriam quatro categorias de máximas da conversação — Quantidade, Qualidade, Relação e Modo [*Manner*], que seriam formuladas também como normas. As que nos interessam para este artigo são as máximas da quantidade e do modo:

4. MÁXIMAS CONVERSACIONAIS DE GRICE

Categoria da quantidade: [...] Têm-se [...] as seguintes [sub-]máximas:

a) Faça sua contribuição tão informativa quanto necessária (aos propósitos do

¹⁶Há na literatura linguística contemporânea uma ampla discussão conceitual e empírica sobre o papel da intenção comunicativa como parte da análise da produção de sentido em situações de uso real de linguagem. Não posso entrar nessa discussão aqui. Como está no texto, vou assumir que a intenção comunicativa de tipo griceano é um conceito útil e que enquadra bem os dados utilizados neste artigo—ciente de que essa assunção não é livre de problemas teóricos e empíricos noutras situações e com outros dados (cf. Haugh, 2008; Terkourafi, 2021). A teoria referida tem o nome de “Teoria do Ato de Fala Racional” [*Rational Speech Act Theory*] (Frank & Goodman, 2012; Goodman & Frank, 2016), em que o significado obedece a uma função de probabilidade bayesiana expressa pela fórmula:

$P_L(w | u) \propto P_s(u | w)P(w)$ onde $P_L(w | u)$ é a probabilidade atribuída pelo ouvinte de que o significado ‘w’ seja condicionado pelo texto ‘u’, $P_s(u | w)$ é a probabilidade de que o falante tenha a intenção de comunicar ‘w’ em virtude de ter enunciado ‘u’ e $P(w)$ é a probabilidade a priori de que o estado de coisas ‘w’ seja comunicado por qualquer pessoa.

Para o que nos interessa neste artigo, esse modelo formaliza as intuições, presentes nas perspectivas neogriceanas, de que o significado produzido de um texto varia em função da atribuição de intenção comunicativa de um interlocutor ao outro, que essa atribuição é proporcional à superfície do texto e ao conhecimento de mundo dos interlocutores e, crucialmente, que tudo isso é “calculado” pelos interlocutores em termos probabilísticos, e, portanto, incertos e abertos a negociações, edições, cancelamentos e outras modificações.

¹⁷Nesse e no próximo excerto utilizo a tradução feita por Hagemeyer Burgo and Araújo (2018), por já estar publicada no Brasil e ser de boa qualidade. A fonte original é Grice (1975, pp. 45–47).

intercâmbio em questão);

b) Não faça sua contribuição mais informativa que o necessário. *Categoria da relação*: há uma única máxima: seja relevante [...].

No modelo griceano, como dito, essas máximas seriam normas que captam a assunção de racionalidade que cada interlocutor atribui ao outro. A violação dessas máximas, quando ocorrida, implicaria que o significado de um texto não se esgotava na sua superfície, mas precisaria de processos de significação derivados das implicaturas conversacionais derivadas da violação ostensiva das máximas.

Em versões neogriceanas, o potencial significativo das máximas da conversação deixou de depender de violações ostensivas a elas. As implicaturas conversacionais continuam a ser condicionadas à assunção de que as máximas estão sendo observadas (cf. Grice, 1975, pp. 49-50), mas podem ser evidenciadas mesmo quando as máximas estão sendo respeitadas, por mecanismos normais de produção de sentido, que são reconhecidos por interlocutores como maneiras de evidenciar a intenção comunicativa uns dos outros (Horn, 1995; Huang, 2017). Por isso, ganham proeminência as máximas da quantidade e da relação, citadas acima em (4), que se estabelecem como limites superiores e inferiores à quantidade de discurso necessária para estabelecer uma comunicação bem-sucedida (Horn, 1995)¹⁸.

Essa releitura das máximas conversacionais como limites numa escala quantitativa de significado permite que a dimensão pragmática se torne mais claramente relevante também para a análise de textos escritos, tais como normas legais e constitucionais, para as quais a falta de dados sobre o contexto interacional torna a tarefa de derivar conclusões sobre o significado mais complexas (Allott & Shaer, 2017). Normalmente, com textos normativos, a única evidência sobre o significado, e a possibilidade de projeção de intenção comunicativa é a superfície textual, especialmente porque textos normativos são concebidos como autônomos; devem poder ser compreendidos e se aplicar a situações muito separadas no tempo e no espaço da sua concepção inicial (Tiersma, 2001).

Uma última característica do modelo neogriceano que será importante para a análise dos dados é que as inferências do nível do que é implicado são canceláveis pela adição de materiais linguísticos na própria superfície textual (Levinson, 2007, pp. 141–144). Aliás a cancelabilidade vem sendo proposta como a principal característica distintiva dos sentidos pragmáticos (Oliveira & Basso, 2007).

¹⁸O modelo griceano clássico, por exemplo, explica bem por que alguém que diga, “está quente aqui” numa sala de reunião vai ser interpretado como implicando um pedido de ligar o ar-condicionado ou abrir a janela (vide nota de rodapé 8, acima), já que falar a esmo um comentário sobre a temperatura numa reunião viola a máxima de relação e possivelmente a máxima da quantidade, submáxima a. Porém, ela tem problemas para explicar por que o tradicional exemplo da norma hipotética, “nenhum veículo é permitido no parque” é tida como implicando uma cláusula de razoabilidade ou autorização do tipo “nenhum veículo *não-autorizado* é permitido no parque”. Em modelos neo-griceanos, essa cláusula de razoabilidade é explicada pelo funcionamento normal do princípio R, pelo qual um falante deve dizer apenas o necessário para se fazer entender e o ouvinte vai recriar implicaturas consistentes com a intenção racional desse falante, condicionado ao contexto de emissão e recepção do discurso (Horn, 1995). Vide a nota ix, acima, para uma formalização dessa ideia, na Teoria do Ato de Fala Racional. E vide seção “Discussão”, abaixo, para mais detalhes sobre o exemplo de Horn.

Cancelabilidade quer dizer que o autor do texto pode expressamente anular uma implicatura ou outra inferência implicada que ele/a sabe que derivará do próprio texto e que não faz parte da sua intenção comunicativa. Um exemplo é dado abaixo, em (5):

5. Está quente aqui!—mas não precisa ligar o ar-condicionado.¹⁹

A implicatura de pedir que alguém ligue o ar-condicionado pode ser cancelada, sem problemas. Mas significados de nível semântico, não. Se tentamos cancelar aspectos ou acarretamentos semânticos, a frase fica sem sentido ou contraditória:

6. #Está quente aqui!, mas não está calor.²⁰

Juntando tudo isso, temos então, um modelo de produção de significado que pode ser esquematizado da seguinte forma:

7. MODELO NEOGRICEANO DE PRODUÇÃO DE SENTIDO²¹

SIGNIFICADO/SENTIDO (x) \iff (*Intenção comunicativa projetada pelo discurso*) (x)

Sentido(x) = *significado semântico* (x) + *significado pragmático* (x)

significado semântico (x) \approx *significado de superfície* ([palavra₁], ..., palavra_n])

significado pragmático (x) \iff *princípio da cooperação*

princípio da cooperação \implies *máxima da quantidade, máxima da relação*

\forall (*significado pragmático*(x))_x[*Cancelável*(*significado pragmático*(x))]

Fixados esses elementos do modelo, não é necessário que cheguemos a pormenores sobre teorias e conceitos sobre produção e negociação de sentido discursivo. O que foi exposto até aqui é suficiente para explicar os resultados da análise, na seção seguinte. Interessantemente, uma das intuições fundamentais desse modelo é capturado por um brocardo comumente invocado nas tarefas interpretativas no campo do direito: “a lei não contém expressões inúteis”.

5. Resultados

Analisando a questão interpretativa na ADI 4.277 à luz do modelo acima, os níveis semânticos da produção de sentido não são problemáticos, nem geraram necessidade de maior justificação da parte do voto do Ministro Relator ou dos demais votos. Nem

¹⁹Na literatura linguística, também foi proposto um teste para diferenciar implicaturas de acarretamentos lógicos do significado semântico que se vale do processo oposto ao cancelamento, nomeadamente, o processo de *reforço* (ou possibilidade de reforço, do Inglês *reinforceability*). Implicaturas podem ser repetidas sem gerar redundâncias (McCawley, 1993, p. 324, citando Sadock).

Por exemplo, “está quente aqui; você pode ligar o ar-condicionado?” é um enunciado normal em Português, mas “está quente aqui; faz calor” é redundante a ponto de ser estranho—inclusive, essa redundância pode, ela mesma, acionar uma implicatura conversacional.

²⁰Uso a convenção comum em análises linguísticas, com o símbolo “#” marcar uma frase-exemplo com anomalias semânticas ou lógico-semânticas.

²¹Apenas para efeito de formalização, o modelo utiliza símbolos da lógica matemática ou lógica formal. Um glossário desses termos pode ser encontrado em https://pt.wikipedia.org/wiki/Lista_de_s%C3%ADmbolos_l%C3%B3gicos, acesso em: 18 abr. 2022. Uma fonte indispensável para a aplicação de modelos lógicos em linguística é McCawley (1993), de onde, em geral, a notação utilizada no modelo é oriunda.

os textos invocados como parâmetros de controle, nem o texto da norma impugnada, o artigo 1.723 do Código Civil, apresentam problemas de vagueza, ambiguidade ou indeterminação semântica (cf., sobre problemas semânticos na interpretação jurídica, Solan, 1998, pp. 78–86; Solan, 2018).

O ponto relevante está no nível pragmático de produção de sentido. Como visto, o texto do artigo 1.723 do Código Civil é quase igual ao do artigo 226, § 3º, da Constituição. Crucialmente, ambos contém o sintagma “entre o homem e a mulher” como qualificador de “união estável”. A colocação desse sintagma no texto do artigo constitucional suscita, pragmaticamente, qual é a relevância do significado dele para o sentido do texto. Isso é comumente verdade com qualquer sintagma modificador (Jurafsky & Martin, 2019, pp. 12-20; Perini, 2006, pp. 340-342). Considere o seguinte exemplo:

8. “É reconhecido o parentesco entre irmãos para efeito de herança.”

A característica inserida pelo sintagma modificador “entre irmãos” é essencial para o sintagma modificado “o parentesco”. Um/a usuário/a da Língua Portuguesa entende que a inserção desse sintagma exclui da intenção comunicativa projetada por essa sentença o parentesco entre outros tipos de parentes—primos, por exemplo—, mesmo que qualquer relação entre parentes esteja dentro do conceito semântico de “parentesco”.

O exemplo (8) demonstra que, normalmente, a inserção de sintagmas modificadores tem o efeito semântico de adicionar circunstâncias ao sintagma modificado e o efeito pragmático de *excluir* outras possibilidades, a partir, principalmente, do mecanismo das máximas griceanas da quantidade e da relação.

Aplicando isso ao artigo 226, §3º, temos que a menção a “entre o homem e a mulher” implica que estão excluídas do escopo de significado da norma, condicionado à intenção comunicativa projetada pelo texto, outras possibilidades. Além do exemplo acima, uma outra maneira de ver isso é notar que o mesmo texto pode ser escrito sem o modificador, preservando seu significado semântico, como abaixo:

9. Para efeito da proteção do Estado, é reconhecida a união estável como entidade familiar, devendo a lei facilitar sua conversão em casamento.

Não só o texto continua a fazer sentido, como parece claro, nessa versão, que não há mais implicaturas que excluem outros tipos de união estável.

O Relator da ADI 4.277 abordou a questão do escopo do artigo 226, §3º, de forma breve. No seu voto de 31 páginas, ele faz um percurso argumentativo que começa com considerações sobre a sexualidade e a homoafetividade como parte dessa sexualidade (ADI 4.277, p. 629-631), passa pela construção do que chama de um “constitucionalismo fraternal” e da vedação ao preconceito (p. 632-634), pela afirmação de que a Constituição “opera por um intencional silêncio” sobre a prescrição de formas de sexo e de orientação sexual (p. 634-639), e chega num resumo parcial de todos esses argumentos (p. 640-643). Em seguida, se dedica a interpretar especificamente o artigo 226, dando ênfase à construção de conceitos de família e casamento (p. 643-650).

Somente na página 651, o Relator enfoca o artigo 226, § 3º, e seu texto. Nesse tema, faz dois argumentos iniciais. O primeiro, de que

essa referência à dualidade básica homem/mulher tem uma lógica inicial: dar imediata sequência àquela vertente constitucional de incentivo ao casamento como forma de reverência à tradição sócio-cultural-religiosa do mundo ocidental

de que o Brasil faz parte (§1º do art. 226 da CF), sabido que o casamento civil brasileiro tem sido protagonizado por pessoas de sexos diferentes, até hoje. (ADI 4.277, p. 652)

E o segundo, de que

a normação desse novo tipo de união, agora expressamente referida à dualidade do homem e da mulher, também se deve ao propósito constitucional de não perder a menor oportunidade de estabelecer relações jurídicas horizontais ou sem hierarquia entre as duas tipologias do gênero humano, sabido que a mulher que se une ao homem em regime de companheirismo ou sem papel passado ainda é vítima de comentários desairosos de sua honra objetiva [...]. (ADI 4.277, p. 652, grifos no original)

Esses dois argumentos iniciais demonstram que o modelo de sentido construído acima é reconhecido intuitivamente pelo Relator como usuário da língua. Ele demonstra que há a necessidade de negociar algum sentido para o modificador “entre o homem e a mulher”; não é possível apenas ignorá-lo. Para dar conta dessa necessidade, o Relator realiza trabalho retórico de fôlego: dá duas explicações não relacionadas entre si e não derivadas da superfície do texto do artigo 226, § 3º, mas construídas como “propósitos” ou “lógicas” extralinguísticas para sua presença no texto.

O Relator ainda demonstrou sentir a necessidade de expressamente afastar outras explicações ou implicaturas possíveis desse modificador:

[E]sse combate mais eficaz ao preconceito que teimosamente persiste para inferiorizar a mulher perante o homem é uma espécie de briga particular ou bandeira de luta que a nossa Constituição desfralda [...], nada tendo a ver com a dicotomia da heteroafetividade e da homoafetividade. [...] [...] [T]anto numa quanto noutra modalidade de legítima constituição da família [isto é, casamento civil e união estável], nenhuma referência é feita à interdição, ou à possibilidade, de protagonização por pessoas do mesmo sexo.” (ADI 4.277, P. 652-654, grifos no original).

Há um somatório de argumentos que parecem redundantes²², já que qualquer um deles, se correto, já seria suficiente para validar a posição do Relator. De um lado, a necessidade de dar duas explicações para a presença do modificador “entre o homem e a mulher”, do outro, a afirmação de que o texto da norma não faz “nenhuma referência à interdição” entre a união homoafetiva²³.

²²Essa redundância, por si só, também pode acionar uma implicatura em relação à máxima da quantidade, submáxima b (“não diga mais do que o necessário”). Vide (4), acima, no texto.

²³No voto do Min. Gilmar Mendes, por exemplo, há apenas a utilização do último desses argumentos: “[C]heguei até a comentar com o Ministro Relator Ayres Britto, tendo em vista, como amplamente confirmado, que o texto do Código Civil reproduz, em linhas básicas, aquilo que consta do texto constitucional. [...] [O] texto, em princípio, reproduzindo a Constituição, não comportaria esse modelo de interpretação conforme. Ele não se destinava a disciplinar outra instituição que não fosse a união estável entre o homem e a mulher, na linha do que estava no texto constitucional. Daí não ter polissemia, daí não ter outro entendimento que não aquele constante do texto constitucional. Talvez o único argumento que pudesse justificar a tese da aplicação ao caso da técnica da interpretação conforme à Constituição seria a invocação daquela previsão normativa de união estável entre homem e mulher como óbice ao reconhecimento da união entre pessoas do mesmo sexo, como uma proibição decorrente daquele dispositivo.

E, de fato, é com base nesse argumento que entendo pertinente o pleito trazido nas ações diretas de inconstitucionalidade”. (ADI 4.277, p. 765-5, voto Min. Gilmar Mender, grifos omitidos).

O efeito desse somatório de argumentos, especialmente o último deles, é um efetivo cancelamento da implicatura sugerida pelo modificador “entre o homem e a mulher”, e explicada pela interação entre as máximas da relevância e da quantidade. Nomeadamente, a implicatura gerada pela inferência de que a presença do modificador não pode ser supérflua; se está lá, ele implica a exclusão das possibilidades alternativas.

Como qualquer cancelamento de implicaturas, também esse pode ser formulado como uma adição expressa ao texto, marcada abaixo em itálico:

10. Para efeito da proteção do Estado, é reconhecida a união estável entre o homem e a mulher como entidade familiar, devendo a lei facilitar sua conversão em casamento, mas não apenas entre o homem e a mulher.

Portanto, o que o STF fez, ao julgar procedente a ADI 4.277, do ponto de vista discursivo, foi efetivamente um cancelamento da implicatura derivada do que foi dito pela superfície textual do artigo 226, § 3º, da Constituição, estendendo esse cancelamento para a produção de sentido do artigo 1.723 do Código Civil.

Esse cancelamento não se deu de forma explícita, nem foi formulado como tal. Foi formulado em termos deextralinguísticos. À luz desses argumentos, na linha do voto do Relator, somados ao argumento linguístico de que o texto do artigo 226, § 3º, não proibiu expressamente outras formas de união que não “entre o homem e a mulher”, o resultado seria o cancelamento da implicatura derivada da superfície do texto como única conclusão normativamente aceitável.

Esse cancelamento, no entanto, apresentado como a óbvia interpretação constitucional do texto do referido artigo, obscurece o importante e ativo trabalho linguístico a que se teve de chegar, a fim de apresentar o sentido da norma como aquela seguida pela unanimidade do/as ministro/as. Também obscurece qual o papel que interpretações literais, limites semânticos e modelos de produção de sentido heterodoxos têm—ou podem ter—sobre a interpretação constitucional. Falarei sobre isso na discussão abaixo.

6. Discussão

A análise acima parte de um modelo descritivo ((7), acima). Modelos descritivos não são de observância “obrigatória” pelo/a usuário/a da língua, senão que procuram captar, explicar, formalizar processos linguísticos normalmente seguidos por esse/as usuário/as. Assumido esse modelo como uma representação simplificada e suficiente de como o/as usuário/as da língua produzem sentido de discursos, especialmente em contextos jurídicos, ele precisa então explicar como uma interpretação constitucional foi empiricamente formulada em um ou mais casos concretos.

É possível assumir que na maioria dos casos em que o sentido é produzido sobre normas jurídicas, o modelo segue linearmente do nível semântico ao nível pragmático, com o significado identificado na superfície do texto indicando sem maiores questões a intenção comunicativa projetada nesse sentido.

Em alguns casos, no entanto, aspectos desse modelo são ativados desproporcionalmente, gerando mecanismos de produção de sentido mais complexos. É razoável supor que esses casos coincidem com aqueles que a teoria do direito chama de “casos difíceis”, nos quais questões morais são mais salientes e há maior incentivo a juízes/as para serem consequencialistas ou instrumentalistas — isto é, para buscarem ativamente interpretações que atinjam objetivos normativos pré-definidos (Struchiner, 2011).

É preciso reconhecer que, ao contrário de uma tarefa de produção de sentido em contextos diferentes ou menos institucionais (como em obras literárias, por exemplo), em contextos jurídicos a interpretação de textos se mistura com questões normativas. Direito é um sistema social de regulação de comportamentos e resolução de conflitos, afinal. Por causa disso, Eskridge e Levi propuseram que o significado de textos normativos seria mais bem caracterizados como (um conjunto de) “variáveis regulatórias”. O autor e autora partem da assunção de que palavras e sintagmas têm um “significado *linguístico*” associado ao seu “significado convencional [isto é, de superfície]”. Porém, seu “significado *jurídico* varia com o propósito da norma, outras políticas públicas e o grau de discricionariedade jurídica” (Eskridge & Levi, 1995, p. 1106, grifos no original). O conceito de variáveis regulatórias procura acomodar esses dois aspectos do processo de produção de sentido de textos normativos jurídicos: “aqueles que aplicam uma norma podem escolher tratar um termo particular da [texto da] norma como se fosse uma variável regulatória, isto é, um termo cuja interpretação jurídica pode variar em função da mudança de circunstâncias” (Eskridge & Levi, 1995, p. 1107, grifo omitido).

Na interpretação da norma do artigo 226, § 3º, da CF/88, o STF parece ter transformado o texto, especialmente o sintagma “entre o homem e a mulher” em uma variável regulatória. O sentido projetado pela presença desse sintagma no texto parece claro a qualquer falante da língua. Implica que a união estável ali regulada é *somente* entre o homem e a mulher. Porém, a mudança nas circunstâncias sociais relativas à homoafetividade desde 1988, que foi objeto da maior parte do voto do Ministro Relator, bem como de outro/as ministro/as, sugere que esse sentido não é mais aceitável para o Tribunal. Por isso, ao tratar o sintagma como uma variável regulatória, o Tribunal pode produzir um sentido parcialmente novo para a norma, cancelando a implicatura.

Comentando o conceito de variável regulatória, Horn (1995) propôs que é possível formular as mudanças que a interpretação jurídica promove no sentido da norma em termos de modificações na própria superfície do texto normativo. Ele usou como exemplo o famoso experimento mental de Herbert Hart sobre uma norma hipotética com o texto “são proibidos veículos no parque”. Discutindo sobre se uma ambulância carregando um doente por dentro do parque violaria ou não a norma, e fazendo referência a um consenso entre juristas de que não violaria, Horn então propõe que essa intuição do jurista pode ser formulada em termos de um “qualificador implícito” no texto (Horn, 1995, 1146–1147):

11. São proibidos veículos não autorizados no parque.

Os dados e a análise deste artigo corroboram essa proposição de Horn. Quando deparados com um caso de controle de constitucionalidade no qual as duas normas (controle e controlada) tinham textos quase idênticos, o STF produziu um sentido do texto-controle (o artigo 226, § 3º, da CF/88) no qual implicitamente adicionou um cancelamento da implicatura conversacional projetada pelo texto (vide (10), acima).

O cancelamento de implicaturas, como vimos, é um fenômeno linguístico comum, usado frequentemente e característico do nível pragmático do significado. A questão interessante no caso de normas jurídicas é que há uma discrepância entre a distribuição do trabalho de produção de sentido em relação a situações de uso ordinário da língua. No uso ordinário, o emissor do discurso é quem cancela a implicatura, como parte da sinalização da sua intenção comunicativa.

No caso da interpretação de normas jurídicas, o emissor das normas comumente é o Poder Legislativo. Quem produz o sentido dessas normas comumente é o Poder Judiciário. Portanto, quando o Judiciário trabalha ativamente na produção de sentido das normas que ele não emitiu—como quando cancela implicaturas, ou adiciona modificadores implícitos, ou cria dispositivos de categorização que mudam a referência da norma (cf. Pádua, 2017, 2019a —o receptor da norma se coloca como seu co-emissor (ou seja, como coautor do texto normativo)).

Essa situação gera uma mudança da divisão de trabalho assumida pela ideia de separação dos poderes, em que um poder cria uma norma geral e abstrata e outro a aplica ao caso concreto, eventualmente criando normas ad hoc apenas em caso de lacunas ou obscuridades. Essa limitação empírica ao desenho institucional da separação de poderes suscita duas outras questões importantes para a teoria do direito.

A primeira é que a postura de coautor da norma pelo Poder Judiciário não é normalmente assumida como tal. Ao contrário, como vimos acima e é claro da leitura do acórdão da ADI 4.277, a interpretação constitucional, mesmo em casos difíceis, é normalmente apresentada como uma demonstração do “correto” sentido da norma, não como uma criação normativa. Em outras palavras: uma mera aplicação da Constituição. Essa postura não só mascara o papel discursivo de coautor da norma do Judiciário, mas também faz supor que é possível, lendo as normas constitucionais, antecipar qual seria essa “correta” aplicação.

A segunda questão importante é qual é o limite da atividade de coautoria que o Poder Judiciário pode arrogar-se na interpretação das normas, no contexto de um desenho institucional de separação de poderes. É intuitivo assumir que normas cuja superfície textual é pobre em modificadores e referências às situações-modelo de sua aplicação suscitam maior espaço para coautoria pelo Poder Judiciário. É o tipo de texto típico da categoria de normas que a teoria do direito vem classificando como “princípios” (cf., em geral, Ávila, 2004). Mas esse não parece ser o caso do artigo 226, § 3º, da CF/88, especialmente se usado como parâmetro de controle do artigo 1.723 do Código Civil.

Quando não se trata de princípios—isto é, quando se trata de regras²⁴ —a doutrina do direito constitucional tem uma resposta para a questão dos limites da atividade interpretativa: os limites são os “limites semânticos do texto” ou a “literalidade do texto” constitucional. A questão passa a ser como esses limites seriam trabalhados e quanta relevância é dada a eles.

Uma maneira de fazer uma primeira exploração sobre isso é vendo como a palavra “literal” aparece no próprio acórdão da ADI 4.277. Usando um código que escrevi em linguagem de programação Python, eu processei o inteiro teor do acórdão, dividi em

²⁴A distinção entre normas e princípios é um tema que, embora antigo, ainda não encontrou formulação conceitual segura no âmbito da teoria do direito. Para o escopo deste trabalho, é suficiente conceituar “princípios” como categorias de normas jurídicas que não definem a situação-modelo (ou “suporte fático”) em que se aplicam e as consequências da sua aplicação, senão enunciam objetivos e valores normativos mais gerais. Normas, por oposição, seriam as categorias de normas que definem sua situação-modelo e as consequências de sua aplicação. O criador da distinção, Dworkin (1967), formulou-a em termos de regras serem de aplicação “tudo-ou-nada” (*all-or-nothing*) e princípios serem “padrões a serem observados” (*standard that is to be observed*) para avançar alguma dimensão de moralidade. Para discussões conceituais sobre a distinção, cf. Ávila (2004).

palavras (isto é, “tokenizei” os dados), eliminei cabeçalhos, números de página e outros ruídos, e procurei isolar dois fenômenos: (1) palavras com as quais a palavra “literal” se combina; e (2) sua frequência de utilização²⁵.

Num universo de quase 75 mil palavras que compõem o inteiro teor do acórdão, a palavra “literal” ocorre apenas 10 vezes²⁶; ou seja, um pouco mais de uma vez a cada 10 mil palavras. Isso por si só não quer dizer muita coisa. Porém, se compararmos a frequência de “literal” (como modificador adjetivo)²⁷ com a frequência de outras palavras na mesma função sintática, notamos, por exemplo, que “literal” ocorre menos de uma vez a cada 500 modificadores adjetivos. E se compararmos o modificador adjetivo mais comum no acórdão—“constitucional”, com 322 ocorrências—com “literal”, vemos que “literal” ocorre com uma frequência 29 vezes menor.

Além disso, ao menos em duas ocorrências de “literal” o contexto do seu uso sugere que essa palavra não fazia referência à interpretação literal como limite ou referência para a produção de sentido do texto normativo. Em ambas, “literal” foi usado como um modificador de reforço (por exemplo, numa referência à “literal categorização” da constituição como “base da sociedade”)²⁸.

Em conjunto, esses dois fenômenos sugerem que o tópico da literalidade como limite da produção e sentido das normas constitucionais esteve presente nas argumentações registradas no acórdão da ADI 4.277, mas com uma importância secundária. O que é algo surpreendente, tendo em vista a peculiaridade já ressaltada deste caso, de que tanto a norma-controle, quanto a norma controlada têm textos quase idênticos. Isto é, era de esperar que, tendo em vista que o STF estava afastando uma interpretação literal de uma norma constitucional, o tema da literalidade fosse ser expressamente tratado com mais atenção, o que se traduziria quantitativamente em mais uso da palavra “literal” como modificador de “interpretação”.

Talvez uma das razões pelas quais isso não tenha ocorrido seja indicada por uma espécie de ato falho retórico do Ministro Relator no seu voto: “que não se faça uso da letra da Constituição para *matar o seu espírito*” (ADI 4.277, p. 653. Grifo no original).

A invocação de uma oposição entre “letra” e “espírito” da constituição, por sua vez, reforça a hipótese aventada acima de que, em casos difíceis como é a ADI 4.277, o STF assume uma postura consequencialista, na qual uma decisão moralmente tida como correta—ou, de qualquer maneira, desejada pela maioria do tribunal—é tomada *a priori*

²⁵Esse dois fenômenos, isto é, frequência de palavras (“tokens”) e a colocação dessas palavras com outras são duas das mais antigas e mais comuns técnicas de pesquisa quantitativa sobre dados de linguagem do campo da linguística chamada linguística de corpus (Sardinha, 2000).

²⁶11 vezes, se contarmos não a palavra “literal”, mas a raiz “literal”, que incluiria palavras derivadas como “literalidade” ou “literalmente”, por exemplo

²⁷Freitas et al. (2022) usaram também esse método, de associar palavras com sua função na relação de predicação para melhor avaliar como a frequência dessas palavras, condicionada à sua função, revelava opções discursivas conectadas a concepções sociais. No caso do trabalho das autoras, a concepção que elas analisavam eram estereótipos associados com o masculino e o feminino em obras literárias. Seu método pode ser estendido para outras concepções, como, no caso deste trabalho, a construção dogmático-jurídica dos “limites semânticos” dos textos normativos.

²⁸Cf. as chamadas linhas de concordância (excertos no qual a palavra aparece, seguida das cinco palavras anteriores e das cinco posteriores) no caderno que documenta os resultados da exploração quantitativa. Vide também, no mesmo caderno, as linhas de concordância da palavra “limite”.

e uma espécie de engenharia discursiva reversa é posta em prática para justificar essa decisão como a única correta à luz do texto constitucional visto como um todo.

Considere-se que casos como o do ADI 4.277 são especialmente interessantes no contexto do consequencialismo, tendo em vista que não se trata de escolher dentre várias interpretações possíveis de um texto vago, aquela que melhor atende ao “espírito” da Constituição. Trata-se de ativamente editar um texto cujo significado é claro na superfície, a fim de adaptá-lo a esse “espírito”.

Esse poder assumido pelo STF nesse caso suscita questões sobre o seu papel como “guardião da Constituição”. “Guardar” a constituição significa aceitar o texto que foi incluído nela pelo legislador constituinte, ou significa, em alguns casos, editar esse texto, a fim de aperfeiçoá-lo com mudanças na leitura moral da Constituição vista como um todo?

Nesse caso, o STF optou pelo segundo caminho, mas o apresentou como se fosse o primeiro. Outros estudos de caso podem averiguar se esse padrão de trabalho discursivo de edição implícita do texto constitucional é recorrente.

7. Conclusão

Neste trabalho, abordei como os domínios da semântica e da pragmática se relacionam com a questão da interpretação—isto é, da produção de sentidos—de normas jurídicas. Comecei mostrando algumas intuições utilizadas por juristas para incorporar esses domínios nas discussões sobre a teoria da interpretação jurídica. Depois, mostrei como essas intuições geralmente estavam fora de sintonia com como os mesmos conceitos vêm sendo teorizados e testados empiricamente no âmbito dos estudos da linguagem—nomeadamente, na linguística e na filosofia da linguagem.

Usei essa revisão conceitual para propor um modelo analítico sobre a produção de sentido de normas jurídicas que tivesse aplicabilidade—isto é, testabilidade—empírica. E usei o modelo para reconstruir o método discursivo utilizado pelo Supremo Tribunal Federal no julgamento da ADI 4.277, com foco no voto do Ministro Relator (que foi seguido à unanimidade). Mostrei que, apesar de construir um argumento interpretativo pelo qual o Tribunal apenas aplicava sistematicamente o texto do artigo 226, § 3º, da CF/88, na verdade o Ministro Relator operou um cancelamento da implicatura conversacional projetada por aquele texto; nomeadamente, pela presença do sintagma “entre o homem e a mulher” como modificador de “união estável”. Mostrei também que é possível formular e demonstrar esse cancelamento em função de modificações previsíveis no texto normativo em questão.

Na seção de discussão, sugeri que esse cancelamento de implicatura demonstra que o STF assumiu a posição de coautor do sentido do mesmo texto, dividindo com o Legislativo Constituinte a tarefa de projetar, no texto, uma intenção comunicativa reconhecível por quem o leia. Também discuti como é provável que, em casos difíceis, o Poder Judiciário fique tentado a utilizar mecanismos discursivos mais ou menos complexos para assumir essa coautoria do sentido da norma, correlacionada a uma postura mais consequencialista na tomada de decisões. Porém também adverti que essa postura não é apresentada como tal; isto é, que o Poder Judiciário, mesmo quando assume coautoria sobre o sentido da norma, formula sua atividade como “apenas aplicando” ou “apenas interpretando” a norma.

Finalmente, ainda nessa parte, apresentei dados quantitativos exploratórios sobre o uso da linguagem pelo STF no acórdão completo da ADI 4.277. Esses dados sugerem que a literalidade como limite à criatividade no processo interpretativo, embora reconhecida pontualmente, não é vista como um tópico fundamental, nem mesmo neste caso, em que há identidade de textos entre as normas-controle e a controlada.

A revisão de literatura e os resultados analíticos desse artigo indicam que a teoria da interpretação jurídica ainda está aquém da tarefa de descrever os mecanismos discursivos empiricamente utilizados pelo Poder Judiciário para produzir o sentido das normas que é incumbido de aplicar. E que, mesmo quando tenta invocar conceitos dos estudos da linguagem, essa teoria, em geral, deixa evidenciada a falta de compreensão do seu escopo e potencial descritivo. A utilização da interface entre a teoria jurídica e estudos linguísticos pode ajudar a superar esse problema e esclarecer melhor como o Poder Judiciário produz sentido das normas jurídicas e o que isso acarreta para a divisão de trabalho normativo entre Legislativo e Judiciário. E pode contribuir para uma melhor compreensão sobre a linguagem da interpretação de normas jurídicas.

References

- Afonso da Silva, J. (2001). *Curso de direito constitucional positivo*. São Paulo: Malheiros.
- Afonso da Silva, V. (2005). Interpretação constitucional e sincretismo metodológico. In V. Afonso da Silva (Ed.), *Interpretação constitucional* (pp. 115–143). Malheiros.
- Afonso Da Silva, V. (2006). Interpretação conforme a constituição: Entre a trivialidade e a centralização judícia. *Revista Direito GV*, 191–210.
- Allott, N., & Shaer, B. (2017). Inference and intention in legal interpretation. In J. Giltrow & D. Stein (Eds.), *The pragmatic turn in law* (pp. 83–118). Berlin: De Gruyter.
- Arribas-Ayllon, M., Sarangi, S., & Clarke, A. (2011). Rhetorical discourse analysis. In M. Arribas-Ayllon, S. Sarangi, & A. Clarke (Eds.), *Genetic Testing* (pp. 55–77). London: Routledge.
- Atienza, M. (2003). *As razões do direito: Teorias da argumentação jurídica* (3rd edition ed.). São Paulo: Landy.
- Auroux, S. (2018). *La philosophie du langage*. Paris: PUF.
- Barroso, L. R., & Barcellos, A. P. d. (2006). O começo da história. A nova interpretação constitucional e o papel dos princípios no direito brasileiro. In L. R. Barroso (Ed.), *A nova interpretação constitucional: Ponderação, direitos fundamentais e relações privadas*. Rio de Janeiro: Renovar.
- Bittar, E. C. B. (2008). *Linguagem jurídica* (3rd edition ed.). São Paulo: Saraiva.
- Britto, C. A. (2006). *Teoria da constituição*. Rio de Janeiro: Forense.
- Carvalho, P. d. B. (2016). O legislador como poeta: Alguns apontamentos sobre a teoria flusseriana aplicado ao Direito. In R. Pinto, A. L. T. Cabral, & M. d. G. S. Rodrigues (Eds.), *Linguaem e direito: Perspectivas teóricas e práticas* (pp. 11–27). São Paulo: Contexto.
- Coulthard, M. (1997). A failed appeal. *Forensic Linguistics*, 4(2).
- Dworkin, R. (1967). The Model of Rules. *Chicago Law Review*, 35(1), 14–46.
- Egbert, J., Larsson, T., & Biber, D. (2020). *Doing Linguistics with a Corpus*. Cambridge: Cambridge University Press.
- Epstein, L., & King, G. (2002). The rules of inference. *The University of Chicago Law Review*, 69(1), 1–133.
- Eskridge, W. N., & Levi, J. N. (1995). Regulatory variables and statutory interpretation. *Washington University Law Quarterly*, 73(3).
- Felder, E., & Friedemann, V. (2014). Sprache im Recht. In E. Felder & A. Gardt (Eds.), *Handbuch Sprache und Wissen* (pp. 358–372). Berlin: De Gruyter.
- Ferraz Jr., T. S. (1994). *Introdução ao estudo do direito: Técnica, decisão, dominação*. (3rd edition ed.). São Paulo: Atlas.
- Ferraz Jr., T. S. (2011). Direito, linguagem e interpretação. In *Direito e interpretação: Racionalidades e instituições* (pp. 103–18). São Paulo: Saraiva.
- Frank, M. C., & Goodman, N. D. (2012). Predicting Pragmatic Reasoning in Language Games. *Science*, 336(6084), 998–998. Retrieved 2024-03-11, from <https://www.science.org/doi/10.1126/science.1218633> doi: 10.1126/science.1218633
- Frege, G. (2007). Über Sinn und Bedeutung. In M. Textor (Ed.), *Funktion - Begriff - Bedeutung* (pp. 23–46). Göttingen: Vandenhoeck & Ruprecht.
- Freitas, C., Martins, F., & Biar, L. (2022). Um ‘olhar discursivo’ sobre predicação e gênero: aproximações metodológicas entre corpus e discurso. *Texto Livre*, 15, e36213. Retrieved 2024-03-11, from <https://periodicos.ufmg.br/index.php/textolivre/article/view/36213> doi: 10.35699/1983-3652.2022.36213

- Geis, M. L. (1995). The Meaning of meaning in the law. *Washington University Law Quarterly*, 73, 1125–1144.
- Goffman, E. (1979). Footing. *Semiotica*, 25(1-2), 1–30. Retrieved 2024-03-11, from <https://www.degruyter.com/document/doi/10.1515/semi.1979.25.1-2.1/html> doi: 10.1515/semi.1979.25.1-2.1
- Goodman, N. D., & Frank, M. C. (2016). Pragmatic Language Interpretation as Probabilistic Inference. *Trends in Cognitive Sciences*, 20(11), 818–829. Retrieved 2024-03-11, from <https://linkinghub.elsevier.com/retrieve/pii/S136466131630122X> doi: 10.1016/j.tics.2016.08.005
- Grice, H. P. (1957). Meaning. *Philosophical Review*, 66(3), 377–388.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics, vol. 3: Speech acts* (pp. 41–58). London: Academic Press.
- Habermas, J. (1984). *The theory of communicative action. Volume one: reason and the rationalization of society*. Boston: Beacon.
- Hagemeyer Burgo, V., & Araújo, C. P. d. E. d. (2018). O princípio da cooperação em audiências judiciais: as máximas conversacionais em casos de violência contra mulheres. *Cadernos Discursivos*, 1(1), 189–204.
- Hart, H. L. A. (1961). *The concept of law*. New York: Oxford University.
- Haugh, M. (2008). Intention in pragmatics. *Intercultural Pragmatics*, 5(2), 99–110.
- Horn, L. (1995). Vehicles of meaning: unconventional semantics and unbearable interpretation. *Washington University Law Quarterly*, 73(3), 1145–1152.
- Hovy, D. (2022). *Text Analysis in Python for Social Scientists: Prediction and Classification* (Vol. 3794).
- Huang, Y. (2017). Neo-gricean pragmatics. In Y. Huang (Ed.), *The oxford handbook of pragmatics*. Oxford, UK: Oxford University.
- Jurafsky, D. (2015). *The language of food: A linguist reads the menu*. New York: Norton & Co.
- Jurafsky, D., & Martin, J. H. (2019). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Retrieved from https://web.stanford.edu/~jurafsky/slp3/edbook_oct162019.pdf
- Kommers, D. P. (1997). *The constitutional jurisprudence of the Federal Republic of Germany* (2nd edition ed.). Durham: Duke University.
- Levinson, S. C. (2003). Putting linguistics on a proper footing: Explorations in Goffman's concepts of participation. In P. Drew & A. Wootton (Eds.), *Erving Goffman: Exploring the Interaction Other* (pp. 161–227).
- Levinson, S. C. (2007). *Pragmática*. São Paulo: Martins Fontes.
- Martins, C. (2002). A Indeterminação do Significado nos Estudos Sócio-Pragmáticos: Divergências Teórico-Methodológicas. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada*, 18(1), 87–116. Retrieved 2024-03-11, from http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-44502002000100004&lng=pt&tlng=pt doi: 10.1590/S0102-44502002000100004
- Maximiliano, C. (2002). *Hermenêutica e aplicação do direito* (19th edition ed.). Rio de Janeiro: Forense.
- McCawley, J. D. (1993). *Everything that linguists have always wanted to know about logic but were ashamed to ask*. (2nd edition ed.). Chicago: University of Chicago.
- Oliveira, R. P. d., & Basso, R. M. (2007). A semântica, a pragmática e os seus mistérios.

- Revista Virtual de Estudos da Linguagem - ReVEL*, 5(8), 1–30.
- Perini, M. A. (2006). *Gramática descritiva do Português* (4th edition ed.). São Paulo: Ática.
- Pádua, J. P. (2017). Audiência de custódia e processo penal oculto: Esboços teóricos de um programa de pesquisa. In A. E. R. Santoro & C. E. Gonçalves (Eds.), *Audiência de Custódia* (pp. 335–353). Belo Horizonte: D'Plácido.
- Pádua, J. P. (2018). Notas semânticas sobre a interpretação da legítima defesa, por ocasião da proposta do novo governador do estado do Rio de Janeiro sobre o abate de criminosos. *Boletim IBCCRIM*, 313, 10–15.
- Pádua, J. P. (2019a). Categorizações como método de produção de sentido de normas constitucionais: Contribuições da interface entre direito e análise do discurso. *Revista de Informação Legislativa*, 56, 87–109.
- Pádua, J. P. (2019b). Discursive devices for inserting morality into law: initial exploration from the analysis of a Brazilian Supreme Court decision. *Language and Law=Linguagem e Direito*, 6(1), 11–29. Retrieved 2024-03-11, from <http://ojs.letras.up.pt/index.php/LLLD/article/view/6261/5896> doi: 10.21747/21833745/lanlaw/6_1a1
- Pádua, J. P. (2020a). Modelos normativos e modelos descritivos na pesquisa sociojurídica. In C. F. Alves et al. (Eds.), *Instituições da justiça e do trabalho, volume 2: Campos de pesquisa*. Rio de Janeiro: Autografia.
- Pádua, J. P. (2020b). O Sistema judicial brasileiro e linguística forense: Linguística para juristas e direito para linguistas. In D. C. d. Almeida, M. Coulthard, & R. Sousa-Silva (Eds.), *Perspectivas em linguística forense* (pp. 8–26). Campinas: Unicamp.
- Pádua, J. P. (2020c). Repristinação interpretativa de normas penais em tempos de crise: Limites discursivos e empíricos aos limites normativos dos tipos penais. In A. J. C. Cunha Filho, C. S. L. Arruda, R. H. Issa, & R. W. Schwind (Eds.), *Direito em tempos de crise: Covid-19, Volume II: Constitucional - Ambiental - Econômico* (pp. 61–84). Bom Retiro, SP.
- Recanati, F. (2001). What Is Said. *Synthese*, 128, 75–91.
- Recanati, F. (2002). Does Linguistic Communication Rest on Inference? *Mind & Language*, 17(1-2), 105–126. Retrieved 2024-03-12, from <https://onlinelibrary.wiley.com/doi/10.1111/1468-0017.00191> doi: 10.1111/1468-0017.00191
- Roberts, C., & Sarangi, S. (2005). Theme-oriented discourse analysis of medical encounters. *Medical Education*, 39(6), 632–640. Retrieved 2024-03-12, from <https://onlinelibrary.wiley.com/doi/10.1111/j.1365-2929.2005.02171.x> doi: 10.1111/j.1365-2929.2005.02171.x
- Rühlemann, C. (2019). *Corpus linguistics for pragmatics: A guide for research*. New York: Routledge.
- Sarangi, S. (2005). The Conditions and Consequences of Professional Discourse Studies. *Journal of Applied Linguistics*, 2(3), 371–394. Retrieved 2024-03-11, from <http://www.equinoxpub.com/journals/index.php/JAL/article/view/4578> doi: 10.1558/japl.v2i3.371
- Sardinha, T. B. (2000). Linguística de Corpus: histórico e problemática. *DELTA: Documentação de Estudos em Linguística Teórica e Aplicada*, 16(2), 323–367. Retrieved 2024-03-11, from http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-44502000000200005&lng=pt&tlng=pt doi: 10.1590/S0102-44502000000200005

- Savigny, F. C. v. (1878). *Sistema del derecho romano actual: Tomo I*. Madrid: Góngora y Campaña.
- Schauer, F. (1998). On The Supposed Defeasibility of Legal Rules. *Current Legal Problems*, 51(1), 223–240. Retrieved 2024-03-11, from <https://academic.oup.com/clp/article-lookup/doi/10.1093/clp/51.1.223> doi: 10.1093/clp/51.1.223
- Schier, P. R. (2005). Novos desafios da filtragem constitucional no momento do neoconstitucionalismo. *Revista de Direito Administrativo & Constitucional*, 5(25), 145–165.
- Sgarbi, A. (2007). *Teoria do direito: Primeiras lições*. Rio de Janeiro: Lumen Juris.
- Shecaira, F. P., & Struchiner, N. (2015). Direito e linguagem: Aspectos filosóficos. In S. Silveira, C. S. Abritta, & A. T. Vieira (Eds.), *Linguística aplicada em contextos legais* (pp. 22–41). Jundiaí: Paco.
- Shuy, R. W. (1993). *Language crimes: the use and abuse of language evidence in the courtroom*. Blackwell.
- Slocum, B. G. (2017). Pragmatics and legal texts: How best do account for the gaps between literal meaning and communicative meaning. In J. Giltrow & D. Stein (Eds.), *The pragmatic turn in law* (pp. 119–144). Berlin: De Gruyter.
- Solan, L. M. (1998). Law, Language, and Lenity. *William and Mary Law Review*, 40(1), 57–144.
- Solan, L. M. (2018). The Interpretation of Legal Language. *Annual Review of Linguistics*, 4(1), 337–355. Retrieved 2024-03-11, from <https://www.annualreviews.org/doi/10.1146/annurev-linguistics-011817-045649> doi: 10.1146/annurev-linguistics-011817-045649
- Solan, L. M., & Gales, T. (2016). Finding ordinary meaning in law: The judge, the dictionary or the corpus? *International Journal of Legal Discourse*, 1(2), 253–276. Retrieved 2024-03-11, from <https://www.degruyter.com/document/doi/10.1515/ijld-2016-0016/html> doi: 10.1515/ijld-2016-0016
- Stein, D. (2017). On inferencing in law. In J. Giltrow & D. Stein (Eds.), *The pragmatic turn in law* (pp. 335–365). Berlin: De Gruyter.
- Streck, L. L. (2006). *Verdade e consenso: Constituição, hermenêutica e teorias discursivas*. Rio de Janeiro: Lumen Juris.
- Struchiner, N. (2002). *Direito e linguagem: Uma análise da textura aberta da linguagem e sua aplicação ao direito*. Rio de Janeiro: Renovar.
- Struchiner, N. (2011). Indeterminação e objetividade. Quando o direito diz o que não queremos ouvir. In R. P. Macedo Jr & C. Barbieri (Eds.), *Racionalidades e Instituições*. São Paulo: Saraiva.
- Struchiner, N., & Hannikainen, I. R. (2020). An experimental guide to vehicles in the park. *Judgment and Decision Making*, 15(3), 312–29.
- Struchiner, N., & Shecaira, F. P. (2012). A distinção entre direito e moral e distinção moral do direito. *Revista de Direito do Estado*, 7(22), 131–45.
- Terkourafi, M. (2021). Pragmatics as an interdisciplinary field. *Journal of Pragmatics*, 179, 77–84. Retrieved 2024-03-11, from <https://linkinghub.elsevier.com/retrieve/pii/S0378216621001466> doi: 10.1016/j.pragma.2021.04.015
- Tiersma, P. M. (2000). *Legal language*. Chicago: University of Chicago.
- Tiersma, P. M. (2001). A message in a bottle: Text, autonomy and statutory interpretation. *Tulane Law Review*, 76(2), 431–82.
- Von Wright, G. H. (1999). Deontic Logic: A Personal View. *Ratio Juris*, 12(1), 26–38.

Retrieved 2024-03-11, from <https://onlinelibrary.wiley.com/doi/10.1111/1467-9337.00106> doi: 10.1111/1467-9337.00106

Voskuhle, A. (2000). Theorie und Praxis der verfassungskonformen Auslegung von Gesetzen durch Fachgerichte. *Archiv des öffentlichen Rechts*, 125, 177–201.

Ávila, H. (2004). *Teoria dos princípios: da definição à aplicação dos princípios jurídicos* (4th edition ed.). São Paulo: Malheiros.

Joke or threat? Competing genre uptakes in a Danish court case

Marie Bojsen-Møller ¹

¹Institut for Nordiske Studier og Sprogvidenskab, mbm@hum.ku.dk

Abstract

This article examines a Danish criminal case concerning a high school student who posted a picture on Instagram with the caption ‘I’ll be the next school shooter guys, lmao, watch out’ (Danish: jeg bliver den næste school shooter guys, lmao, watch out). Reactions to – or uptakes on (Freadman, 1994, 2002) – the Instagram post were split between two types; some took it as a threat and some took it as a joke. Building on theory on threatening communications (e.g. Bojsen-Møller, Auken, Devitt, & Christensen, 2020; Fraser, 1998; Gales, 2019; Solan & Tiersma, 2005), on humor (Billig, 2005; Norrick, 2010; Tsakona, 2017) and on genre theory (Devitt, 2009; Freadman, 2012, 2020; Miller, 1984), this article studies the diverging uptakes as they are presented within the court case through examination of both transcripts of audio recordings from the court room and court documents. The findings of this study show that while linguistic interpretation is mostly limited to the netronym ‘lmao’, other non-linguistic behaviors are often used as justification for a particular interpretation of the linguistic behavior. Not only the prosecuted utterance itself, but also the utterances’ uptakes, i.e. an essential aspect of its context, have a bearing on the outcome of the case and the assessment of the defendant’s intent. As such, this study offers important insight into an area that Gales (2019) argues is in need of further interdisciplinary research, namely how contextual features can influence the outcome of criminal cases concerning disputed threats.

Keywords: Threatening Communications, Jokes, Uptake, Legal Genres, Textual Travels.

Resumo

Este artigo analisa um processo penal dinamarquês relativo a um estudante do ensino secundário que publicou uma fotografia no Instagram com a legenda "Serei o próximo atirador da escola, pessoal, lmao, cuidado" (em dinamarquês: jeg bliver den næste school shooter guys, lmao, watch out). As reações à publicação no Instagram - ou UPTAKES (Freadman, 1994, 2002) - dividiram-se em dois tipos: uns consideraram-na uma ameaça e outros uma piada. Com base na Teoria de Ameaça à Face (por exemplo, Bojsen-Møller et al., 2020; Fraser, 1998; Gales, 2019; Solan & Tiersma, 2005), no humor (Billig, 2005; Norrick, 2010; Tsakona, 2017) e na teoria de género (Devitt, 2009; Freadman, 2012, 2020; Miller, 1984), este artigo estuda as

reações divergentes apresentadas no processo judicial através da análise das transcrições das gravações áudio da sala de audiências e dos documentos do tribunal. As conclusões deste estudo mostram que, embora a interpretação linguística se limite sobretudo ao netronym "lmao", outros comportamentos não linguísticos são frequentemente utilizados como a justificação para uma determinada interpretação do comportamento linguístico. Não só o próprio enunciado processado, mas também as suas consequências, ou seja, um aspeto essencial do seu contexto, têm influência no resultado do processo e na avaliação da intenção do arguido. Como tal, este estudo fornece uma visão importante sobre uma área que Gales (2019) defende necessitar de mais investigação interdisciplinar, nomeadamente, a forma como as características contextuais podem influenciar o resultado de processos penais relativos a ameaças contestadas.

Palavras-chave: *Comunicações de ameaças, piadas, aceitação, géneros jurídicos, viagens textuais.*

1. Introduction

On September 11, 2018 around midnight a Danish high school student posted a picture of herself on Instagram with the following caption: ‘I’ll be the next school shooter guys, lmao, watch out’ (Danish: *jeg bliver den næste school shooter guys, lmao, watch out*)¹. This case study examines data from the criminal case that ensued, consisting of police record, court judgment and transcriptions of audio recordings from the trial. The data set shows that the court representatives discussed whether the Instagram post was to be interpreted as a threat or a joke, the latter in part with reference to the netronym ‘lmao’ (an abbreviation of the slang idiom ‘laughing my ass off’). My analysis shows that not only the formulation itself, but also the immediate and later reactions to the post were decisive for the outcome of the case. This article examines these reactions as genre uptakes (Freadman, 1994, 2002, 2020), meaning the responses and reactions to an utterance that show that an utterance is considered to be an instantiation of a given genre.

In this study, I analyze the different *uptakes* of the utterance as they appear in the data set, including the immediate uptakes of witnesses, as well as the later uptakes of the court’s representatives (prosecutor, defense lawyer, judges). Furthermore, I analyze the defendant’s *own* uptakes of her utterance as *accounts* of untoward behavior (Atkinson & Drew, 1979; M. L. Komter, 1994; Scott & Lyman, 1968).

Very little research has been conducted on the uptakes of disputed threatening communications by sender, recipients, victims and court systems (however, see Bojsen-Møller, 2022; Bojsen-Møller et al., 2020). This is an important empirical aspect to consider when working with criminal genres such as threatening communications, as it

¹All data in this article are translated by me (unless otherwise explicitly stated). See appendix A for a transcription in Danish of the original audio recording and of the original data from the police report and the court judgment that are analyzed in this study. Also, see appendix B for an English translation of the transcriptions of the audio recording.

offers insight into how these genres function and how contextual factors such as uptakes of an utterance can influence the outcome of a criminal case. In fact, Gales (2019) addresses the need for such interdisciplinary research on contextual aspects of threatening communications as well as on threatening words (locution), threatener intent (illocution) and effects on recipients (perlocution) (cf. Austin, 1962).

It is beyond the scope of this article to resolve the question of whether the utterance in this case study is, in fact, (meant as) a threat or a joke. Instead, this study has two primary and interrelated aims. The first aim is to show how one utterance within its specific and largely undisputed context can have distinctly diverging uptakes. Importantly, these diverging uptakes have entirely different sets of evidence supporting their interpretations. The prosecutor focused on the fact that the high school student had angrily flipped over a table in class on the same day as she posted the Instagram post, while the defense lawyer focused on the fact that she had a cosy afternoon with her two friends.

The second aim of the study is to show that these varying uptakes can have an important impact on how the intent of the sender of the utterance is debated and negotiated in court and court documents, including in judges' verdicts.

When these uptakes are presented in court and court documents, they will often have undergone different kinds of *textual travels* (Heffer, Rock, & Conley, 2013), which will be discussed when relevant. As such, this article is an interdisciplinary study that combines Forensic Linguistics with Rhetorical Genre Studies in incorporating research on threatening communications, textual travels within the legal system and genres and their uptakes.

1.1. Background

To situate the Instagram post, some background knowledge of the incidents leading up to and following from it will be useful. Earlier in the day, the high school student had felt bullied by a classmate with whom she was supposed to do a joint presentation. In frustration, she flipped over a classroom table. Two classmates summoned a teacher who knew her well and who had a chat with her. Later in the day, she went to a friend's house where she hung out with two friends. It was when she had returned home that she posted the picture of herself with the caption 'I'll be the next school shooter guys, lmao, watch out' in large, red letters. Subsequently, a classmate's mother wrote an e-mail to a school representative about the Instagram post and the earlier incident in class, and the school representative then reported the post to the police. The next day, armed police met the girl on the school grounds and arrested her for having threatened to commit a school shooting. The case was brought to trial by 'the Prosecution Service' (Danish: *Anklagemyndigheden*), first in district court and later in high court. Denmark has a three-tiered court system with 24 district courts, two high courts (the Western High Court and the Eastern High Court) and one Supreme Court. In the local district court, the case was judged by a 'court with lay judges' (Danish: *domsmandsret*) consisting of one professional judge and two lay judges. In high court, three professional judges and three lay judges adjudicated the case. The high school student herself explained to the police and to the courts that she meant the Instagram post as a joke. There were, thus,

two distinct kinds of genre uptakes to the post; some took it as a threat and some took it as a joke.

2. Theory

2.1. Genre, uptake and textual travels

Genres are utterances that are considered a type or a kind through perceived *recurrence* in their usage. As such, Rhetorical Genre Studies (RGS) describe genres as communicative *social actions* that are *typified* (Miller, 1984). The focus of RGS is typically more on the functional aspects of a genre and less on its formal traits, as communicative function directs genres and thereby influences their form. However, as a linguist, I agree with RGS scholars Devitt (2009) and Auken (2015) that both function *and* linguistic form have to be considered, as “[a]ny complete understanding of genre will need to include the language forms that serve to achieve [its] purposes and effects, the forms that make generic action happen” (Devitt, 2009, p.27). Genres are recognized through actual or perceived recurrence in form, function, content and context, but genre recognition is not necessarily only contingent on homogeneity in form (Medway, 2002). For instance, Bojsen-Møller et al. (2020) argue that threatening communications are in fact an exceedingly heterogeneous genre, but nonetheless a highly recognizable one. Some genres are much more stringent and formulaic than others (cf. Bojsen-Møller, 2022; Heffer, 2013; Heydon, 2013) on the formulaic nature of legal genres. For an overview of different kinds of genres, see Miller (2017).

In accordance with the above, Miller argues that “in genre we are seeking not the stabilization of [practices and performances] but rather the stabilization of shared recognitions and social agreements, as represented by naming, replication, and meta-commentary” (Miller, 2017, p.5) Responding to a genre either by labelling it (Nyboe, 2016), by commenting on it or by referring to it through different intertextual techniques is in genre terms called uptake (Freadman, 1994, 2002, 2020). Borrowing Austin (1962) term “uptake”, Freadman (2020) explains that uptake is to be understood in light of the Bakhtinian notion of utterances preceding and succeeding each other in “a very complexly organized chain of other utterances” (Bakhtin, 1986, p.69). In other words, the context of an utterance, including its uptakes, consists of “a sequence of events in time and across a variety of temporal sequences” (Freadman, 2012, p.558). Through their individual usages and uptakes, genres are changed and adapted (Devitt, 2009). Uptake is conditioned by “a relation of reciprocal difference” between genres (Freadman, 2020: 105), but genres also intertwine and mix through their uptakes, sometimes creating new genres if the reciprocal differences between the “source” genre and its uptakes become too great. Every utterance is in a complex dialogic, intertextual relationship with previous utterances. The most explicit way of taking up an utterance as a particular genre is by labelling it, and alternatively by describing its perceived purposes and effects (metacommentary) or simply responding to it in a way that shows that it is considered an instantiation of a given genre. However, uptake can also happen more indirectly through *textual travels* (different kinds of intertextuality; Heffer, 2013) of one genre into another genre (cf. Bojsen-Møller (2022), who links the notion of uptake to textual travels). As such, genres or parts of genres will often be *embedded* within other

genres (Auken, 2021). See figure 1 below for an overview of the genre embedding in this case beginning with the posting of the utterance and ending with the high court judgment and court record:

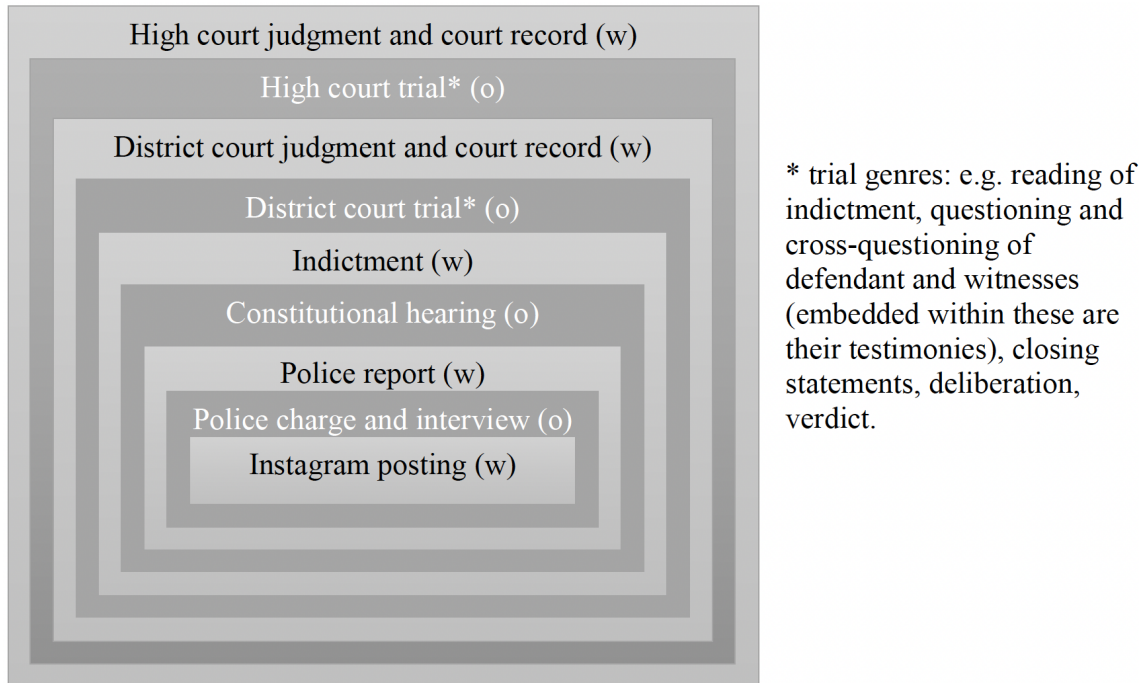


Figure 1. Embedding of genres in the case study (w=written, o=oral)

Figure 1 above is not meant to convey that an entire genre is embedded within another genre², but that traces from the previous genres are embedded within the encompassing genre. These traces can consist of different kinds of textual travels (of words, phrases, whole paragraphs or utterances), or they can consist more indirectly of sentiments, attitudes or norms from the previous genres. The fact that this legal *genre set* (Devitt, 1991) switches between written and oral genres means that different types of textual travels (e.g. summaries, direct and indirect discourse) are used throughout.³

Legal representatives are in possession of overt but also tacit knowledge (Auken, 2018; Freadman, 1994) of the genres in the legal genre set. The preceding genres in the legal genre set therefore lend themselves in design to the succeeding genres to secure a smooth uptake. As such, an utterance can project its desired uptake, but “no speaker or writer can completely secure an uptake” (Freadman, 2012, p.560). This is especially true when it comes to illicit genres and vernacular genres, as these will often be less homogeneous than for instance institutional genres (Bojsen-Møller et al., 2020; Miller, 2017). Genre labelling can then become a “discursive battleground” (Auken, 2019), where some people will take up an utterance as an instantiation of one genre and other people will take up the same utterance as an instantiation of another genre. A person’s uptake will (supposedly) be a reflection of their interpretation of the utterance in question –

²For example, the entire police report is not reproduced within the indictment. However, the entire district court judgment is reproduced within the high court judgment (for an overview of the structure of records of judgments in Denmark, see Bojsen-Møller, 2022).

³For an overview of some of the different textual travels that happen in the Danish legal system, see Bala (2020) and Bojsen-Møller (2022).

and also likely a reflection of the *perlocutionary* effect (Austin, 1962) that the utterance has had on them or others – but their uptake will also reflect the social action they themselves wish to accomplish.

2.2. Uptake on own utterance: Accounts

For the purposes of this article, it is important to distinguish between uptakes made on someone else's utterance and uptakes made on one's own utterance – what might be termed self-uptake. In a Bakhtinian sense, the “boundaries of an utterance⁴ are delimited solely by a change in speaking subject” (Bakhtin, 1986, p. 71), and in light of that, an uptake on one's own utterance would in principle have to happen *after* another person has taken up the utterance. In a courtroom setting, the notion of *accounts* (Scott & Lyman, 1968) becomes particularly relevant as a specific form of uptake that defendants often have on their own previous utterances or behavior (cf. Atkinson & Drew, 1979; M. L. Komter, 1994). An account is “a statement made by a social actor to explain unanticipated or untoward behavior” (Scott & Lyman, 1968, p. 46), and an account will primarily be generated as a response to an accusation. There are two main types of accounts, namely *excuses* and *justifications* (Scott & Lyman, 1968). One subtype of excuses relevant to this study is called ‘appeals to defeasibility’, which is when the person accused of untoward behavior makes a claim that would invalidate the accusation. This could for instance be that they did not intend for the behavior to be untoward or offensive, or that they had no knowledge of it being possibly understood as such. For the sake of clarity, the narrower term ‘appeals to intent’ is proposed and used throughout this article. Threatening communications can indeed be categorized as untoward—or what Bojsen-Møller et al. (2020) term “illicit”—behavior, especially if they are prosecuted as illegal threatening communications.

2.3. Threatening communications and the Danish threat statute

Existing Forensic Linguistic literature on threatening communications has identified salient linguistic features in threatening communications (Gales, 2010; Muschalik, 2018; Nini, 2017) and has offered important tools with which to analyze and assist in recognizing the speech act of threatening (Christensen, 2019; Fraser, 1998; Gingiss, 1986; Salgueiro, 2010; Yamanaka, 1995). Within Forensic Linguistics, threatening communications are described in terms that highlight

- their **illegality** (language crimes (Shuy, 1993), crimes of language (Solan & Tiersma, 2005), illegal speech acts (Christensen, 2019), and
- their **social disruptiveness** (linguistic acts of power (Gales, 2010), illicit genre (Bojsen-Møller et al., 2020)).

The general consensus among threat scholars is that the main function of threats is to intimidate their targets (Bojsen-Møller et al., 2020; Christensen, 2019; Fraser, 1998; Gales, 2010; Solan & Tiersma, 2005; Walton, 2000) which is part of the reason many threats are considered illegal acts in and of themselves in most societies. This act of intimidation is achieved by communicating that a future event will happen that is to the detriment of a victim and for which the threatener takes responsibility (Bojsen-Møller et al., 2020; Christensen, 2019; Fraser, 1998; Shuy, 1993). These aspects (futurity,

⁴This is similar to the delimitation of *turns* in conversational analysis (CA), which is unsurprising since both RGS and CA have roots in ethnomethodology (cf. Miller, 1984; Watson, 1992).

future harm, victim, sender's responsibility) can either be communicated directly or indirectly (Bojsen-Møller et al., 2020; Christensen, 2019; Gales, 2010; Muschalik, 2018). Threats that are indirectly phrased (such as 'I know where you live') are considered more amenable to *plausible deniability* in court (Solan & Tiersma, 2005) than direct threats (such as 'I will kill you'), since more aspects will have to be inferred from the context (Christensen, 2019). However, not only indirect threats require attention to contextual features: Gales (2019) emphasizes the fact that "the historical and conversational context in which a threat is allegedly made cannot be excluded from a threat's core definition" (475). Threat scholars have to consider form, function, effect and context in their research on threatening communications.

The Danish Criminal Code's § 266 criminalizes threats to *carry out an illegal act* that are *fit to provoke serious fear*, which means that it is the illegality of the threatened act and the potential of the threat for creating fear that determines whether a threat is illegal or not:

Whosoever threatens to carry out an illegal act in a way that is fit to provoke serious fear in someone for their own or other people's lives, health or wellbeing, shall be penalized by fine or imprisonment of up to 2 years (the Danish Criminal Code, Chapter 27, § 266. The translation is from Christensen, 2019).

As such, the act of intimidation and the *potential* perlocutionary effect of fear are central to Danish law. The victims of a threat need not actually have been scared, but the threatening utterance in its specific context will have to be considered fit – suitable – in general to elicit fear (Toftegaard Nielsen, Elholm, & Jakobsen, 2017).

The Danish threat statute does not explicitly mention intent (Danish: *forsæt*) but there are in fact three degrees of intent in Danish jurisprudence that also apply to threats: direct intent, probability intent, and *dolus eventualis* (which can be translated to 'possible deceit').

As the highest degree of intent, **direct intent** means that the offender is believed to have acted with 'knowledge and will' (cf. *mens rea*, 'criminal aforethought', 'a guilty mind') when they committed their offence (Waaben & Langsted, 2015, pp.171-172). **Probability intent** expresses a lower degree of intent, where the offender is believed to have considered it 'predominantly probable' (Danish: *overvejende sandsynligt*) that their action would have a certain consequence (Waaben & Langsted, 2015, p.173). **Dolus eventualis** is the lowest - and rarely used - level of intent in Danish law, where the offender is believed to have considered it 'possible' (Danish: *muligt*) that their action would have a certain consequence, but they consciously disregarded this and accepted the risk. *Dolus eventualis* is sometimes describes as 'conscious recklessness' (Danish: *bevidst uagtsomhed*) (Waaben & Langsted, 2015, pp.174-175).

Establishing intent is a notoriously difficult task (Fuller, 2015; Hurt & Grant, 2019; Shauer, 2003; Waaben & Langsted, 2015), since it is inherently a mental state that can thus only be assessed indirectly. Therefore, even though it is a necessary part of legal proceedings, it is not sufficient to consider intent, since such an assessment will always rest on premises that indirectly point to a person's state of mind. In most cases, there will be corroborating evidence that is important to consider. For instance, specific contextual features that may help clarify whether it is probable that the defendant had an intent to threaten. In threat cases, one of the contextual features may be the uptake by others in the specific situation since it shows a situated reaction to the possible threat.

In cases where diverging uptakes exist, like in the present case, the task of assessing intent becomes even more complicated.

In the case of threatening communications, direct intent would be proven if the offender were judged to have understood and accepted that their utterance was a threat that was fit to create fear, meaning they had knowledge and will. Probability intent would be proven if the offender were judged to have found it predominantly likely that the result of their words would be fear. *Dolus eventualis* would be proven if the offender were judged to have found it possible that the result of their words would be fear but accepted it and acted in the same way regardless.

2.4. Threats that are (purportedly) meant as jokes

An interesting uptake on the genre of threats are utterances that have the form of a threat (a future, harmful act towards a victim that the sender of the utterance takes responsibility for), but that the sender purports are intended as jokes. In such cases, context is key, since giving a threat a different genre label or negating the genre label – what Bojsen-Møller et al. (2020) call *re-labelling* and *anti-labelling* – can be a possibly useful explanation in a courtroom in terms of appealing to plausible deniability. The sender's *intent* may of course in fact have been to make a joke (see Threatening communications and the Danish threat statute), and recipients may or may not have perceived it as such. This is why uptake is such an important notion when pertaining to genres. The uptakes of a genre utterance – especially the immediate uptakes – can help establish how the genre utterance functioned within the situation it was made. In some instances, such utterances will still be taken up as threats (acts of intimidation), and in other instances, they will be taken up as jokes (acts of humor), and as we see in this case study, sometimes both uptakes will be present. In order to better understand this dichotomy, it is important to examine the notion of humor.

Humor is frequently considered *non-bona-fide* or non-serious communication (Raskin, 1985), and it is often used “to create humorous versions of genres or to manipulate [or violate] generic conventions” (Toftegaard Nielsen et al., 2017, p.489). Much humor thus relies on tacit knowledge of genres, and in such cases, humor can be seen as pretense where the sender expects their pretense recognized by others and does not intend it to be taken seriously (Norrick, 2010, p. 233; cf. the Gricean notion of implicature). However, such pretense will not always be recognized by others, which will then result in miscommunication and failed humor (Bell, 2017). Especially if humor occurs in a genre where it is not normally expected, “some members of the audience could interpret it as non-serious and inaccurate discourse, while some others could seek for ‘hidden’ serious, offensive, aggressive, etc. meanings” (Tsakona, 2017, p.501). However, instances of humor can also be *intended* as aggressive actions that are concealed under a humorous veneer, perhaps in order to gain or retain social control (Norrick, 2010, p. 234 and p. 240). Humor here becomes an either purposeful or self-deceptive strategy, where the sender can detach themselves from the aggressive or serious parts of their utterance and claim that they were joking (Billig, 2005, p. 185 and p. 192).

Humor scholars often divide humor into positive humor and negative humor (cf. Billig, 2005); i.e. humor that has a positive aim and uptake versus humor that has a negative aim and uptake. Positive humor is associated with humor with positive affect

that “builds camaraderie” and leads “to further joking” (Norrick, 2010, p.233). Here, the aim of the sender would be to *align* themselves with their target audience (cf. the notion of audience design; N. Mortensen J. and Coupland, 2018) in order to create a humorous effect. On the other hand, negative humor is associated with humor with negative affect that “can demean, cause suffering to and humiliate its victims” (Billig, 2005, p. 26), in which alignment with a target audience is at least not the only aim of the sender. However, Billig (2005) further argues that the categories of positive and negative humor are not as clear-cut as they may seem, since humor can “result in just a small number of people laughing and with many more people registering outrage” (160).

Some of the genres that are typically considered negative humor are *dark humor* (making “fun of situations usually regarded as tragic, such as death, sickness, disability, and extreme violence”, Bucaria, 2008, pp. 218-219), *killing jokes* (jokes about killing or “dehumanizing the human form”, Lewis, 1997) and *aggressive humor* (jokes where hatred and humor is linked; where “jokers derive pleasure from the expression of aggression against the target”, Billig, 2001, p.268).

Negative humor can have a harmful effect on potential targets, but can also create indignation or fear in some audiences. Different audiences may take up the same utterance in different ways, either because there are different norms and genre expectations in different communities *or* because the utterance has different functions for different audiences. For instance, there is a great difference between being the target of aggressive humor and being a participant in laughing at the target of aggressive humor.

3. Data and Method

The data material for this case study includes:

- a) the original utterance from the Instagram posting (see (1) below),
- b) my transcripts of audio recordings conducted by The Danish Broadcasting Corporation DR during the district court trial, and
- c) police report and district court and high court judgments⁵ (Danish, *dombøger*) from the case (see figure 1 above for an overview of how the different genres are embedded within each other in this threat case).

All names have been pseudonomised, and no information about court dates or court venues is disclosed. The original picture containing the utterance was publicized by different Danish media with the face of the defendant blurred to ensure her anonymity. The defendant’s facial expression in the picture could perhaps have been an indicator of her intent, but the picture itself was not discussed as an important factor, at least not in the parts of the data set that were available to me. As such, no multimodal analysis was conducted, but a formal analysis of the utterance follows in section 4.1.

Parts of the audio recordings from the trial can be accessed through a radio documentary about ‘school threats’ by DR’s radio channel P1 (DR, The Danish Broadcasting Corporation, P1 Documentary). Since I only have access to the parts of the audio recording broadcasted by P1, I cannot account for the remaining parts of the trial. The documentary is in two parts of 26.32 and 26.05 minutes respectively. The audio recording was transcribed in detail using Transcriber and following the transcription conventions

⁵The police and the prosecution’s office granted access (Danish, *aktindsigt*) to the police report and the district court judgment. The high court judgment is accessed through the judicial database Karnov.

of CA (for transcription conventions, see Jefferson, 2004). I only transcribed the segments of the radio documentary that include audio recordings from the trial itself. One witness (whom I have given the pseudonym 'the school representative') did not wish to have her voice recorded by DR, and I therefore did not transcribe any parts of the documentary that contained audio from the prosecutor's or the defense lawyer's questions to this witness. The parts of the audio recording transcribed amount to approximately 15.5 minutes, predominantly from part one of the radio documentary.

The excerpts from the transcription, the police report and the court judgments that are analyzed in this study were selected, because they include uptakes on the defendant's utterance and the justifications for these interpretations. The uptakes can either consist of direct labeling of the genre or descriptions of or metacommentary on the social action that the genre represents and enacts.

My primary data source for this is the transcripts of the audio recordings, since the uptakes of the people involved here are presented verbatim and not in a summarized form (as is the case in the police reports and court judgments). If no data is available in the audio recording on a particular uptake or if additional information about a person's uptake is present in the police report or the court judgments, then excerpts from police report or court judgments are included. Since the data from police report and court judgment consist of written summaries of oral statements, they are of course not verbatim renditions, and as such, it is not possible to know how much they diverge from the original statements. However, they can still be used to give indirect insight into uptakes, since general and central themes (the gist of things) will often be what people recall (e.g. Sachs, 1967). For an overview of the data sources for each of the uptakes on the utterance, see table 1 below:

Table 1. Overview of data sources for uptakes

	Through audio recordings from district court	Through district court and high court judgments	Through police report
The prosecution (P)	Prosecutor's <ul style="list-style-type: none"> • reading of indictment • questioning of defendant and witness teacher • closing statement 	Indictment	-
The defense lawyer (L)	Defense lawyer's: <ul style="list-style-type: none"> • questioning of defendant and witness teacher • closing statement 	-	-
The defendant (D)	Her testimony in district court	Summaries of her testimonies in district court and high court	Summary of her statement to the police (she read it through with her lawyer)
Witness teacher (T)	Her testimony in district court	Summaries of her testimonies in district court and high court	Summary of her statement to the police (it was read aloud to her)
The school representative who reported it to the police	- (She did not wish her testimony to be recorded, but <i>DR PI</i> summarizes it)	Summaries of her testimonies in district court and high court	Summary of her statement to the police (she did not wish to read it through)
The classmate and her mother who reported the IG posting to the school	The prosecutor reads aloud an e-mail from the mother in court + mentioned in D's testimony*	Mentioned in the summaries of the testimonies by D and the school representative*	Mentioned in the summary of the school representative's statement to the police*
The defendant's two friends	Mentioned in D's testimony*	Mentioned in the summaries of the testimonies by D*	Mentioned in the summary of D's statement to the police*
Other classmates	-	Mentioned in the summaries of the testimonies by the school representative*	-
People who responded to the IG post online	One of the posts is read aloud in court + mentioned in D's testimony*	Mentioned in the summaries of the testimonies of D*	-

The data material also exemplifies the complex textual travels that happen within the legal system. This aspect will be considered when relevant, for instance if data material from either police report or court judgments is used instead of audio recordings. However, even in the audio recordings, textual travels are prevalent, for instance whenever anyone offers hearsay evidence (indicated in table 1 by an asterisk *) or every time the purported threatening communication is alluded to or quoted.

4. Analysis

In the following analysis, I first briefly outline the formal aspects of the original utterance that the young female defendant is on trial for writing. Next, I concentrate on the uptakes of the utterance that take it as a threat, and then on those that take it as a joke. Finally, I discuss the uptakes of the district judges and the high court judges in their verdicts.

4.1. The Instagram utterance

The utterance from the Instagram post in (1) below has the form of a direct threat, since both a future, harmful event towards a target and the sender's responsibility for the event happening are explicitly stated (cf. section 2.3 See appendix A for all original data in Danish):

(1) 'I'll be the next school shooter guys, lmao, watch out'

The harmful event against a target is explicitly stated through the noun phrase *næste school shooter* 'next school shooter', i.e. the target is the school and the harmful event is the shooting (the imperative *watch out* also points to a harmful event). The sender's responsibility is made explicit through the agentive 1SG *jeg* 'I' together with the present tense of the copula verb *bliver*, which has future reference and therefore can be translated into English as the modal present with future reference 'will be'.

The abbreviation *lmao* can be seen as a *contextualization cue* (Gumperz, 1982) that is supposed to contravene the propositional content of the utterance⁶, i.e. 'this is not really a threat – it is a joke'. Alternatively, it may be interpreted as an intensifying cue that is supposed to strengthen the propositional content of the utterance, i.e. 'this is a threat – and I am having fun writing it'. Jefferson (1984, 1985) studies of laughter in *oral* conversations show that laughter particles are not only placed at the end of jokes, but also at the end of more serious utterances, sometimes as "signals to the listener that the speaker is aware that [something] might give offence" (from Billig, 2005).

The defendant alternates between Danish and English in the utterance, notably using English mainly to express the proposed harmful event (school shooter, watch out). Such code switching may function as a way of distancing herself from parts of her utterance (Gumperz, 1982), perhaps making the message less offensive or directly threatening. However, the laughter cue *lmao* is also written in English. That makes the intention of the code switching more unclear, since the defendant purportedly is not trying to distance herself from the humorous aspects of her utterance. In addition, seeing that the court case shows that many adult Danes are apparently not familiar with the abbreviation *lmao*, it makes a unified uptake of the utterance as a joke less likely. However, her intended addressees (her peers who are her own age) are more likely to know this type

⁶Thank you to Associate Professor Martha Sif Karrebæk for suggesting the term 'intensifying cue' for this context.

of netspeech, even if it is in English, but since she had an open profile, it is more difficult to control who her audience is.

4.2. The diverging uptakes of the utterance

The following analysis focuses on the different uptakes on the Instagram post as they appear in the audio recordings from the trial and as they are represented in the police report and the court judgment. An overview of these uptakes is shown in table 2 below (cf. table 1 for an overview of the data sources for this analysis), with abbreviations of the parties from the trial that appear in the transcripts:

Strategy	Freq.	Feature in the analysis	Type of feature
Emoticons/emojis ^a	62	Number of emojis ^b	scale
Opening	47	Type of first emoji ^b	categorical ^c
		Full opening	qualitative (string)
		Greeting	qualitative (string)
		Addressee	qualitative (string)
Punctuation	45	Punctuation	qualitative (string)
		Number of full stops	scale
		Number of commas	scale
Word use ^d	43	-	-
Writing style ^d	30	-	-
Sentence/message length	21	Number of separate text messages sent	scale
		Number of sentences	scale
		Number of words	scale
		Average sentence length	scale
Formality ^d	21	-	-
Sentence structure	18	Subject deletion	binary
		Article deletion	binary
Exclamation marks ^e	17	Number of excl. marks	scale
Closing	15	Full closing	qualitative (string)

Table 2. Overview of uptakes (excl. the district court and high court judges' verdicts)

As table 2 above shows, the uptakes of the defendant's Instagram post can be divided into two main categories, namely the ones who take up the utterance as a threat, and the ones who take up the utterance as a joke. The following sections will be organized according to these two distinct types of uptakes.

4.3. Uptake as threat

4.3.1. Uptake as threat by the classmate's mother and the school representative

Two of the most significant uptakes in this case are the uptakes by the classmate's mother and the school representative, in that their combined uptakes are what causes the defendant to be arrested by the police. The classmate's mother writes to the school representative, and the school representative then reports the defendant to the police. In sequence 1 below from the audio recording of the district court trial, the prosecutor reads aloud an e-mail from the classmate's mother to the school representative (see appendix A for the Danish transcription of the original audio recording from the district court trial).

Sequence 1: prosecutor's reading aloud of e-mail from classmate's mother

117 P: subject reg- imPORTant please call A S A P that is as soon as possible at telephone number; regarding a student from {NAME OF CLASS} who went aMOK in CLASS yesterday during a presentation (0.5) she has subsequently posted a: >my story< ((said in English)) (.) where she writes i'll be the next school shooter.

Sequence 1 not only shows the uptake of the classmate's mother, but also demonstrates how genres are embedded within each other through textual travels; the written e-mail is presented orally within the trial (see figure 1 for a general overview of the genre embedding in this case).

This uptake is a more or less immediate response to the utterance (the e-mail was sent the day after the utterance was posted). As such, it is more likely to reflect the immediate and impulsive affect of the people involved, as opposed to uptakes that happen later, which will often be more likely to include metacommentary or genre labels, since such later reactions often serve to evaluate an utterance from a temporal distance⁷.

In her e-mail, the mother does not explicitly label the genre a threat in sequence 1, but instead implicitly uses the incident in class to identify the utterance as dangerous or at least immediately demanding action ('please call A S A P'). In that way, she offers a possible explanation for the defendant having written the Instagram post, namely that she 'went aMOK' (Danish: *gik aMOK*) in class on the same day. In Danish, this phrase means that someone 'gets wild and uncontrollable, often out of anger or excitement and often followed by violence' (my translation of the definition from the Danish Dictionary, DDO). The mother clearly links the posting of the utterance to the incident in class where the defendant flipped a table in anger, thereby treating it as indirect evidence of the utterance being a threat. The sense of urgency in her e-mail suggests that she is worried or scared, but the e-mail does not directly mention fear.

Even though it is not necessary for a threat to have caused *actual* fear in order to be deemed illegal according to § 266 (see section 2.3 on 'fitness to create fear'), the presence of actual fear in threat cases is nonetheless often considered by Danish prosecutors and judges (cf. Bojsen-Møller, 2022). The defendant's utterance had in fact created fear among the students in school, as evidenced in excerpt 2 below from the high court judgment's summary of the testimony given by the school representative who received the e-mail from the classmate's mother. The excerpt refers to a conversation between the school representative and the classmate's mother:

(2) From the high court judgment:

'The mother said to her that her daughter and other students did not dare go to school.'

According to the court judgment's summary, both the daughter and other students in school were scared – in fact, they were so scared that they did not dare go to school the next day. There is no direct labelling of the genre of threats in either sequence 1 or in excerpt (2) above, but the mother's uptake of the utterance is still clearly reflected as that of a threat. Instead of labelling the genre a threat, the perceived (perlocutionary) *effect* of a threat is described, i.e. that threats often cause fear.

⁷However, there are examples of *in-situ* uptakes to threatening utterances where a recipient asks the sender of the utterance 'is this a threat?', i.e. they request that the sender labels their utterance themselves.

Aside from the students at the school, the school representative also clearly took up the Instagram post as a threat. In excerpt (3) below, which is taken from the high court judgments' summary of her testimony in court, she ('the witness') explains her reasons for reporting the defendant to the police:

(3) From the high court judgment:

'The witness' prior knowledge of the defendant and awareness of the incident on the day before were part of her considerations before calling the police. She cannot answer whether she would have called, if it had been a student without this background. Before contacting the police, she looked up the abbreviation »lmao«, which she didn't know in advance, but which means something like »I am laughing my ass off«. However, one cannot know whether the funny part was sending the message or carrying out the act. Since 2013, school shootings have been part of the emergency plan.'

Similar to the classmate's mother, the school representative here clearly links the incident in class and her prior knowledge of the defendant to her reporting the defendant's Instagram post to the police. The incident again serves as indirect evidence or justification of the utterance being a threat. The fact that she reported the defendant to the police for threatening makes her uptake of the utterance as a threat particularly strong. In addition, she takes up on the possible act of school shooting (they are 'part of the emergency plan'), meaning that she considered it a real possibility that the defendant would carry out the threat communicated in her post. School shootings are almost non-existent in Denmark, but they are of course well-known from media reports from other countries, especially the United States. One can assume that a history of school shootings would have made the threatened act appear much more probable.

The school representative explains that she looked up 'lmao', but the meaning of the abbreviation did not change her uptake of the utterance from threat to joke, since she still called the police and reported her for threatening. Thus, she must have interpreted the laughter cue as an intensifying cue rather than a contextualization clue (see further in the section 4.1). Seeing it as an intensifying cue would mean that the defendant was serious about frightening people at school (and perhaps also serious about carrying out her threat), but that she simultaneously found it funny. In all the data available in the case, the school representative is the only one who addresses this double-sided nature of humor: the fact that humor is not always non-serious and positive, but that you can also use humor to disguise a serious and aggressive intent (see section 2.4).

4.3.2. Hesitant uptake as threat by the teacher

The teacher's uptake is more ambiguous than the remaining uptakes presented in this article. However ambivalent, her statements still correlate more with an uptake of the utterance as a threat than as a joke. In sequence 2 below, in a verbatim textual travel from the police report (turns 121 and 125), in which the summary of the teacher's testimony to the police is found, the prosecutor asks the teacher of her impression of the defendant after the incident at school where she got angry and flipped a table. The teacher did not witness the incident in class, but she talked to the defendant immediately after, having been summoned by some of the defendant's classmates. Similarly to the classmate's mother, though more indirectly, the prosecutor attempts to link the intense incident in class to the writing of the utterance later that day, thereby making the defendant's state of mind be evidence of her having threatened.

Sequence 2: from prosecutor's questioning of teacher

- 118 **T:** and she she's sitting like holding (.) herself right?
and she says that she's (0.8) very very SORRY
- 119 **P:** did she say that she was- f-felt put down by someone?
- 120 **T:** YES (0.4)
- 121 **P:** ^ookay^o (.) then I HAVE just one more thing
the defendant SAID that she had had a lot of NASTY thoughts
about that student that she (.) felt had put DOWN the defendant
the defendant said that she didn't want to FOLLOW these thoughts
>but would rather take it out on< an object.
have you [said] that to the police
- 122 **T:** [YES] YES
- 123 **P:** okay. (.)
>when asked how the defendant is as a student
the subject describes her< as a LOVELY girl
uh who very much wants all the RIGHT and GOOD things
- 124 **T:** YES
- 125 **P:** the defendant is a FRAGILE girl who can get angry
if others don't underSTAND what she MEANS (0.3)
the defendant SAYS that she has angry and nast-nasty thoughts (.)
she doesn't want to be someone (.) who has such thoughts
- 126 **T:** YES
- 127 **P:** is that also what uh you've ex[plained]
- 128 **T:** [YES]
- 129 **P:** (0.5) does that ONLY relate to that day?
o[r: (.) how should i understand it?]
- 130 **T:** [hhhhhhhhhhhhhhhhhhhh ((loud sigh))]
(0.9)
well really I haven't had any experiences with her (.)
uh:: (0.3) that is bad experience ON THE CONTRARY (.)
umm: i really haven't (0.5)
- 131 **P:** okay.
(2.9) ((turning of pages))

The teacher is hesitant about making negative statements that can be made to link the incident in class plus the defendant's general disposition to the writing of the utterance. The teacher agrees to having described the defendant both in positive terms as 'a LOVELY girl' (Danish: *en DEJlig pige*) (turn 123) and in terms that highlight more negative aspects of the defendant, such as her having issues with anger and having 'nasty thoughts' (Danish: *grimme tanker*) that she does not want to follow or have (turns 121 and 125). However, when the prosecutor asks her if that is in fact a general pattern for the defendant, the teacher is at first very hesitant about answering the question, as indicated by a very lengthy sigh of 1.3 seconds followed by a 0.9 second pause (turn 130). Afterwards, she vehemently denies having had any other negative experiences with the defendant. In effect, she thereby contradicts part of the statement that she agreed to having given to the police, namely the general statement that the defendant 'can get angry if others don't understand what she MEANS' (Danish: *som ka blive vred hvis andre ikke forstår hva hun MENER*). Since textual travels from oral to written will always involve (more or less acceptable) transformations (Heffer, 2013; Jönsson & Linell, 1991), it is not possible to know whether she did in fact speak of a general tendency during the police interview or whether she was only referring to the incident in class.

Even though the teacher's uptake is more ambivalent than the rest of the uptakes in the case, her descriptions of the defendant and the utterance correlate better with seeing

the utterance as a threat than a joke, as seen in excerpt (4) below from the summary of her testimony to the police:

(4) From the police report:

‘Confronted with the fact that the accused has put up a post on Instagram saying that the accused could become the next school shooter, the subject explained that she knows that the accused does not want to carry out these threats. The subject believes that the accused has not been able to control her thoughts when she wrote it.’

Since excerpt (4) contains a summary of an interaction between the teacher (‘the subject’) and a police officer, it is not possible to know whether the label ‘these threats’ (Danish: *disse trusler*) was produced by the teacher herself during the police interview or whether the label was introduced by the police, either in their questioning or in the report itself. Nevertheless, the teacher evidently did not describe the utterance as humorous, funny or a joke. She instead proposes a reason for the defendant in fact having threatened, namely that she was not able to ‘control her thoughts when she wrote it’ (Danish: *kontrollere sine tanker da hun skrev det*), and she denies that the defendant would ever act on her threats⁸.

The defense lawyer is actually the one who lets this part of the police report travel into the courtroom, as can be seen in sequence 3 below. He focuses on the fact that the teacher at least does not think the defendant capable of acting on her threats (even though that is not a requirement in order to be found guilty of threatening):

Sequence 3: defense lawyer’s questioning of teacher

132 L: you had said to the police that you were sure (0.3)
 that uh (0.8) the defendant wouldn’t carry out her threats
 133 T: YES i [am c-]ABSOLUTELY SURE that she would NEVER do that
 134 L: [()]
 nope

Notice how the defense lawyer curiously presupposes that the utterance in fact was a threat through the genre label ‘her threats’ (Danish: *hendes trusler*), the reason for this probably just being that he is directly quoting the police report. The teacher again does not label the utterance herself, but accepts the presupposition that they are threats, once more only focusing on the fact that the defendant would never ‘carry out her threats’ (Danish: *gøre alvor af sine trusler*). Despite her hesitation and evident warm feelings for the defendant, the teacher’s witness statements – especially her statements to the police that were taken closer to the event itself – clearly indicate that she did not take up the utterance as a potential joke but rather as a potential threat.

4.3.3. Uptake as threat by the prosecution in the indictment

One of the powerful and official uptakes a society can have to an utterance is when it is taken up as an example of a prosecutable, illegal act by a prosecutor (in Denmark by ‘the Prosecution Service’). The indictment is one of the required genres in the criminal trial genre set (see figure 1), which has the function of presenting the case to the court

⁸Actually, the defendant is not on trial for having intent to carry out a threat or even planning to *carry out* a threat (that would have required a much more serious charge than § 266); she is on trial for having *threatened* ‘to carry out an illegal act in a way that is fit to provoke serious fear’.

and the defense. It is part of the genre requirements of an indictment to be exact and precise when relaying the overall accusation of the prosecution. This includes naming the criminal act, the appropriate legislature as well as stating the specific details of the case (Kistrup, Poulsen, Røn, & Rørdam, 2018, pp. 125-127), in this case including the exact verbatim wording of the Instagram utterance and labelling the utterance a threat.

In sequence 4 below from the audio recording of the district court trial, the prosecutor orally presents the written indictment. Sequence 4 not only shows the Prosecution's official uptake of the utterance as a threat but also again demonstrates the intricate genre embedding and the shifts in written to oral mode that happen continuously throughout court cases:

Sequence 4: prosecutor's reading aloud of indictment

001 P: the indictment is dated {DATE} >and reads that<
 {NAME} is charged by the court of {NAME OF CITY}
 seeking punishment for (0.6)
 threats under the criminal code's section twohundredsixtysix (0.7)
 by on the {DATE} around midnight (0.5)
 as a student at {NAME OF CITY} high school (.)
 having threatened to commit a punishable act (0.5)
 in that the defendant put a picture of herself
 on instagram with the TEXT (.)
 i'll be the next school shooter (.) guys (0.3)
 L M A O (0.3) watch out! (.)

In the indictment, the Prosecution's uptake of the utterance as a threat is demonstrated clearly and explicitly through the genre label 'threats under the criminal code's section twohundredsixtysix' (Danish, *trusler efter straffelovens paragraf tohundredogsekstretes*). It is also demonstrated through the more detailed genre description 'having threatened to commit a punishable act' (Danish: *at ha truet med at foretage en strafbar handling*), which is close to a verbatim textual travel from the Danish threat statute (§ 266). The indictment not only formalizes and categorizes the utterance and embeds it within the legal system; it also serves as an official accusation that is inherently present throughout the case.

4.3.4. Uptake as threat in the prosecutor's cross-examination of the defendant

Since the indictment introduces the overall accusation to the court and to the defendant, the accusation is not mentioned explicitly during the prosecutor's cross-examination of the defendant (at least not in any of the parts on the recording published by *DR P1*). This means that many of the defendant's answers to the prosecutor's questions include accounts wherein she attempts to deny the overall accusation that is made explicit by the prosecutor's reading of the indictment in the beginning of the trial.

In the examinations and cross-examinations of witnesses and defendants, Danish prosecutors have a duty to adhere to 'the principle of objectivity' (Danish: *objektivitet-sprincippet*) (Toftgaard Nielsen, 2016, p. 44). They are therefore obligated to seek out evidence that point away from as well as towards from the defendant's guilt (the same applies to the Danish Police). Even though Danish courtrooms have "an adversarial participation structure similar to that of the Anglo-American courtroom" (S. S. Mortensen & Mortensen, 2017, p.404), Danish courtroom interaction is often described as less confrontational and more casual and informal than the classic adversarial courtroom

interaction and even than the Continental inquisitorial system (S. S. Mortensen & Mortensen, 2017; cf. Anderson, 1992).

The Danish courtroom interaction is instead more indirect and subtle, as can also be seen in sequence 5 below. In this sequence, the prosecutor asks the defendant about the intense incident in class where she flipped a table, still as an indirect attempt to link the incident to the writing of the utterance later that day:

Sequence 5: prosecutor's cross-examination of defendant

- 004 P: and and did something happen then? (0.7)
 005 D: umm i got a bit ANGRY. (0.3) and flipped a table, (0.3)
 006 P: °yes°, what were you angry about?
 007 D: umm >I had got a bit< angry about >what's it called< (0.3)
 umm {NAME} who (0.6) eh came with proVoking °like°
 (1.0)
 references to me
 (1.4)
 008 P: was it something about a presentatio[n o]r:?
 009 D: [YES]
 010 P: °yes° try and tell me what it was=
 011 D: =well i was about to present and we had also rehearsed it also
 and then he took what i was supposed to say (0.8)
 and afterwards afterwards then he said
 >i'm just waiting for her to open her MOUTH< (0.6)
 aloud in class ((voice trembling))
 umm
 (1.4)
 which e- made >me a bit< irritated
 012 P: were you practically humiliated on that occa[sion]?
 013 D: [YES]
 014 P: yes yeah° where- was it in front of the WHOLE class and the teacher?
 015 D: uh-huh (0.8)
 016 P: and how did you react to it. (0.7)
 017 D: yeah that was- the:n then i flipped the table
 018 P: (0.3) °yes° were you also umm: (0.5) otherwise (0.5)
 reactive towards {NAME} or [>towards one of the others<]
 019 D: [NO: not at all]
 I quickly exited afterwards

The prosecutor does not explicitly label the utterance a threat, but he treats the incident in class as indirect evidence of her intent to threaten, since she acted negatively and violently on the day that she wrote the post.

In her account, in sequence 5 above, the defendant downplays the seriousness surrounding that incident, as can be seen by the mitigating modifier 'a bit' (Danish, *lidt* in front of the adjectives 'angry' (Danish: *sur*⁹ and 'irritated' (Danish, *irriteret*) (turns 005, 007 and 011). The word choice and choice of modifier in her account conveys that the situation in class was a minor and non-consequential dispute. In this way, she indirectly rejects that the incident in class was the reason for her posting the picture on Instagram later that day.

In comparison, the prosecutor intensifies the sense of seriousness of the incident. For instance, he uses the adjective 'angry' without any mitigating modifier in response to her 'a bit angry' (turn 006 and 005). Additionally, when asking about the defendant's state of mind during the incident, he uses the adjective phrase 'practically humiliated'

⁹The Danish adjective *sur* has a very broad meaning potential, varying from 'cross' to 'angry'.

(Danish: *nærmest ydmyget*) (turn 012) after she describes herself as feeling ‘a bit irritated’. He is thereby introducing a much stronger negative emotion than irritation, one that may feasibly correspond more easily with her having intended to threaten later that day. Furthermore, the prosecutor introduces the adjective phrase ‘otherwise reactive’ (Danish, *i øvrigt reagerende*) (turn 018) when he asks her if she did anything else to her classmates other than flip the table. He is actually referring to the specific situation in class; whether she did anything else in class except flipping the table, but the phrase ‘otherwise’ (*i øvrigt*) could also refer to other situations where she reacted negatively towards her classmates. The inference being that perhaps the Instagram she posted later that day was in fact *intended* as a negative reaction towards her classmates and the school—or in other words, she intended to threaten the school. Furthermore, the fact that he uses the adjectival predicative ‘reactive’ (in Danish, the present participle *reagerende*) instead of the verb ‘react’ (e.g. as in ‘did you react in any other way towards your classmates’) invites an additional inference. The use of the present participle in Danish is used to refer to an ongoing action or a continuous state. In this context, this would invite the inference that instead of the defendant being a person who simply reacted in a specific way in a singular, specific situation, the defendant may be a person who tends to react in this manner in general. Therefore, she could plausibly be viewed by the school as an actual threat—a person who is fit to provoke fear. The police report does in fact mention previous negative incidents concerning the defendant that could indicate a pattern, which the high court judgment also alludes to in excerpt 3 above. In excerpt 3, the school representative is referenced saying that she partly based her decision to report the defendant to the police on ‘prior knowledge of the defendant’. The defendant fervently denies the idea of her having been ‘otherwise reactive’, her answer even overlapping with his question, which does not happen often throughout the audio recording. No doubt, it is important for the defense’s case that she be viewed as a non-reactive, non-violent girl who could not foresee the consequences of her Instagram post, and not a reactive, violent adult who should know the possible consequences of posting such an utterance on an open online platform.

In sequence 6¹⁰ below, the prosecutor asks whether the defendant has experienced similar incidents to the one in class earlier. The link between the intense incident in class and her possibly having a general destructive disposition is made explicit in this sequence (as opposed to the more indirect insinuations in sequence 5 above):

Sequence 6: prosecutor’s cross-examination of defendant

060 P: this outburst uh that you had
have you HAD more of those.
(0.7)

061 D: no::

062 P: can you- uh: can you say that you sometimes get immensely angry (.)

063 D: ye:s i can get a bit angry

¹⁰The Danish dialogic particle *jo*, which signals (that there should be) consensus between speaker and audience, has no direct English equivalent, but can be translated into ‘of course’ or ‘as you know’.

- 064 **P:** yes (.) and of course {Danish: *jɔ*¹} what I am interested in uh that is of course {Danish: *jɔ*} kinda how the others perceive you and ho-how YOU somehow maybe KNOW that the others perceive you (0.7) uh: and yes according to count ONE appendix seven it's a police report (0.7) is it tru- can you remember that you've spoken with the police?
- 065 **D:** YES
- 066 **P:** yes umm it says on page four in the report the subject then explained that the subject has PROBLEMS WITH HER MOOD (1.0) the subject can get immensely ANGRY when she gets angry (1.0) can you uh: have you said anything like that to the police?
- 067 **D:** yes (.)
- 068 **P:** °yes° is that also how it was. (0.5)
- 069 **D:** yes
- 070 **P:** yes ((sharp voice)) >has there been other incidents where you got angry in class< and perhaps could have [<been a little>] (0.4) [enraged?]
- 071 **D:** [NO:::] [nonono<] i haven't been angry in class at all or anything (.)
- 072 **P:** NO (.) so this was the first time this
- 073 **D:** yes it w- (0.5)
- 074 **P:** >THANK YOU fn-for now i have no further questions for the defendant<

In sequence 6 above, the prosecutor continues to focus on the defendant's state of mind that day and on her overall disposition, in an indirect attempt to depict her as a person who is not able to control her emotions and therefore capable of threatening.

By taking up the defendant's "own" words from the police report, the prosecutor is attempting to render it probable that the situation surrounding the Instagram post was not consistent with making a joke but rather consistent with making a threat. In fact, the defendant was 'angry', 'humiliated' and 'reactive' that day, and in addition, she is by her own admission generally a person who has 'PROBLEMS WITH HER MOOD' (Danish: *PROBLEMER MED SIT HUMØR*) and who can get 'immensely ANGRY' (Danish: *vældig SUR*) (turn 066). Both of these phrases are verbatim textual travels from the defendant's statement as reproduced in the police report.¹¹

In a rhetorical question, the prosecutor addresses the question of how much the defendant knows or understands about how others perceive her and her actions (turn 064) – in fact, her knowledge or understanding of the probable or possible generic up-takes of her utterance. He continues by reading aloud her own acknowledgement to the police of her having what may be seen as an unstable and aggressive disposition. He implies that if *she* sees herself in that light, then she must also have understood that *others* would perceive her as being capable of either threatening someone or of carrying out a violent threat. The prosecutor is actually, though rather indirectly, addressing the

¹¹Notice how the prosecutor both uses the phrase 'immensely angry' (Danish: *vældig sur*) as a direct quote in turn 066, and also as a hidden quote in 062. The Danish adverb of degree *vældig* is an unusual word to use for a contemporary Danish teenager, especially when produced orally, as it would have been in a police interview ('immensely' has been chosen as an English correlation to that). In Denmark, oral police interviews consist of dialogues between police officers and suspects (or witnesses), but written police reports are recontextualized into monologues (Jönsson & Linell, 1991; cf. M. Komter (2019) for similar recontextualization techniques in the Netherlands). As such, it is not possible to know whether this specific word was used by the defendant herself at the oral police interview or whether it was produced by the police officer who wrote the police report.

question of intent. If a defendant is believed to have either known, found it *probable* or found it *possible* that others would perceive their utterance as a threat ‘fit to cause fear’ then intent would be proven (see section 2.3 on the different levels of criminal intent). In other words, the indirectly stated argument is that she must have understood that her posting the utterance shortly after the incident in class would mean that some people would take it up as a threat.

4.3.5. Uptake as threat in the prosecutor’s closing statement

The prosecutor sums up the findings presented during the case and offers the Prosecution’s recommendations to the judges during the prosecutor’s ‘closing statement’ (Danish: *procedure*) (Toftgaard Nielsen, 2016, pp. 198-199). The prosecutor still has to adhere to the principle of objectivity here, but a prosecutor’s “closing statement is in practice shaped by his point of view”¹² (Toftgaard Nielsen, 2016, pp. 198). Seeing that this trial genre allows for the prosecutor to take a more direct and somewhat stronger stance, the prosecution’s uptake of the utterance as a threat culminate in the closing statement (see sequence 7 below):

Sequence 7: from prosecutor’s closing statement

135 P: in class was a violent incident.
 we also know (0.5) that she threw or flipped a table in ra:ge
 or in (.) frustration
 and that she (0.4) was in a very BAD state (0.6)
 and that was the LAST the w-uh persons in the CLASSroom saw of her
 (0.3) and some of these persons (0.5)
 she must of course {Danish: jo} to some degree have realized (.)
 would probably go and look
 when fourty FIFTy would normally look at her posts (0.6)
 and that she then on that same night (0.3) put up a post where she wrote
 i'll {said in English} (.) uh i'll be the next <school shooter> guys
 (0.4) we also know that uh the teacher was
 familiar with her having NASTY thoughts
 she also had a FEAR that these NASTy thoughts would be REAlized (0.5)
 and because of that she would AIM it against objects
 (1.3)
 it is ACTUALLY WELL-KNOWN that uh school shootings HAPPEN other places
 (0.3) we all know that (0.6)
 but tha:t the reason that i say that is NOT
 to uh accuse uh: the defendant OF wanting to DO this here (0.5)
 as such i don't think she would do that EITHER
 that's not deCIasive for the conviction either (.)
 in reality there is of course {Danish: jo} actually no NEED
 in order for uh the defendant to be convicted here
 for anyone a:ctually to be scared
 it only needs to be FIT to creATE this serious fear
 for life health or welfare (0.7)
 but there was (0.6)
 also {THE SCHOOL REPRESENTATIVE} took it quite seriously
 >and before we knew it< (0.3)
 the police had arrived (0.9)
 can one then (.) do away with all this by writing L M A O (0.8)
 can one be TOTALLY sure that everyone is thinking
 this is just for fun? (0.3)
 yes there ARE some (0.6) uh: who have <another> (0.6)
 insider uh: immediate understanding of THIS (.)
 when they write something to each other
 but she hadn't limited it to them (0.9)
 when all that's (.) said

¹²Original quote (my translation): "I praksis er hans procedure dog naturligt præget af hans synsvinkel."

i'm really trying to sum this UP and recommend
 that the pro- that the court (0.7) establishes
 that she is to be convicted in accordance with the indictment (.)
 and found guilty of violation of section twohundredandsixty
 six (1.3)
 this is really something that AFFECTS a lot of people (.)
 it is something that has had a BIG effect on the SCHOOL (0.5)
 and it is also somethi:ng that has a SIGNalling effect (0.3)
 uh that which (0.4) uh becomes the reAction from the court toDAY
 (1.4)

The prosecutor again addresses the defendant's negative state of mind during the incident in class, now using stronger terms ('violent incident', 'rage'), and concludes in unambiguous terms that she 'was in a very BAD state' (Danish: *havde det rigtig DÅRLigt*). He links her bad state to the question of criminal intent to threaten by concluding that the defendant must 'of course to some degree have realized' (Danish: *jo i et eller andet omfang have indset*) that her classmate's would get scared when she wrote the purported threatening communication shortly after having been in a self-admitted bad state. In other words, she must have had some degree of criminal intent to threaten (see section 2.3 on criminal intent).

The prosecutor addresses the question of intent to *carry out* the threat through a textual travel from the teacher's statement to the police (see sequence 2, turn 121). The teacher said that the defendant did not want to act on her negative thoughts, which the prosecutor here interprets as the defendant having 'a FEAR that these NASTy thoughts would be REalized' (Danish: *en FRYGT for at de GRImme tanker de blev OMsat*). This particular textual travel is thus influenced by the stance of the prosecutor, since not wanting to act and fearing to *act* are not the same – the latter could potentially be much more dangerous. The prosecutor concedes that he himself does not really think that the defendant had intent to *carry out* a school shooting (this is probably a result of his adherence to the principle of objectivity). However, he emphasizes that such intent to carry out the threat is not 'decisive' (Danish: *afgørende*) in order to be found guilty of threatening according to § 266, only intent to threaten to carry it out.¹³

Even actual fear is not required according to § 266, but the prosecutor emphasizes that people at the school were in fact scared and 'took it quite seriously' (Danish: *tog det ganske alvorligt*). He uses the fact that many people were very affected and that it had a 'BIG effect on the SCHOOL' (Danish: *STOR effekt på SKOlen*) as evidence of the utterance having been a threat that was 'fit to provoke serious fear'. He never actually labels the utterance 'a threat' but simply refers to § 266 and explains the legal requirements that this statute imposes.

Similarly to the school representative's uptake (see section 4.3.1), the prosecutor does not take up 'lmao' as an instantly recognizable contextualization cue that shifts the meaning of the utterance from threat to joke. He argues that the defendant could not be sure that other witnesses to the communication would not see it as a threat simply because she wrote 'lmao', even though there were people on Instagram who had

¹³DR P1 who published the audio recording apparently found it distressing that the prosecutor did not think the defendant would have carried out her threat but still wanted her to be found guilty. At least, they play sad music throughout the remainder of the prosecutor's closing statement from the moment he says 'as such i don't think she would do that either' (Danish: *det tror jeg sådn set heller ik hun ville gøre*).

an ‘insider’ (Danish: *intern*) understanding of the utterance as a joke. The uptakes of these people with ‘insider’ understandings and the uptake of the defendant herself will be the focus of the following sections.

4.4. Uptake as joke

4.4.1. Uptake as joke by the defendant and her friends

The defendant’s uptake of her own utterance is reflected in the accounts she gives to the police, the prosecutor and the defense lawyer (see section 2.2 on accounts). The defendant’s uptake is of course different from the uptakes of all other people involved in the case, seeing that she herself produced and posted the utterance. Her uptake therefore becomes a self-evaluative action, but not one that necessarily requires frankness or truthfulness.¹⁴ However, it is not relevant to this article to attempt to determine whether the defendant’s uptake in her accounts reflects her *actual* intent in writing the utterance or not. If she indeed intended to threaten the school, admitting that would earn her a prison sentence, and as such, the stakes are much higher for the defendant than for any of the witnesses or legal representatives.

In sequence 8 below, in an attempt to disassociate the incident in class from the writing of the Instagram post, the defendant explains to the prosecutor how her state of mind changed after exiting the school:

Sequence 8: defendant’s account to prosecutor of her state of mind after school

025 **D:** yes she then drove me home and said THAT i could (0.3)
 start (0.4) over again tomorrow
 i felt much better after having talked to {TEACHER}† (.)
 and then >i then chose to be with< a couple of friends (.)
 after it happened (0.4)

026 **P:** mm

027 **D:** >where we then just made lasagna and enjoyed ourselves
 and i thought that it was over< (0.6)

028 **P:** yeah did you talk about anything specific? (0.5)

029 **D:** we had talked about many things and we had also talked a bit about
 li:ke school shootings and things like that (.)
 in the other countries (0.4)

The defendant distances herself from the negative parts of the intense incident in school. She instead focuses on the redeeming aftermath of the incident, thus placing emphasis on the positive aspects of her day instead of the negative situation in class. This focus on the shift to positivity can be seen through phrases like ‘start over again’ (Danish: *starte forfra igen*) and ‘I felt much better’ (Danish: *jeg fik det meget bedre*) (turn 025) and ‘[we] enjoyed ourselves and I thought that it was over’ (Danish: *[vi] hyggede os og jeg tænkte at det var ovre*) (turn 027). She focuses on the positivity and camaraderie she had later that day with her two friends, which means that she uses that more positive context as indirect evidence of her having meant the utterance as a joke instead of the negative context in class. She also distances herself from the act of school shooting by referring to it as something that happens ‘in the other countries’ (Danish: *i de andre lande*) (turn 029)—not in Denmark.

The prosecutor has placed great emphasis on the defendant’s negative disposition and negative state of mind (see sequences 5 and 6) in order to argue that she had intent

¹⁴Defendants are not obliged to speak the truth, according to Danish law. Only witnesses are instructed to speak the truth under penalty of law (Kistrup et al., 2018, p. 657).

to threaten or at least would have realized that others would perceive her utterance as a threat. Instead, she attempts to render it probable that the utterance was meant as a joke and could not plausibly be taken up as a threat to commit something as foreign as a school shooting. However, here the prosecutor could have argued that her statement regarding her talk with her friends about school shootings could be seen as evidence that she had in fact had the actual act of school shooting in mind on the same day that she posted the Instagram post. In that way, she could thereby be viewed as having been aware that her utterance could be taken up as a threat to commit such a school shooting, which would constitute direct intent to threaten according to Danish law (see section 2.3 on the different degrees of criminal intent).

In sequence 9 below, the defendant gives an account of why she posted the picture containing the purported threatening communication:

Sequence 9: defendant's account to prosecutor of posting picture

043 D: just like for FUN ((forceful voice))
 044 P: [mm]
 045 D: [and] >we had taken pictures together and things like that< (0.4)
 umm and then i went home and thought that i would post a picture (0.3)
 046 P: [mm]
 047 D: [then] i took that picture where i sat on the toilet
 048 P: mm
 049 D: and so then i wrote and then-th >i had to find something to write<
 and that became it then.
 (1.2)
 050 P: what did you mean by it?
 (1.0)
 051 D: i meant like, it was something li:ke (0.7)
 especially for THOSE two basically (.) ⁰that⁰ (0.3)
 >they would be able to understand it because we just talked about it<
 (.) so it was like <a joke> (0.5)
 052 P: so when you <write> >i'll be the< next <school shooter guys>
 lmao! watch out.
 053 D: yes (0.3)
 054 P: just so we all understand it
 [what] does lmao stand for
 055 D: [yes] uh laughing my ass off ((said in English))
 056 P: yes umm >laughing my ass off< ((said in Danish))
 057 D: [YES EXACTLY]
 058 P: [yes yes] (0.3)
 059 D: which makes it a joke.

The indictment (see sequence 4 further above) still serves as an implicit accusation of illegal and untoward behavior for which the defendant gives her account, for instance in sequence 9 above. The defendant's uptake of her own utterance as a joke involves references to intent to be humorous and lack of intent to threaten, in other words she mainly uses the subtype of excuse 'appeals to intent' in her account (see section 2.2 on accounts). The defendant describes her intention behind posting the picture in phrases that highlight the non-seriousness and lightheartedness of her action. This can for instance be seen in phrases like 'just like for FUN' (Danish: *bare sån for SJOVT*) (turn 043), '[i] thought that i would post a picture' (Danish: *jeg tænkt at jeg ville lægge et billed op*) (turn 045) and 'i had to find something to write and that became it then' (Danish: *så sku jeg find på noget at skrive og det blev så det*) (turn 049). The implication seems to be that this was just something she wrote in the spur of the moment, nothing of any consequence – i.e. it was only a joke. She directly labels the utterance 'a joke' (Danish: *en joke*) (turns 051 and 059), which is in contrast to the absence of such direct labelling of the utterance as a 'threat' by the prosecutor (except in the indictment).

In turn 051, the defendant identifies her desired audience, namely the two friends that she was having fun with after school on the day that she wrote the utterance, making lasagna and talking about school shootings. According to the defendant, her two friends would also take up the utterance as a joke, which was in agreement with her intention behind writing it: ‘they would be able to understand it because we just talked about it’ (Danish: *de ville ku forstå det fordi vi lige havde snakket om det*) – ‘it’ referring to school shootings. She uses their (purported) uptake as indirect evidence of her own account being true. In that light, ‘lmao’ would then have been written in order for the defendant to align herself with her target audience (see section 2.4), in an attempt to procure an uptake of the utterance as a joke. In sequence 10 below (turn 076), the defendant agrees with the defense lawyer that she and her friends have a particular sense of humor, namely a ‘DARK sense of humor’ (Danish: *SORT humor*):

Sequence 10: Defense lawyer’s questioning of defendant

075 L: you ha- you:- >i think you’ve explained<
that they could have a DARK sense of humor? (.)

076 D: yes (0.5)

077 L: uh try and explain who- what could that wha-what can THAT BE

078 D: uh we can write things like that uh
we will kill ourselves the next day (.)
and then we write lol or l m a o (.) aft-afterwards (.)
as if like >of course we won’t do that< it’s a joke, (0.3)

079 L: yeah
(1.6)
SO when you wri- when you express L M A O (0.7)
uh you then mean what (0.6)

080 D: that it’s a JOKE↑

081 L: that it’s a joke.

082 D: yes

083 L: i-is that something everyone underSTANDS? (.)
i [me:an]

084 D: [all] teenagers and (0.6)
<people> who are online un- should understand it,

085 L: basically all those (.) who who see (0.3) [your] posts (0.4)

086 D: [yes] yes

087 L: they know what it me[ans]

088 D: [yes]

The defendant goes from explaining this dark sense of humor in an abstract way to referring to the specific utterance that she wrote. She does this by first using a fictitious example ‘we can write things like that uh we will kill ourselves the next day’ (Danish: *vi kan skrive ting som om at øh vi vil dræbe os selv næste dag*) to then including the particular abbreviation ‘lmao’ in combination with the similar netronym ‘lol’ (i.e. laughing out loud) (turn 078). In that way, the purported threatening communication is framed as an acceptable and ordinary genre, at least within the small discourse community (Swales, 2016) of her and her two friends. The abbreviation ‘lmao’ would thus have been used as a contextualization cue (see section 4.1), which was to signal to the audience that the utterance is not to be taken seriously. In her own words, ‘lmao’ means: ‘of course we won’t do that it’s a joke’ (Danish: *det gør vi selvfølgelig ikke det en joke*) (turn 078).

The defense lawyer moves away from her opinion on the more localized uptakes of her friends to asking whether everyone in general would understand ‘lmao’ as an indication of a joke (turn 083). In her answer, the defendant expands the small ‘insider’ discourse community of three friends to instead being a large ‘insider’ discourse community of ‘all teenagers’ (Danish: *alle teenagere*) and ‘people who are online’ (Danish:

folk som er på nettet) (turn 084). They should all understand it – and therefore the entire utterance – as a joke. Through this line of questioning, the defense lawyer attempts to render it probable that the defendant was certain that people would only take this up as a joke, which would mean that she did not have intent to threaten (see section 2.3). The genre label ‘joke’ is repeated several times (turns 078, 080 and 081), making this uptake particularly strong and explicit.

In sequence 11 below, the defense lawyer and the defendant discuss the actual and immediate uptakes of people online. One of the immediate responses to the Instagram post is read aloud by the defense lawyer in court (again displaying the abundance of textual travels and genre embedding that happen in court cases):

Sequence 11: Defense lawyer’s questioning of defendant

- 100 L: you GET some reACTions on this post isn’t that true,
 101 D: yeah (0.4)
 102 L: uh try and just try and explain what it- what it says THERE
 103 D: yes he has written can you school shoot at my school then? x d (0.5)
 and xd that is like writing (0.5)
 uh it’s like a smiley (.) that’s laughing (0.7)
 so tha:t p- is part o:f forexample lol and these things (.)
 104 L: a-and when HE says that
 can you school shoot at MY school [also] X D SMiley
 105 D: [YES] YES
 106 L: >what THEN does he MEAN by THAT. he doesn’t mean it seri[ously]<
 107 D: [no]
 then it’s just like for fun
 108 L: >for fun he says it< (0.6)
 you said before that (.) there were also OTHERs who
 <wrote like you< on THESE >medias<
 109 D: yes (0.5) yes <lots of other> you know on FACEbook
 where in groups- where they write that they will DO: the sa:me
 and ^othings like that^o
 110 L: for fun.
 111 D: for fun.
 112 L: you have explained to the poLIce (0.4)
 you are uh Cited saying
 that it was supposed to be understood as <a> JOKE
 or a s- or sarcastically (0.4)
 >and then you write< (0.3) >or then you say< (0.3)
 all who know the subject (.) knows that the subject
 >would never THINK of< <HARming ANYone>
 113 D: yes
 (1.2)

The defense lawyer and the defendant co-construct the argument that people on Instagram took up her utterance as a joke and that it is a normal way of making a joke online (e.g. ‘there were also others who wrote like you on these medias’, turn 108). In other words, they argue that this is an existing genre online that is generally accepted within particular communities – and therefore not seen as an illicit genre among them. One of the responses to her utterance ‘can you school shoot at MY school then? x d’ (Danish: *ka du skoleskyde på MIN skole så? x d*) (turn 103) was exactly like her own utterance ‘just like for fun’ (Danish: *bar sådant for sjovt*) (turn 107). The defendant links the use of the symbol of a laughing smiley face ‘XD’ in the response to her own use of the abbreviation ‘lmao’ – they both mean this as a joke. The defense lawyer and the defendant use the immediate uptakes of others online (as well as references to similar

conduct by others online) as indirect evidence that her utterance was in fact a joke. Thus, they portray her utterance in combination with the online responses as an example of positive humor that seeks to build affinity and camaraderie in a community and that leads to “further joking” (Norrick, 2010, p.233).

The defense lawyer attempts to strengthen the defendant’s claims of innocence by referencing her “own” words to the police that everyone knows she would never think of ‘HARming ANYone’ (Danish: *GØRE NOgen forTRÆD*) (turn 112), again focusing on the fact that she would never carry out a threat. He does this through a textual travel from the written police report, which causes him some trouble to reference as an originally oral testimony, again underlining the complexity of the textual travels in this genre chain. In sequence 12 below, as her final words to the judges, the defendant vehemently negates that she meant the utterance as a threat:

Sequence 12: Defendant’s final word before deliberation

140 D: umm uh i would just say that i never meant it as a THREAT
(0.5) and that i’m of course SORRY about ^owhat has happened^o

In that way, her final words become a final and clear repudiation of the genre label stated in the indictment.

4.4.2. Uptake as joke in the defense lawyer’s closing statement

Unlike the prosecutor, the defense lawyer should not adhere to the principle of objectivity, but is instead obligated to plead the defendant’s case to the court (Toftegaard Nielsen, 2016, p. 198). The defense lawyer’s uptake will therefore naturally emulate the defendant’s uptake closely (unless this would harm the defendant rather than help them). In an even more direct sense than what we see in the prosecutor’s closing statement (see section 4.3.5), the defense lawyer’s closing statement in sequence 13 below shows the stance of the defense and the uptake of the utterance as a joke most clearly and strongly:

Sequence 13: From defense lawyer’s closing statement

136 L: >the a- the defendant still pleads< not guilty (0 .6)
and i umm am of course {Danish: jo} in a NUMBER of (0.3)
esSEntial points in DISAGREEMENT with the prosecutor
uh- one has to interpret <the formuLation itself> (.)
AND (.) >one has to< take into considerAtion (.) in what CONtext (.)
it is stated. (0.4)
it is STATED in- on social media (.) on instagram (0.4)
it ISn’t STATED (0.9) towards the school† (0.3)
it is stated (.) uh in a forum (0.3)
where they SPEAK to each other and use a:n- a (0.3)
>yeah well almost< for: those who DON’T know it (0.7)
which apparently is {THE SCHOOL REPRESENTATIVE} (0.5)
they then almost SPEAK in a CODED LANGUAGE† (0.7)
and if you’re NOT <CAPABLE> (0.3)
of uh- of reading and underSTANDING that <conversational form↓> (0.6)
then of course {Danish: jo} you’re not really capable of ASSESSING
whether we’re talking↓ about a THREAT† or a joke. (0.3)
>and then you can say< it can be a: STUpid joke
(1.3)

it can be a <RECKLESS> joke (0.6)
 >you can't convict that after twosixtysix< (0.6)
 BUT but you have to (0.3) underSTAND (0.4)
 what L M A O MEANS in this <CONtEXT> (0.4) it is s- uh stated↑ within
 to the (0.4) uh young people to the (.) TEENAGERS
 who ALL↓ communicate in the same way (0.3)
 and when they use that EXPRESSION (.)
 and it's of course {Danish: jo} important the court UNDERSTANDS this
 (0.6) then it MEANS (0.6)
 i'll be the next <SCHOOL shooter> (0.8) just for fun (0.8)
 and it's clear (.) >that if you don't< UNDERSTAND this. (0.7)
 then of course {Danish: jo} it can be a PROblem↑ (0.5)
 but that is of course {Danish: jo} not (0.5) your honors↓ (.)
 in (.) the criminal code's section tohundredsixtySIX' san-sense (.)
 it's not supposed to be <held against> the DEFENDANT
 (1.0)
 because the SECOND article (0.4) in TWOsixtySIX is (0.7)
 that the DEFENDANT has had inTENT (0.6) to proDUCE a threat
 (1.1)
 a-and that's not an article that can be inflected (0.6)
 one can't REALIZE AFTERwards that this was too STUPID (.)
 or (0.3) <this could have been> something or other
 or SOMEONE could PERHAPS understand it-
 (1.1)

The defense lawyer shifts the attention from the context of the intense incident in school (which the prosecutor focused on in his closing statement) to the immediate context of 'the formulation itself' (Danish: *selve den formulering*), namely that it was stated on social media and not sent directly to the school.

He creates a sharp divide between insider people who understand 'that conversational form' (Danish: *den samtaleform*) and outsider people who do not understand it and for whom it becomes a 'CODED LANGUAGE' (Danish: *kodesprog*). In what may be seen as an *ad hominem* argument (Walton, 2000, p. 10), he singles out the school representative as one of the people who do not understand the 'coded language'. According to the defense lawyer, she is therefore 'of course not really capable of assessing whether we're talking about a THREAT or a joke' (Danish: *jo reelt ikke i stand til at vurdere om der er tale om en TRUSSEL eller en spøg*). Young people and teenagers in fact 'ALL communicate in the same way' (Danish: *kommunikerer [ALLEsammen] på samme måde*), and it is not the defendant's fault if other people take it up in a different manner. As a result of this argument, the judges would then have to become one of the outsiders who do not understand, if they are to find the defendant guilty.

The defense lawyer emphasizes that the defendant meant that she would be the next school shooter 'just for fun' (Danish: *bare for sjov*), i.e. as a joke. According to him, that would completely negate an intent to threaten, since she cannot be found guilty of producing a 'STUpid joke' (Danish: *ÅNDssvag spøg*) or a 'RECKLESS joke' (Danish: *uagtsom spøg*). If he had changed the genre label into 'reckless threat', then that would possibly have been precarious for the defendant, since the seldom used degree of intent *Dolus Eventualis* is sometimes interpreted as a form of 'conscious recklessness' (see section 2.3).¹⁵ However, besides saying that the school representative would not be capable

¹⁵The defense lawyer refers to the 'SECOND article' (Danish: *ANDET led*) of § 266 and claims that it explicitly references 'intent'. He is referring to § 266b (known as the 'racism section'), which (unlike §

of assessing whether the utterance was a threat or a joke, the defense lawyer seems very careful not to utter the label ‘threat’ in his closing statement. This strategy is especially clear at the end of the closing statement, where he says that it is not enough to find someone guilty if they first realized afterwards that ‘this could have been something or other or SOMEONE could PERHAPS understand it-’ (Danish: *det ku godt være etellerandet eller NOGEN ku MÅSKE forstå det-*). Instead of saying ‘threat’, he uses the indefinite pronominal phrase ‘something or other’ (Danish: *et eller andet*). He also cuts himself off before specifying what precisely someone might understand the utterance as, namely a threat. On the other hand, he repeats the genre label ‘joke’ three times, conceivably to emphasize the strength of that particular uptake in an attempt to persuade the judges.

4.5. And the verdict is

The verdict offers the most important uptake of the case, as it proclaims the outcome of the case. In their verdict, the district court judges emphasize lack of intent as the deciding factor in their decision to acquit the defendant (see sequence 14 below):

Sequence 14: from reading of verdict

141 J: and i can say to you that uh we agree that you are to be acquitted
(0.9) no:t because (0.3) we are not talking about an (0.8)
objectively seen illegal threat but because we believe
that uh you didn’t have the required intent. uh: (0.5)

In effect, the judges’ verdict combines the two different uptakes presented in the case. On the one hand, they label the utterance an ‘objectively seen illegal threat’ (Danish: *objektivt set strafbar trussel*), using the double negation ‘not because we’re not talking about’ (Danish: *ikke fordi der ikke er tale om*). On the other hand, they do not believe the defendant had ‘the required intent’ (Danish: *det fornødne forsæt*) to threaten. The summary of the district court verdict in excerpt (5) below from the high court judgment includes more details from the judges’ verdict that further underline the duality of the judges’ uptake:

(5) From summary of district court verdict in high court judgment:

‘The City Court stated that the post was fit to provoke fear for life or health, cf. the Criminal Code’s § 266, even if it was not meant literally or seriously and included the slang phrase »lmao«. The court did however find that D [the defendant] meant the post as a joke that would be read by her friends, and that these would also perceive the post as a joke. The court therefore found that the defendant had not understood that the post would be perceived as a threat, and that consequently she had not had intent to violate the Criminal Code’s § 266.’

The utterance is deemed ‘fit to provoke fear’ (Danish: *egnet til at fremkalde frygt*) as stated in § 266, and the laughter cue ‘lmao’ does not in itself change this fact. Aside from that, the summary does not offer any additional information about the judges’ stance on how ‘lmao’ should be understood – either as an indicator of a joke or as an intensifier

266) does mention intent explicitly. However, §266 and not § 266b is the statute that the defendant is indicted under (see section 4.3.3). Nevertheless, as mentioned in section 2.3, the three degrees of intent still apply to § 266 (as well as to § 266b). So even though the defense lawyer says that intent cannot be ‘inflected’ (Danish: *bøjes*), the three degrees of intent might be seen as exactly that – three inflections of intent.

in a threat. Nevertheless, they believe that the defendant intended the utterance to be a 'joke' (Danish: *vittighed*) that would also be perceived as a joke by her friends. The fact that her friends take it up as a joke serve as evidence for the judges that she did not have intent to threaten. They find that she had not understood (and in that, perhaps they mean that she had not even seen it as a possibility) that it would be taken up as a threat, thus excluding any intent. Their primary uptake of the utterance is therefore that of a joke, but the duality of the verdict is perhaps what causes the case to be appealed and taken up by the high court.

The summary from the verdict by the high court judges in excerpt (6) below from the high court judgment instead shows an even split between the two uptakes. Three judges vote for a conviction because they believe the defendant had intent to threaten, and three judges vote for an acquittal because they did not believe the defendant had intent to threaten:

(6) From summary of high court verdict in high court judgment:

'In the high court, three judges voted for acquittal for the reasons that the district court had stated, while three judges voted for conviction, because they found that the defendant had had the necessary intent to make the threat, including that it was fit to provoke the mentioned fear in the Criminal Code's § 266. Because of the distribution of votes, the defendant was then acquitted.'

Interestingly, the opposing uptakes of the utterance that we see throughout the case (via the prosecution and the defense) are morphed together in the district trial verdict (see sequence 14 and excerpt (5) further above), but in the high court verdict, they are instead split into two distinct and opposite uptakes (see excerpt (6) above). Because of the legal principle in *dubio pro reo* (from Latin: '[when] in doubt, in favor of the accused'),¹⁶ the defendant is acquitted when there is a parity of votes (Kistrup et al., 2018, p. 31 and 781). The case never went to the Supreme Court, and as such, the defendant was found not guilty.

5. Discussion and conclusion

One of the contributions of a genre perspective to Forensic Linguistics is that it offers a systematic examination of the context of an utterance, including how an utterance and its context is interpreted by various parties through their uptakes. Gales (2019) has specifically identified the need for greater attention to the social context of disputed threatening communications. This study has delved into the interrelation between a genre's uptake and the function thereby attributed to it. The disputed utterance is taken up – and therefore effectively also functions – as two different genres, as a threat and as a joke. These uptakes first happen within the immediate situation, and later – through various textual travels – the uptakes are recontextualized as part of the opposing sides in a courtroom.

The specific uptakes of an utterance can reveal important facets of genre conventions and genre understandings, since these are not fixed or immovable but rather (to a greater or lesser extent) flexible and negotiable. The two uptakes reflect the participants' understanding of the genres in question. As evident from the data, threats are

¹⁶ *In dubio pro reo* is often explained as an equivalent to the standard of proof *beyond a reasonable doubt*.

associated with a negative and destructive state of mind in the threatener, with seriousness (whether about intending to carry out a threatened act or about uttering a threat), and with the consequent effect of fear in victims. As a stark contrast, jokes are associated with a positive state of mind, non-seriousness, camaraderie and fun. The two genres are thus treated as each other's direct opposites in this case.

The view that a joke can only be funny and positive is not grounded in facts, however, as witnessed by negative humor. This case raises the question of what determines whether something is a joke – is it when the sender of an utterance claims to be joking or when both the sender and the audience see it as being funny. Moreover, humor can be part of destructive and extremely serious acts such as bullying or hate speech (e.g., Billig, 2001; Holmes & Marra, 2002). The tacit knowledge of genres that allows the defendant to write her utterance includes a tacit knowledge of threats, and even of the usual or recurrent form of a classic, direct threat. On top of this, it may also be a special kind of uptake on the genre of threats; one that includes that it is not meant to be taken seriously. As Tsakona describes it, “humor may result from incongruities in terms of genre expectations, that is from the fact that the boundaries between genres are often blurred as speakers deliberately and more or less playfully combine discursive features coming from different genres” (2017, p. 494). If in fact two such different genres were intended to be hybridized in this case, the effect of it did not just create a “humorous effect”, but prompted two entirely different uptakes, one of them extremely negative and serious. If, on the other hand, the defendant intended the utterance only to be funny, then the negative uptake is a confirmation of Freadman's (2012) assertion that no one can secure a specific uptake and therefore cannot even decide what genre their utterance represents. This is at issue particularly with social media, since it is difficult to know who one's entire audience or uptake community is. If, on the other hand, the defendant also intended to threaten or intimidate (while not necessarily intending to carry out the threat), then the negative uptake of the law is a response in kind, unwelcome as it may be for her. In a legal context, when a sentence is to be passed, the judges have to consider whether the utterance is an illegal threat or not – there is no in-between. As seen in this case, the dualistic potential of the utterance and its uptakes are reflected in the paradox within the district court verdict and also in the division of the high court judges' opinions.

In the two diverging uptakes, non-linguistic behavior is used as indirect evidence of or justification for the preferred interpretation of the linguistic behavior. As such, the context of the utterance becomes a means to justify the uptake of the utterance itself, but different parts of the generally undisputed context is used to serve each interpretation. The prosecution side focuses on the defendant's disruptive actions in class on the day that she wrote the Instagram post, on the fact that the utterance actually induced fear (and therefore was demonstrably ‘fit to provoke fear’) and to a minor degree on the fact that school shootings do happen. The defense side instead focuses on the positive situation later that same day where the defendant had a nice afternoon with two friends and on the fact that the utterance was posted on Instagram to friends who write in a similar manner. When referring to the act of school shooting, they are only talked about in relation to what happens in other countries, not as something that happens in Denmark. To explain the meaning of the utterance (its function, its genre), both sides refer to her ancillary actions and to whether school shootings are realistic or

not, which means that non-linguistic acts and behavior are an important part of their interpretations of the utterance.

In contrast, the linguistic interpretation is limited to the abbreviation 'Imao'. To my knowledge, the prosecutor does not explain anything about the linguistic content of the utterance itself, except in his closing statement, when he mentions that 'Imao' cannot function as a disclaimer of all other content in the utterance. In that way, he treats the language and genre knowledge as shared knowledge, i.e. everyone must know what a threat looks like (also see Bojsen-Møller, 2022). The only feature that needs explaining is the presumed apparent genre anomaly 'Imao'¹⁷, which the school representative in particular does not consider to be an unambiguous signpost of a joke. Aside from that, the prosecution takes the linguistic elements for granted. Considering the crucial role that different kinds of textual travels have in court cases, it is curious how the travel of the disputed utterance itself happens without any greater assessment of the content of the utterance. This substantiates one of the common claims about genres, namely that genre knowledge is tacit. Often, there is no inclination – and for non-linguists typically no acquired terminology – to explain explicitly what the linguistic or generic features of a specific genre are, because it is treated as shared and unproblematic, common knowledge that is intuitively understood. The district judges seem to agree with prosecution that no further explanation is needed, since they conclude that the utterance is an 'objectively seen illegal threat', which seems to be an indirect comment on the linguistic content of the utterance. They are not swayed by the defense's explanation of 'Imao' as a sufficient cue to change the genre from a threat to a joke (and we do not know what the high court judges' stance on the laughter cue is).

Despite the categorization and labelling of the utterance as an 'objectively seen illegal threat', the defense claim that the defendant lacked an intent to threaten becomes the deciding factor for the district court judges. In this way, the case clashes rather surprisingly with the established legal norm of cases tried under § 266, which focuses more on the perceived "character" of the utterance itself and less on the criminal intent of the defendant (Bojsen-Møller, 2022). This reflects the difficulty in objectively verifying what a person's actual intent is (see Threatening communications and the Danish threat statute). It is not uncommon for people accused of committing a language crime to disagree with the genre label that alleged victims, police or prosecution use and instead use a non-criminal or non-illicit genre label (Bojsen-Møller et al., 2020). The alternative label can either be a deceptive, or even self-deceptive, re-labelling technique to avoid legal repercussions, or a true reflection of their original intent behind the words. As this case study illustrates, determining intent will not only include an assessment of the credibility of the defendant, but also an assessment of contextual factors, such as the uptakes of others. This demonstrates that the uptakes of a disputed criminal utterance can have a deciding impact on the outcome of a criminal trial. In this case, the two diverging uptakes from two different groups of people make the situation more complicated and less straight-forward than if there had been one unanimous uptake of the utterance as either a threat or a joke. If everyone had taken up the utterance as a joke, then of course there would not have been a trial, but if everyone had taken up the

¹⁷People may not expect humor in threatening communications, but laughter cues do in fact occur in threats (cf. KORP threats), both as a way of intensifying the scary aspects of the threat and as a way of making other people laugh while the victim gets scared.

utterance as a threat, the outcome of the case would probably have been less divided. Instead, the dual (and duelling) uptakes are reflected in the ambivalence expressed in the verdicts. In other words, the differing uptakes of the utterance contribute to the conclusion of sufficient (reasonable) doubt for the defendant to be found not-guilty.

Sometimes genre interpretations have very real consequences. Sometimes they are the difference between guilty and not-guilty, prison and freedom.

Acknowledgements

This work was funded by the Carlsberg Foundation under grant number CF17-0907 for the research project *Understanding Threats: Language and Genre*. Thank you to PI of the project, Professor Tanya Karoli Christensen for her invaluable feedback on this article. I also thank Professor Tammy Gales, Professor Mary-Jo Reiff, Professor Emeritus Frans Gregersen, Professor Emerita Amy J. Devitt, Associate Professor Sune Auken, Professor Martha Sif Karrebæk, Professor Janus Mortensen and the anonymous reviewer for their helpful comments on this study.

References

- Anderson, S. (1992). The transition from inquisitorial to adversarial criminal procedure in Denmark. *Scandinavian Studies*, 64(2), 181–198.
- Atkinson, J. M., & Drew, P. (1979). *Order in Court: the Organisation of Verbal Interaction in Judicial Settings*. London: Macmillan.
- Auken, S. (2015). Utterance and function in genre studies: a literary perspective. In J. Andersen (Ed.), *Genre theory in information studies* (1. ed ed.). Bingley: Emerald.
- Auken, S. (2018). Understanding genre. *Journal of Zhejiang International Studies University*, 3(2), 14–27.
- Auken, S. (2019). Summing up. In T. K. Christensen & M. Bojsen-Møller (Eds.), *Illicit genres. Symposium*. Department of Nordic Studies and Linguistics, University of Copenhagen.
- Auken, S. (2021). Genres inside genres. A short theory of embedded genre. *Discourse and Writing/Rédactologie*, 31, 163–178. doi: 10.31468/dwr.883
- Austin, J. L. (1962). *How to Do Things with Words*. London: Oxford University Press.
- Bakhtin, M. (1986). The problem of speech genres. In C. Emerson & M. Holquist (Eds.), *Speech Genres and Other Late Essays* (pp. 60–102). Austin: University of Texas Press.
- Bala, P. (2020). ‘Tiltalte var ikke sur eller ophidset, da han gik.’ På sporet af tiltaltes stemme [‘The defendant wasn’t angry or upset when he left.’ Searching for the defendant’s voice]. *Spindet Sprogpsykologisk Information og Debat*, 16–19.
- Bell, N. D. (2017). Failed Humor. In S. Attardo (Ed.), *The Routledge Handbook of Language and Humor* (1st ed., pp. 356–370). New York, NY: Routledge. doi: 10.4324/9781315731162-25
- Billig, M. (2001). Humour and hatred: the racist jokes of the Ku Klux Klan. *Discourse and Society*, 12, 267–89.
- Billig, M. (2005). *Laughter and Ridicule: Towards a Social Critique of Humour*. London: Sage Publications.
- Bojsen-Møller, M. (2022). Fit to provoke fear?: Uptakes and textual travels of threatening communications in legal genres. *International Journal of Speech, Language and the Law*. doi: 10.1558/ijll.18869
- Bojsen-Møller, M., Auken, S., Devitt, A. J., & Christensen, T. K. (2020). Illicit Genres: The Case of Threatening Communications. *Sakprosa*, 12(1), 1–53. doi: 10.5617/sakprosa.7416
- Bucaria, C. (2008). Dubbing Dark Humour: A Case Study in Audiovisual Translation. *Lodz Papers in Pragmatics*, 4(2), 215–240. doi: 10.2478/v10016-008-0014-2
- Christensen, T. K. (2019). Indirect threats as an illegal speech act. In K. Ramshøj Christensen, J. Wood, & H. Jørgensen (Eds.), *The Sign of the V. Papers in Honour of Sten Vikner Festschrift* (pp. 113–130). Aarhus: Aarhus University.
- Devitt, A. J. (1991). Intertextuality in tax accounting: Generic, referential, and functional. In C. Bazerman & J. Paradis (Eds.), *Textual Dynamics of the Professions: Historical and Contemporary Studies of Writing in Professional Communities* (pp. 336–57). Madison: University of Wisconsin.
- Devitt, A. J. (2009). Re-fusing form in genre study. In J. Giltrow & D. Stein (Eds.), *Genres in the Internet: Issues in the Theory of Genre* (pp. 27–48). Amsterdam: John Benjamins.

- Fraser, B. (1998). Threatening revisited. *Forensic Linguistics*, 5(2), 159–173. doi: 10.1558/sll.1998.5.2.159
- Freadman, A. (1994). Anyone for tennis. In A. Freedman & P. Medway (Eds.), *Genre and the New Rhetoric* (pp. 43–66). London: Taylor & Francis.
- Freadman, A. (2002). Uptake. In R. Coe, L. Lingard, & T. Teslenko (Eds.), (pp. 39–53). Cresskill: Hampton Press.
- Freadman, A. (2012, December). The Traps and Trappings of Genre Theory. *Applied Linguistics*, 33(5), 544–563. doi: 10.1093/applin/ams050
- Freadman, A. (2020). A Tardy Uptake. *Discourse and Writing/Rédactologie*, 30, 105–132. doi: 10.31468/cjsdwr.781
- Fuller, P. B. (2015). Evaluating intent in true threat cases: The importance of context in analyzing threatening Internet messages. *Hastings Communications and Entertainment Law Journal*, 37, 37–58. Retrieved from https://repository.uchastings.edu/hastings_comm_ent_law_journal/vol37/iss1/2
- Gales, T. (2010). *Ideologies of Violence: A Corpus and Discourse Analytic Approach to Stance in Threatening Communications* (Doctoral dissertation). University of California, Davis, CA.
- Gales, T. (2019). Threatening contexts. In M. Evans, L. Jeffries, & J. O’Driscoll (Eds.), *The Routledge Handbook of Language in Conflict* (1st ed., pp. 472–490). Routledge. doi: 10.4324/9780429058011-27
- Gingiss, P. (1986). Indirect threats. *WORD*, 37(3), 153–158. doi: 10.1080/00437956.1986.11435774
- Gumperz, J. J. (1982). *Discourse Strategies*. Cambridge: Cambridge University Press.
- Heffer, C. (2013). Communication and magic: Authorized voice, legal-linguistic habitus, and the recontextualization of “beyond reasonable doubt”. In C. Heffer, F. Rock, & J. Conley (Eds.), *Legal-Lay Communication, Textual Travels in the Law* (pp. 206–225). New York: Oxford University Press.
- Heffer, C., Rock, F., & Conley, J. (2013). *Legal-Lay Communication, Textual Travels in the Law*. New York: Oxford University Press.
- Heydon, G. (2013). From legislation to the courts: Providing safe passage for legal texts through the challenges of a police interview. In C. Heffer, F. Rock, & J. Conley (Eds.), *Legal-Lay Communication, Textual Travels in the Law* (pp. 55–77). New York: Oxford University Press.
- Holmes, J., & Marra, M. (2002). Over the edge? Subversive humor between colleagues and friends. *International Journal of Humor Research*, 15, 65–87.
- Hurt, M., & Grant, T. (2019). Pledging to harm: A linguistic appraisal analysis of judgment comparing realized and non-realized violent fantasies. *Discourse & Society*, 30(2), 154–171. doi: 10.1177/0957926518816195
- Jefferson, G. (1984). On the organization of laughter in talk about troubles. In *Structures of Social Action*. Cambridge: Cambridge University Press.
- Jefferson, G. (1985). An exercise in the transcription and analysis of laughter. In T. A. van Dijk (Ed.), *Handbook of Discourse Analysis, vol. 3*. London: Academic Press.
- Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. H. Lerner (Ed.), *Pragmatics & Beyond New Series* (Vol. 125, pp. 13–31). Amsterdam: John Benjamins Publishing Company. doi: 10.1075/pbns.125.02jef
- Jönsson, L., & Linell, P. (1991). Story generations: From dialogical Interviews to written

- reports in police interrogations. *Text - Interdisciplinary Journal for the Study of Discourse*, 11(3). doi: 10.1515/text.1.1991.11.3.419
- Kistrup, M., Poulsen, J. L., Røn, J., & Rørdam, T. (2018). *Straffeprocessen. [Criminal Procedure]* (3rd edition. ed.). Copenhagen: Karnov Group.
- Komter, M. (2019). *The Suspect's Statement: Talk and Text in the Criminal Process*. Cambridge University Press.
- Komter, M. L. (1994). Accusations and Defences in Courtroom Interaction. *Discourse & Society*, 5(2), 165–187. doi: 10.1177/0957926594005002002
- Lewis, P. (1997). The killing jokes of the American eighties. *Humor: International Journal of Humor Research*, 10, 251–83.
- Medway, P. (2002). Fuzzy genres and community identities: The case of architecture students' sketchbooks." In R. Coe, L. Lingard, & T. Teslenko (Eds.), *The Rhetoric and Ideology of Genre: Strategies for Stability and Change* (pp. 123–153). Cresskill NJ: Hampton.
- Miller, C. R. (1984, May). Genre as social action. *Quarterly Journal of Speech*, 70(2), 151–167. doi: 10.1080/00335638409383686
- Miller, C. R. (2017). "Where Do Genres Come From?". In C. R. Miller & A. R. Kelly (Eds.), *Emerging Genres in New Media Environments* (pp. 1–34). Cham: Springer International Publishing. doi: 10.1007/978-3-319-40295-6_1
- Mortensen, N., J. and Coupland. (2018). Style and styling. In *Handbook of Pragmatics*.
- Mortensen, S. S., & Mortensen, J. (2017). Epistemic Stance in Courtroom Interaction. In F. Poggi & A. Capone (Eds.), *Pragmatics and Law* (Vol. 10, pp. 401–437). Cham: Springer International Publishing. (Series Title: Perspectives in Pragmatics, Philosophy & Psychology) doi: 10.1007/978-3-319-44601-1_16
- Muschalik, J. (2018). *Threatening in English: A Mixed Method Approach*. Amsterdam: John Benjamins Publishing Company.
- Nini, A. (2017, June). Register variation in malicious forensic texts. *International Journal of Speech Language and the Law*, 24(1), 99–126. doi: 10.1558/ijsl.30173
- Norrick, N. R. (2010, April). Humor in Interaction: Humor in Interaction. *Language and Linguistics Compass*, 4(4), 232–244. doi: 10.1111/j.1749-818X.2010.00189.x
- Nyboe, J. (2016). The Game of the Name: Genre Labels as Genre and Signature. *Scandinavian Studies*, 88(4), 364–392. doi: 10.5406/scanstud.88.4.0364
- Raskin, V. (1985). *Semantic Mechanisms of Humor*. Dordrecht: D. Reidel.
- Sachs, J. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception & Psychophysics*, 2, 437–444.
- Salgueiro, A. B. (2010). Promises, threats, and the foundations of speech act theory. *Pragmatics. Quarterly Publication of the International Pragmatics Association (IPrA)*, 213–228. doi: 10.1075/prag.20.2.05bla
- Scott, M. B., & Lyman, S. M. (1968). Accounts. *American Sociological Review*, 33(1), 46. doi: 10.2307/2092239
- Shauer, F. (2003). Intentions, conventions, and the First Amendment: The case of cross-burning. *Supreme Court Review*, 197–230.
- Shuy, R. W. (1993). *Language Crimes: The Use and Abuse of Language Evidence in the Courtroom*. Cambridge: Blackwell.
- Solan, L. M., & Tiersma, P. M. (2005). *Speaking of Crime, the Language of Criminal Justice*. Chicago: University of Chicago Press.

- Swales, J. (2016). Reflections on the concept of discourse community. *ASp*(69), 7–19. doi: 10.4000/asp.4774
- Toftegaard Nielsen, G. (2016). *Straffesagens Gang. [The Course of the Criminal Trial]*. Copenhagen: Jurist- og Økonomforbundet.
- Toftegaard Nielsen, G., Elholm, T., & Jakobsen, M. N. (2017). *Kommenteret Straffelov, Speciel Del. [The Annotated Criminal Code, Special Part.]* (11th ed.). Copenhagen: Jurist- og Økonomforbundet.
- Tsakona, V. (2017). Genres of Humor. In S. Attardo (Ed.), *The Routledge Handbook of Language and Humor* (1st ed., pp. 489–503). New York, NY: Routledge. doi: 10.4324/9781315731162-34
- Waaben, K., & Langsted, L. (2015). *Strafferettens Almindelige Del [The Criminal Law's Ordinary Part]*.
- Walton, D. (2000). *Scare Tactics, Arguments that Appeal to Fear and Threats*. Dordrecht, Netherlands: Kluwer Academic Publishers.
- Watson, D. R. (1992). Ethnomethodology, conversation analysis and education: An overview. *International Review of Education*, 38(3), 257–274. doi: 10.1007/BF01101432
- Yamanaka, N. (1995). On indirect threats. *International Journal for the Semiotics of Law*, 8(1), 37–52. doi: 10.1007/BF01677089

Appendix A: Danish data

- (1) Fra Instagram-opslag, citeret i anklageskriftet og lagt op som billede på dr.dk:
jeg bliver den næste school shooter guys, lmao, watch out
- (2) Fra resumeet i Landsrettens dombog af skolerepræsentantens vidneforklaring:
Moren sagde til hende, at hendes datter og andre elever ikke turde møde i skolen.
- (3) Fra resumeet i Landsrettens dombog af skolerepræsentantens vidneforklaring:
Vidnets forudgående kendskab til tiltalte og kendskabet til episoden dagen i forvejen indgik i hendes overvejelser, inden hun ringede til politiet. Hun kan ikke svare på, om hun ville have ringet, hvis der havde været tale om en elev uden denne forhistorie. Inden hun kontaktede politiet, slog hun forkortelsen »lmao« op, som hun ikke på forhånd kendte, men som betyder noget i retning af, at »jeg griner min røv i laser«. Man kan dog ikke vide, om det sjove var at sende beskeden ud eller at udføre handlingen. Siden 2013 har skoleskyderier været med i beredskabsplanerne.
- (4) Fra resumeet i politirapporten af lærerens forklaring:
Foreholdt at sigtede har lagt et opslag ud på Instagram omkring at sigtede kunne blive den næste skoleskyder, forklarede afhørte, at hun ved at sigtede ikke ønsker at gøre alvor af disse trusler. Afhørte mener, at sigtede ikke har haft kunne kontrollere sine tanker da hun skrev det.
- (5) Fra resumeet af byretsdommen i Landsrettens dombog:
Byretten anførte, at opslaget var egnet til at fremkalde frygt for liv og helbred, jf. straffelovens § 266, selv om det ikke var ment bogstaveligt eller alvorligt og indeholdt slangudtrykket »lmao«. Retten lagde dog til grund, at T [tiltalte] mente opslaget som en vittighed, der ville blive læst af hendes venner, og at disse også ville opfatte opslaget som en vittighed. Retten fandt derfor, at T ikke havde forstået, at opslaget ville blive opfattet som en trussel, og at hun derfor ikke havde haft forsæt til overtrædelse af straffelovens § 266.
- (6) Fra resumeet af landsretsdommen i Landsrettens dombog:
I landsretten stemte tre voterende for frifindelse af de grunde, byretten havde anført, mens tre voterende stemte for domfældelse, da de fandt, at T havde haft det fornødne forsæt til at fremsætte truslen, herunder at den var egnet til at fremkalde den i straffelovens § 266 nævnte frygt. Efter stemmefordelingen blev T herefter frifundet.

Transskription af den originale danske lydoptagelse fra byretten

P=prosecutor (anklager)

D=defendant (tiltalte)

L= defense lawyer (forsvarer)

T=teacher (lærer)

Tidsinterval: 2,042-2,357 minutter

001 **P:** anklageskriftet er dateret ((DATO)) >og lyder at<
 ((NAVN)) tiltales ved retten i ((BYNAVN)) med påstand om straf for (0.6)
 trusler efter straffelovens paragraf tohundredogseksogtres (0.7)
 ved den ellefte september totusindatten omkring midnat (0.5)
 som elev på ((BYNAVN)) gymnasium (.)
 at ha truet med at foretage en strafbar handling (0.5)
 idet tiltalte lagde et billede af <sig> selv
 på instagram med TEKSten (.)
 jeg blir den næste school shooter (.) guys (0.3)
 L M A O (0.3) watch out↑ (.)

((imellem 001 og 002 forklarer anklageren og journalisten skiftevis om sagen))

Tidsinterval: 5,351-6,347 minutter

002 **P:** dagen før det her skete
 var du så i: gymnasiet dagen før

003 **D:** JA

004 **P:** og og skete der noget der? (0.7)

005 **D:** øhm jeg blev lidt SUR. (0.3) og væltede et bord, (0.3)

006 **P:** °ja°, hvad var du blevet sur over?

007 **D:** øhm >jeg var blevet lidt< sur over at >hva hedder det< (0.3)
 øhm ((NAVN)) som (0.6) æh kom med provoKERne °altså°
 (1.0)
 hentydninger til mig
 (1.4)

008 **P:** var det noget med en fremlæggels[e el]ler:?

009 **D:** [JA]

010 **P:** °ja° prøv at fortæl hva det var=

011 **D:** =jamen jeg sku fremlægge og vi havde øvet det osse
 og så tog han så hva jeg sku sige (0.8)
 og derefter sagde han så
 >jeg venter bare på hun klapper MUNDEN op< (0.6)
 højt i klassen ((sagt med skælvende stemme))
 øhm
 (1.4)
 hvilket e- fik >mig til at være lidt< irriteret

012 **P:** blev du nærmest ydmyget ved den lejligh[hed]?

- 013 D: [JA]
- 014 P: °ja jaeh° hvor- var det over for hele klassen og læreren?
- 015 D: mm (0.8)
- 016 P: og hvordan reagerede du på det. (0.7)
- 017 D: ja det var så: så vælted jeg bordet (0.3)
- 018 P: °ja° var du os øh: (0.5) i øvrigt (0.5)
reagerende over for ((NAVN)) eller [>over for en af de andre<]
- 019 D: [NE:J slet ik]
jeg gik hurtigt ud bagefter
- ((imellem 019 og 020 forklarer journalist at to veninder gik efter hende og en lærer kom også og snakkede med hende))
- Tidsinterval: 6,442-8,266 minutter
- 020 P: og den (0.3) øh: LÆRER du så snakked med EFTER
hvem var det.
- 021 D: øhm ((NAVN))
- 022 P: var det ((FORNAVN EFTERNAVN))? (0.7)
som kommer og afgiver forklaring i dag oss.
- 023 D: [jae]
- 024 P: [ja] okay.
tog du så hjem efter at ha snakket med ((NAVN))
- 025 D: ja så kørt hun mig hjem og sagde at jeg ku (0.3)
starte (0.4) forfra igen i morgen
jeg fik det meget bedre efter jeg havde snakket med ((NAVN))† (.)
og så >valgt jeg så at vær sammen med< nogen venner (.)
efter det sket (0.4)
- 026 P: mm
- 027 D: hvor vi så bare laved lasagne og hyggede os
og jeg tænkt at det var ovre (0.6)
- 028 P: jah snakked i om noget bestemt? (0.5)
- 029 D: vi havde snakket om mange ting og vi havde også snakket lidt om
så:n skoleskyderier og sånoget (.) i de andre lande (0.4)
- 030 P: mm (0.7) hvorfor faldt snakken på det.
ka du husk det?
- 031 D: ja altså vi var sån på nettet (0.4)
- 032 P: [°ja°]
- 033 D: [og] så SÅ vi sådn nogen ARTI-ARTIKLER
(1.1)

- 034 **P:** ja (0.8) hvorhenne på nettet så i det?
- 035 **D:** facebook
- 036 **P:** ja (0.5) og så ved titiden tog du hjem. (0.4)
- 037 **D:** ja
- 038 **P:** ja. og hva sket der så?
(1.0)
- 039 **D:** så spilled jeg computer indtil klokken blev tolv
(1.0)
og så sku jeg gå i SENG† (0.5)
og jeg havde taget det billede da jeg var sammen med dem
- 040 **P:** mm
- 041 **D:** hvor jeg sad på TOILETTET† (0.3) ((kraftfuld stemme))
- 042 **P:** mm
- 043 **D:** bare sån for SJOVT ((kraftfuld stemme))
- 044 **P:** [mm]
- 045 **D:** [og] >vi havde taget billeder sammen og sådn noget< (0.4)
øhm og så tog jeg hjem og tænkt at jeg ville lægge et billed op (0.3)
- 046 **P:** [mm]
- 047 **D:** [så] jeg tog det billed hvor jeg sad på toilettet
- 048 **P:** mm
- 049 **D:** og så skrev jeg så og så-d >sku jeg find på noget at skrive<
og det blev så det.
(1.2)
- 050 **P:** hva ment du MED det?
(1.0)
- 051 **D:** jeg ment altså:, det var lissom sådn:(0.7)
specielt til DE to altså (.) °at° (0.3)
>de ville ku forstå det fordi vi lige havde snakket om det< (.)
så det var lissom <en joke> (0.5)
- 052 **P:** der hvor du <skriver> >jeg blir den< næste <school shooter guys>
lmaO‡ watch out.
- 053 **D:** ja (0.3)
- 054 **P:** bar så vi allesammen forstår det
[hva] står lmao for
- 055 **D:** [ja] øh laughing my ass off
- 056 **P:** ja øh >griner min røv i laser<
- 057 **D:** [JA PRÆCIS]
- 058 **P:** [ja ja] (0.3)

059 **D:** hvilket er med til det en joke.

((imellem 059 og 060 forklarer journalist at tiltalte her uploader billedet hvor alle hendes følgere kan se det))

Tidsinterval: 8,316-9,246 minutter

060 **P:** det her udbrud øh som du havde

(0.7)

har du haft flere af dem.

061 **D:** na:rj

062 **P:** ka du-øh: kan du sige du nogen gange blir vældig sur (.)

063 **D:** ja: jeg ka godt blive lidt sur

064 **P:** ja (.) og det jeg jo interesserer mig for øh

det er jo lidt hvordan de andre opfatter dig

og hvorda-hvordan DU etelandetsted måske VED at de andre opfatter dig

(0.7) øh: og ja i forhold ET bilag syv

det en afhøringsrapport (0.7)

er det ri-kan du huske du har snakket med politiet?

065 **D:** JA

066 **P:** ja øhm der står på side fire i rapporten

afhørte forklarede videre at afhørte har PROBLEMER MED SIT HUMØR

(1.0)

afhørte kan blive vældig SUR når hun bliver sur

(1.0)

ka du øh: har du sagt sådn til politiet?

067 **D:** ja (.)

068 **P:** °ja° er det osse sådn det var. (0.5)

069 **D:** ja

070 **P:** ja ((skarp stemme))

har der været andre episoder hvor du blev sur i klassen

og måske ku ha [<været lidt>] (0.4) [vred?]

071 **D:** [ne:j] [>nejnejnej<]

jeg har overhovedet ikke været sur i klassen eller noget (.)

072 **P:** nej (.) så det var første gang det her

073 **D:** ja det v- (0.5)

074 **P:** TAK jeg har fl-foreløbigt ikke flere spørgsmål til tiltalte

((imellem 074-075 forklarer journalist at det nu er forsvarerens tur til at udspørge tiltalte og at hun siger opslaget var en joke så forsvarer vil høre mere om humor og sproget på nettet))

Tidsinterval: 11,531-12,355 minutter

- 075 **L:** du ha- du:- >jeg sys du har forklaret<
at de godt ku ha en SORT humor? (.)
- 076 **D:** ja (0.5)
- 077 **L:** øh prøv at forklare hve-hva ku den hva-hva ka DET VÆRE
- 078 **D:** øh vi ka skrive ting som om at
øh vi vil dræbe os selv næste dag (.)
og så skriver vi lol eller l m a o (.) bagv-bagefter (.)
som om at >det gør vi seføli ikke< det en joke (0.3)
- 079 **L:** jaeh
(1.6)
SÅ når du skr-når du udtrykker L M A O (0.7)
øh så mener du hvad? (0.6)
- 080 **D:** at DET en JOKE†
- 081 **L:** at det en joke.
- 082 **D:** ja
- 083 **L:** e-er det noget alle forSTÅR? (.)
alts[å:]
- 084 **D:** [al]le teenagere og (0.6)
<folk> som er på nettet burde fu- ku forstå det,
- 085 **L:** altså alle dem (.) som som ser (0.3) [dine] opslag (0.4)
- 086 **D:** [ja] ja
- 087 **L:** de ved hvad det bety[der]
- 088 **D:** [ja]

((imellem 088-089 forklarer journalist at en elev fra klassen blev utryg ved opslaget))

Tidsinterval: 12,396-13,075 minutter

- 089 **L:** den pige hvis mor (0.4) vi kan se (0.3)
kontakter ((SKOLEREPRÆSENTANT)) om morgenen angiveligt
var det en af dem her?
- 090 **D:** nej det var ikke nogen af mine venner: det var en fra klassen (0.7)
som ik (0.5) KENDT mig °så godt°

091 L: men men øh kontakter HUN dig om natten [eller om morgenen]?

092 D: [nej slet ik]

093 L: nej

094 D: der ikke nogen (0.8) altså ik [ik hende]

095 L: [>der ik nogen kontakt<]

096 D: [nej]

097 L: [imel]lem dig og hende på nogen måder

(0.9)

ved du om hun har SET_↑ dit opslag

098 D: ja vi var VENner på instagram jeg så hendes og hun så osse mine

099 L: jaeh

((imellem 099-100 forklarer journalist at tiltalte fik andre typer reaktioner på opslaget i form af venner der bryggede videre på hendes spøg))

Tidsinterval: 13,126-14,214 minutter

100 L: du FÅR nogen reakTIONer på det her opslag er det ikke korrekt,

101 D: jo (0.4)

102 L: øh prøv og prøv lige og forklar hva det-hva der står DER

103 D: ja han har skrevet ka du skoleskyde på min skole så? x d (0.5)

og xd det er lissom at skriv (0.5)

øh det lissom en smiley (.) som griner (0.7)

så det: m-er med til: foreksempel lol og de der ting (.)

104 L: o-og når HAN sir det ka du skoleskyde på MIN skole [osse] X D SMiley

105 D: [JA]

JA

106 L: hva MENer han SÅ med DET. det mener han ik alv[orligt]

107 D: [nej]

så det bar sådant for sjovt

108 L: >for sjovt han sir det< (0.6)

du sagde før at (.) der osse var ANDre der

<skrev som dig> på DE >medier<

109 D: ja (0.5) ja <mange andre> altså på FACEbook

hvor i grupper hvor de skriver at de vil GØ:re det sam:

°og sån noget°

110 L: for sjov.

111 D: for sjov.

091 L: men men øh kontakter HUN dig om natten [eller om morgenen]?

092 D: [nej slet ik]

093 L: nej

094 D: der ikke nogen (0.8) altså ik [ik hende]

095 L: [>der ik nogen kontakt<]

096 D: [nej]

097 L: [imel]lem dig og hende på nogen måder

(0.9)

ved du om hun har SET₁ dit opslag

098 D: ja vi var VENner på instagram jeg så hendes og hun så osse mine

099 L: jaeh

((imellem 099-100 forklarer journalist at tiltalte fik andre typer reaktioner på opslaget i form af venner der bryggede videre på hendes spøg))

Tidsinterval: 13,126-14,214 minutter

100 L: du FÅR nogen reakTIONer på det her opslag er det ikke korrekt,

101 D: jo (0.4)

102 L: øh prøv og prøv lige og forklar hva det-hva der står DER

103 D: ja han har skrevet ka du skoleskyde på min skole så? x d (0.5)

og xd det er lissom at skriv (0.5)

øh det lissom en smiley (.) som griner (0.7)

så det: m-er med til: foreksempel lol og de der ting (.)

104 L: o-og når HAN sir det ka du skoleskyde på MIN skole [osse] X D SMiley

105 D: [JA]

JA

106 L: hva MENer han SÅ med DET. det mener han ik alv[orligt]

107 D: [nej]

så det bar sådant for sjovt

108 L: >for sjovt han sir det< (0.6)

du sagde før at (.) der osse var ANDre der

<skrev som dig> på DE >medier<

109 D: ja (0.5) ja <mange andre> altså på FACEbook

hvor i grupper hvor de skriver at de vil GØ:re det sam:

°og sån noget°

110 L: for sjov.

111 D: for sjov.

112 **L:** du har til poliTIet forklaret (0.4)
 du er øh ciTEret for at sige
 at det skulle forstås som <en> SPØG
 eller en s- eller sarkastisk (0.4)
 >og så skriver du< (0.3) >eller så siger du< (0.3)
 alle som kender afhørte (.) ved at afhørte
 >ikke kunne finde PÅ< <at GØRE NOgen forTRÆD>

113 **D:** ja
 (1.2)

114 **L:** o-og hvorfor sir du hvorfor mener du det? (0.4)
 altså hva BYGger du det PÅ?

115 **D:** øh jeg bygger det PÅ at folk ved at jeg (.)
 ik ku find på at lave sådn noget
 og at jeg bare en følsom lille pige ((stemme skælver))

116 **L:** jae:rh

((imellem 116 og 117 forklarer journalist at det er skolerepræsentanten der afgiver forklaring om at have modtaget en mail fra en bekymret forælder))

Tidsinterval: 16,043-16,260 minutter

117 **P:** emne ved-VIGtigt ring meget gerne A S A P altså så snart som muligt
 på teleFONnummer, vedrørende en elev fra ((NAVN PÅ KLASSE))
 som gik aMOK i KLASSEN i GÅR i forbindelse med en fremlæggelse (0.5)
 hun har efterfølgende lagt en: >my story< OP (.)
 hvor hun skriver jeg blir den næste school shooter.

((imellem 117 og 188 gengiver journalist dele af skolerepræsentantens vidneforklaring og dernæst forklarer journalist at tiltaltes lærer er næste vidne))

Tidsinterval: 18,435-19,417 minutter

118 **T:** og hun hun sidder sådn og holder (.) på sig selv ik?
 og hun siger at hun er (0.8) meget meget KED af det

119 **P:** fortalt hun om hun var f-følt sig nedgjort af nogen?

120 **T:** JA (0.4)

121 **P:** °okay° (.) så HAR jeg lige en enkelt ting mere
 sigtede forTALTE at hun havde haft en masse GRIMME tanker
 om den eLEV som hun (.) havde følt havde NEDgjort sigtede
 sigtede fortalte at hun ik ville FØLGE de tanker

- >men hellere la det gå ud over< et objekt.
 har du [sagt] sådan til politiet
- 122 **T:** [JA] JA
- 123 **P:** Okay. (.)
 >adspurgt om hvordan sigtede er som elev
 beskriver afhørte hende< som en DEJlig pige
 øh som meget gerne vil alle de RIGtige og GODE ting
- 124 **T:** JA
- 125 **P:** sigtede er en SKRØbelig pige som ka blive VRED
 hvis andre ikke forSTÅR hva hun MENER (0.3)
 sigtede Siger at hun har vrede og gramme-grimme tanker (.)
 hun ønsker ikke at være en (.) som har sådanne tanker
- 126 **T:** JA
- 127 **P:** er det osse sådn øh du har for[klaret]
- 128 **T:** [JA]
- 129 **P:** (0.5) relaterer det sig så KUN til den her dag?
 el[ler (.) hvordan ska jeg forstå det?]
- 130 **T:** [hhhhhhhhhhhhhhhhhhhh ((langt suk))]
 (0.9)
 ahm altså jeg har ik nogen oplevelser med hende (.)
 øh:: (0.3) altså dårlig oplevelse TVERTIMOD (.)
 øhm: det har jeg altså ik (0.5)
- 131 **P:** Okay.
 2.9 ((bladren af papirer))
- Tidsinterval: 19,417-19,512 minutter
- 132 **L:** du havde sagt til politiet at du var sikker på (0.3)
 at øh (0.8) tiltalte ik ville gøre alvor af sine trusler,
- 133 **T:** JA jeg [er f-]STENSIKKER på det ville hun ALDRIG gøre
- 134 **L:** [()]
 arj
- ((imellem 134 og 135 forklarer journalist at det er tid til procedurerne))
- Tidsinterval: 20,039-22,191 minutter
- 135 **P:** vi VED fra vidnerne og osse fra hende selv
 at der jo altså har været nogle episoder og

øh: det ka være mere eller mindre alvorligt sådn som man opfatter det
 men vi ved i HVERT fald at der dagen FØR (0.7)
 i klassen var en voldsom episode.
 vi ved osse (0.5) at hun kastede eller væltede et bord i vrede
 eller i (.) frustration
 og at hun (0.4) havde det rigtig DÅrligt (0.6)
 og det var det SIDste de v-øh personer i KLASseværelset så til hende
 (0.3) og nogen af de personer (0.5)
 må hun jo i et eller andet omfang have indset (.)
 formentlig ville gå ind og kigge
 da fyrre halvtREDS normalt kigged på hendes opslag (0.6)
 og hun så samme aften (0.3) slog et opslag op hvor hun skrev
 i'lll (.) øh jeg blir den næste <school shooter> guys (0.4)
 vi ved osse at øh læreren blev gjort bekendt med
 at hun havde GRIMME tanker
 hun havde også en FRYGT for at de GRIMme tanker de blev OMSat (0.5)
 og derfor så ville hun RETte det mod objekter
 (1.3)
 det er FAKTISK VELKENDT at øh skoleskyderier FINDER sted andre steder
 (0.3) det ved vi allesammen godt (0.6)
 men det: grunden til jeg siger det er IKKE
 at øh beskyldte øh: tiltalte FOR at ville GØRE det her (0.5)
 det tror jeg sådn set HELLER ik hun ville gøre
 det heller ik AFgørende for domfældelsen (.)
 i virkeligheden BEHØVer jo det faktisk ikke
 for at øh tiltalte ka blive dømt her være sådn
 at der er nogen der for: alvor blir bange
 det ska bare være EGNET til at SKAbe denne her alvorlige frygt
 for liv helbred eller velfærd (0.7)
 men det blev der (0.6)
 også ((SKOLEREPRÆSENTANTEN)) tog det ganske alvorligt
 >og inden vi vidst af det var< (0.3)
 politiet kommet til stede (0.9)
 ka man så (.) AFLive det hele ved at skrive L M A O (0.8)
 ka man være HELT sikker på at alle tænker
 det bare for sjov? (0.3)
 ja der ER nogen (0.6) øh: som har <en anden> (0.6)

intern øh: umiddelbar forståelse af det HER (.)
 når de skriver noget til hinanden
 men hun havde ikke begrænset det til dem (0.9)
 når det så er sagt (.) altsammen
 så er det altså at jeg prøver at summere det OP til og anbefale
 at an-at domsmandsretten (0.7) lægger til grund
 at hun skal dømmes i overensstemmelse med anklageskriftet (.)
 og findes skyldig i overtrædelse af paragraf tohundredogseksogtres
 (1.3)
 det er altså noget der BERØRER mange mennesker (.)
 det er noget der har haft en STOR effekt på SKolen (0.5)
 og det er noget: som også har en sigNALværdi (0.3)
 øh hva det er (0.4) øh der blir reaktIonen fra domsmandsretten i DAG
 (1.4)
 på DEN baggrund (.) så er det min ANbefaling til domsmandsretten
 >at det her ska< AFGøres m:ed en
 kortere fængselsstraf på tyve til tredive dages fængsel;

((imellem 135 og 136 opsummerer journalisten anklagerens procedure og forklarer at det nu er forsvarerens tur))

Tidsinterval: 22,378-24,448 minutter

136 L: >s-tiltalte nægter sig< FORTsat skyldig (0.6)
 og jeg øhm er jo på en RÆKKE (0.3)
 VÆsentlige områder UENIG med anklageren
 øh- man er nødt til og fortolke på <selve den formulering> (.)
 OG (.) >man er nødt til< at ta HENSyn TIL (.) i hvilken SAMmenhæng (.)
 at den er fremsat. (0.4)
 den er FREMsat i de-på de sociale medier (.) på instagram (0.4)
 den er IKke FREMsat (0.9) til skolen; (0.3)
 den er fremsat (.) øh i et forum (0.3)
 hvor man TALER til hinanden og bruger e:n-et (0.3)
 >ja vel næsten< for: dem der IKKE ved det (0.7)
 som åbenbart er ((SKOLEREPRÆSENTANTEN)) (0.5)
 så TALER man jo nærmest i et KODESPROG; (0.7)
 og hvis IK man er i <STAND TIL> (0.3)
 og øh- og læse og forSTÅ den <samtaleform; > (0.6)

så man jo reelt ikke i stand til at VURDERE
om der er tale om en TRUSSEL eller en spøg. (0.3)
>og så kan man sige< det kan være e:n ÅNDSSVAG spøg
(1.3)
det kan være en <UAGTSOM> spøg (0.6)
>det kan man ik dømmе for efter toseksogtres< (0.6)
MEN men man er nødt til (0.3) og forSTÅ (0.4)
hvad L M A O beTYder i den <SAMMENhæng> (0.4) den er f- øh fremsat i
til de (0.4) øh unge mennesker til de (.) TEENAGERE
som ALLEsammen kommunikerer på samme måde (0.3)
og når man bruger det UDTRYK (.)
og det jo VIGtigt at domsmandsretten FORSTÅR det (0.6)
så BETYDER det (0.6)
jeg blir den næste <SKOLEskyder> (0.8) bare for sjov (0.8)
og det klart (.) >at hvis ik man< FORSTÅR det. (0.7)
så ka det jo være et proBLEM (0.5)
men det jo ikke (0.5) høje domsmandsret (.)
i (.) straffelovens paragraf tohundredeseKSogtres forstæn-forstand (.)
ska det jo ikke <gå udover> TILTALTE
(1.0)
fordi ANDET led (0.4) i toSEKSogTRES er (0.7)
at TILtalte har haft FORsæt (0.6) til at FREMkomme med en trussel
(1.1)
o-og det ik et led der ka bøjes (0.6)
man ka ikke BAGEfter blive KLAR over at det var for DUMT (.)
eller (0.3) <det ku godt være> etellerandet
eller NOGEN ku MÅSKE forstå det-
(1.1)
med hensyn til (0.4) hvis det høje domsmandsret må nå frem til
der skal ske en (.) DOMFÆLDELSE i nærværende sag (0.3)
så-så må det være rettens ABSOLUT mildeste DOM

((imellem 136 og 137 forklarer journalist at tiltalte får sidste ord over for dommeren))

Tidsinterval: 24,509-25,058 minutter

137 J: er der noget du vil sige før vi går ud (.)

og drøfter sagen?

og kommer tilbage og træffer en afgørelse?

138 D: øh (0.7) lidt (0.7)

139 J: værsgo

140 D: øhm øh jeg ville bare sige at jeg aldrig havde ment det som en TRUSSEL

(0.5) og jeg selvfølgelig er KED af ⁰hva der er sket⁰

((imellem 140-141 forklarer journalist at dommer og domsmænd voterer og kommer ind igen))

Tidsinterval: 25,236-25,474 minutter

141 J: thi kendes for ret (.) tiltalte ((NAVN)) skal frifindes. (0.6)

statskassen ska betale sagens omkostninger (0.5)

ta plads

(3.1)

og jeg kan sige til dig at øh vi enige om at du skal frifindes

(0.9) ikke: fordi (0.3) der ikke er tale om en (0.8)

objektivt set strafbar trussel men fordi vi mener

at øh du ikke havde det fornødne forsæt. øh: (0.5)

Appendix B: Transcription of Danish audio recording from district court

P=prosecutor

D=defendant

L= defence lawyer

T=teacher

Time interval: 2.042-2.357 minutes

001 P: the indictment is dated ((DATE)) >and reads that<
 ((NAME)) is charged by the court of ((NAME OF CITY))
 seeking punishment for (0.6)
 threats under the criminal code's section twosixtysix (0.7)
 by on the eleventh of september twothousandeighteen around midnight (0.5)
 as a student at ((NAME OF CITY)) high school (.)
 having threatened to commit a punishable act (0.5)
 in that the defendant put a picture of herself
 on instagram with the TEXT (.)
 i'll be the next school shooter (.) guys (0.3)
 L M A O (0.3) watch out! (.)

((the prosecutor and the journalist take turns in explaining about the case))

Time interval: 5.351-6.347 minutes

002 P: the day before this happened
 where you then a:t high school the day before

003 D: YES

004 P: and and did something happen then? (0.7)

005 D: umm i got a bit ANGRY. (0.3) and flipped a table, (0.3)

006 P: °yes°, what were you angry about?

007 D: umm >I had got a bit< angry about >what's it called< (0.3)
 umm ((NAME)) who (0.6) eh came with proVoking °like°
 (1.0)
 references to me
 (1.4)

008 P: was it something about a presentatio[n o]r:?

009 D: [YES]

010 P: °yes° try and tell me what it was=

011 D: =well i was about to present and we had also rehearsed it
 and then he took what i was supposed to say (0.8)
 and afterwards then he said
 >i'm just waiting for her to open her MOUTH< (0.6)
 aloud in class ((voice trembling))
 umm
 (1.4)
 which e- made >me a bit< irritated

012 P: were you practically humiliated on that occa[sion]?

013 D: [YES]

014 P: °yes yeah° where- was it in front of the WHOLE class and the teacher?

015 D: uh-huh (0.8)

016 P: and how did you react to it. (0.7)

017 D: yeah that was- the:n then i flipped the table

018 P: (0.3) °yes° were you also umm: (0.5) otherwise (0.5)
 reactive towards ((NAME)) or [>towards one of the others<]

019 D: [NO: not at all]
 I quickly exited afterwards

((the speaker from The Danish Broadcasting Corporation explains that two of the defendant's girlfriends went after her and that a teacher came out to talk to her as well))

Time interval: 6.442-8.266 minutes

020 P: and the (0.3) uh: TEACHER you then talked to AFTERwards
who was that.

021 **D:** umm ((NAME)) ((voice trembling))
 022 **P:** was it ((NAME SURNAME))? (0.7)
 who is coming to give testimony today also.
 023 **D:** [yeah]
 024 **P:** [yes] okay.
 did you then go home after having talked to ((NAME))
 025 **D:** yes she then drove me home and said THAT i could (0.3)
 start (0.4) over again tomorrow
 i felt much better after having talked to ((NAME))† (.)
 and then >i then chose to be with< a couple of friends (.)
 after it happened (0.4)
 026 **P:** mm
 027 **D:** >where we then just made lasagna and enjoyed ourselves
 and i thought that it was over< (0.6)
 028 **P:** yeah did you talk about anything specific? (0.5)
 029 **D:** we had talked about many things and we had also talked a bit about
 li:ke school shootings and things like that (.)
 in the other countries (0.4)
 030 **P:** mm (0.7) why did you get to talk about that.
 do you remember that?
 031 **D:** yes so we were like online (0.4)
 032 **P:** [⁰yes⁰]
 033 **D:** [and] then we saw like some ARTI-ARTICLES
 (1.1)
 034 **P:** yes (0.8) where online did you see that?
 035 **D:** facebook
 036 **P:** yes (0.5) and then around ten ((PM)) you went home. (0.4)
 037 **D:** yes
 038 **P:** yes. and then what happened?
 (1.0)
 039 **D:** then i played computer until it was twelve o'clock
 (1.0)
 and then i had to go to BED† (0.5)
 and i had taken that picture when i was with them
 040 **P:** mm
 041 **D:** where i sat on the TOILET† (0.3) ((forceful voice))
 042 **P:** mm
 043 **D:** just like for FUN ((forceful voice))
 044 **P:** [mm]
 045 **D:** [and] >we had taken pictures together and things like that< (0.4)
 umm and then i went home and thought that i would post a picture (0.3)
 046 **P:** [mmm]
 047 **D:** [then] i took that picture where i sat on the toilet
 048 **P:** mm
 049 **D:** and so then i wrote and then-th >i had to find something to write<
 and that became it then.
 (1.2)
 050 **P:** what did you mean by it?
 (1.0)
 051 **D:** i meant like, it was something li:ke (0.7)
 especially for THOSE two basically (.) ⁰that⁰ (0.3)
 >they would be able to understand it because we just talked about it<
 (.) so it was like <a joke> (0.5)
 052 **P:** so when you <write> >i'll be the< next <school shooter guys>
 lmao† watch out.
 053 **D:** yes (0.3)
 054 **P:** just so we all understand it
 [what] does lmao stand for
 055 **D:** [yes] uh laughing my ass off ((said in English))
 056 **P:** yes umm >laughing my ass off< ((said in Danish))
 057 **D:** [YES EXACTLY]
 058 **P:** [yes yes] (0.3)
 059 **D:** which makes it a joke.

((the speaker explains that the defendant uploads the picture so that all her followers can see it))

Time interval: 8.316-9.246 minutes

- 060 **P:** this outburst uh that you had
have you HAD more of those.
(0.7)
- 061 **D:** no::
- 062 **P:** would you- uh: would you say that you sometimes get immensely angry (.)
- 063 **D:** ye:s i can get a bit angry
- 064 **P:** yes (.) and of course ((Danish: jo)) what I am interested in uh
that is of course ((Danish: jo)) kinda how the others perceive you
and ho-how YOU somehow maybe KNOW that the others perceive you
(0.7) uh: and yes according to count ONE appendix seven
it's a police report (0.7)
is it tru- can you remember that you've spoken with the police?
- 065 **D:** YES
- 066 **P:** yes umm it says on page four in the report
the subject then explained that the subject has PROBLEMS WITH HER MOOD
(1.0)
the subject can get immensely ANGRY when she gets angry
(1.0)
can you uh: have you said anything like that to the police?
- 067 **D:** yes (.)
- 068 **P:** °yes° is that also how it was. (0.5)
- 069 **D:** yes
- 070 **P:** yes ((sharp voice))
>has there been other incidents where you got angry in class<
and perhaps could have [<been a little>] (0.4) [angry?]
- 071 **D:** [NO::] [nonono<]
i haven't been angry in class at all or anything (.)
- 072 **P:** NO (.) so this was the first time this
- 073 **D:** yes it w- (0.5)
- 074 **P:** >THANK YOU fn-for now i have no further questions for the defendant<

((the journalist explains that it is now the defense lawyer's turn to question the defendant and that she says the post was a joke so the defense lawyer wants to hear more about the humor and language online))

Time interval: 11.531-12.355 minutes

- 075 **L:** you ha- you:- >i think you've explained<
that they could have a DARK sense of humor? (.)
- 076 **D:** yes (0.5)
- 077 **L:** uh try and explain who- what could that wha-what can THAT BE
- 078 **D:** uh we can write things like that uh
we will kill ourselves the next day (.)
and then we write lol or l m a o (.) aft-afterwards (.)
as if like >of course we won't do that< it's a joke, (0.3)
- 079 **L:** yeah
(1.6)
SO when you wri- when you express L M A O (0.7)
uh you then mean what (0.6)
- 080 **D:** that it's a JOKE↑
- 081 **L:** that it's a joke.
- 082 **D:** yes
- 083 **L:** i-is that something everyone underSTANDS? (.)
i [me:an]
- 084 **D:** [all] teenagers and (0.6)
<people> who are online un- should understand it,
- 085 **L:** basically all those (.) who who see (0.3) [your] posts (0.4)
- 086 **D:** [yes] yes
- 087 **L:** they know what it me[ans]
- 088 **D:** [yes]

((the speaker explains that one of the defendant's classmates felt unsafe seeing the instagram post))

Time interval: 12.396-13.075 minutes

- 089 **L:** the girl whose MOTHER (0.4) we can see (0.3)
contacts ((THE SCHOOL REPRESENTATIVE)) in the morning allegedly
was it one of these here?

090 **D:** no it wasn't any of my frie:nds it was someone from the class (0.7)
 who didn't (0.5) KNOW me ^oso well^o
 091 **L:** but but uh does SHE contact you at night [or in the morning]?
 092 **D:** [no not at all]
 093 **L:** no
 094 **D:** there's no (0.8) like no [not her]
 095 **L:** [>there's no contact<]
 096 **D:** [no]
 097 **L:** [be]tween you and her in any way
 (0.9)
 do you know if she has SEEN[†] your post
 098 **D:** yes we were FRIends on instagram i saw hers and she also saw mine
 099 **L:** yeah

((the speaker explains that the defendant got other kinds of reactions to the
 instagram post namely friends who continued her joke))

Time interval: 13.126-14.214 minutes

100 **L:** you GET some reACTIONS on this post isn't that true,
 101 **D:** yeah (0.4)
 102 **L:** uh try and just try and explain what it- what it says THERE
 103 **D:** yes he has written can you school shoot at my school then? x d (0.5)
 and xd that is like writing (0.5)
 uh it's like a smiley (.) that's laughing (0.7)
 so that p- is part o:f forexample lol and these things (.)
 104 **L:** a-and when HE says that
 can you school shoot at MY school [also] X D SMILEY
 105 **D:** [YES] YES
 106 **L:** >what THEN does he MEAN by THAT. he doesn't mean it seri[ously]<
 107 **D:** [no]
 then it's just like for fun
 108 **L:** >for fun he says it< (0.6)
 you said before that (.) there were also OTHERs who
 <wrote like you> on THESE >medias<
 109 **D:** yes (0.5) yes <lots of other> you know on FACEbook
 where in groups- where they write that they will DO: the sa:me
 and ^othings like that^o
 110 **L:** for fun.
 111 **D:** for fun.
 112 **L:** you have explained to the poLIce (0.4)
 you are uh CITED saying
 that it was supposed to be understood as <a> JOKE
 or a s- or sarcastically (0.4)
 >and then you write< (0.3) >or then you say< (0.3)
 all who know the subject (.) knows that the subject
 >would never THINK of< <HARming ANYone>
 113 **D:** yes
 (1.2)
 114 **L:** a-and why do you say- why do you mean that? (0.4)
 i mean what do you BASE this UPON?
 115 **D:** uh i'm basing it UPON that people know that i (.)
 would never think of doing such a thing
 and that i'm just a sensitive little girl ((voice trembling))
 116 **L:** yea:h

((the journalist explains that the school representative gives her statement in
 court about having received an e-mail from a worried parent but she does not want
 to be recorded))

Time interval: 16.043-16.260 minutes

117 **P:** subject reg- imPORTant please call A S A P that is as soon as possible
 at telephone number, regarding a student from ((NAME OF CLASS))
 who went aMOK in CLASS yesterday during a presentation (0.5)
 she has subsequently posted a: >my story< ((said in English)) (.)
 where she writes i'll be the next school shooter.

((the journalist reports parts of the witness statement of the school
 representative og then explains that the defendant's teacher is the next witness))

Time interval: 18.435-19.417 minutes

- 118 **T:** and she she's sitting like holding (.) herself right?
and she says that she's (0.8) very very SORRY
- 119 **P:** did she say that she was- f-felt put down by someone?
- 120 **T:** YES (0.4)
- 121 **P:** ^ookay^o (.) then I HAVE just one more thing
the defendant SAID that she had had a lot of NASTY thoughts
about that student that she (.) felt had put DOWN the defendant
the defendant said that she wasn't going to FOLLOW these thoughts
>but would rather take it out on< an object.
have you [said] that to the police
- 122 **T:** [YES] YES
- 123 **P:** okay. (.)
>when asked how the defendant is as a student
the subject describes her< as a LOVELY girl
uh who very much wants all the RIGHT and GOOD things
- 124 **T:** YES
- 125 **P:** the defendant is a FRAGILE girl who can get angry
if others don't underSTAND what she MEANS (0.3)
the defendant SAYS that she has angry and nasty-nasty thoughts (.)
she doesn't want to be someone (.) who has such thoughts
- 126 **T:** YES
- 127 **P:** is that also what uh you've ex[plained]
- 128 **T:** [YES]
- 129 **P:** (0.5) does that ONLY relate to that day?
o[r: (.) how should i understand it?]
- 130 **T:** [hhhhhhhhhhhhhhhhhhhh ((loud sigh))]
(0.9)
well really I havn't had any experiences with her (.)
uh:: (0.3) that is bad experience ON THE CONTRARY (.)
umm: i really haven't (0.5)
- 131 **P:** okay.
(2.9) ((turning of pages))

Time interval: 19.417-19.512 minutes

- 132 **L:** you had said to the police that you were sure (0.3)
that uh (0.8) the defendant wouldn't carry out her threats
- 133 **T:** YES i [am c-]ABSOLUTELY SURE that she would NEVER do that
- 134 **L:** [()]
nope

((the journalist says that it is time for the closing statements))

Time interval: 20.039-22.191 minutes

- 135 **P:** we KNOW from the witnesses and also from herself
that there actually have been some incidents and
uh: that can be more or less serious depending on how you look at it
but we do IN FACT know that the day BEFORE (0.7)
in class was an violent incident.
we also know (0.5) that she threw or flipped a table in ra:ge
or in (.) frustration
and that she (0.4)was in a very BAD state (0.6)
and that was the LAST the w-uh persons in the CLASSroom saw of her
(0.3) and some of these persons (0.5)
she must of course ((Danish: jo)) to some degree have realized (.)
probably would go and look
when fourty FIFTy would normally look at her posts (0.6)
and that she then on that same night (0.3) put up a post where she wrote
i'll ((said in English)) (.) uh i'll be the next <school shooter> guys
(0.4) we also know that uh the teacher was
familiar with her having NASTY thoughts
she also had a FEAR that these NASTY thoughts would be REalized (0.5)
and because of that she would AIM it against objects
(1.3)
it is ACTUALLY WELL-KNOWN that uh school shootings HAPPEN other places
(0.3) we all know that (0.6)
but tha:t the reason that i say that is NOT

to uh accuse uh: the defendant OF wanting to DO this here (0.5)
 as such i don't think she would do that EITHER
 that's not deCIIsive for the conviction either (.)
 in reality there is of course ((Danish jo)) actually no NEED
 in order for uh the defendant to be convicted here
 for anyone a:ctually to be scared
 it only needs to be FIT to creATE this serious fear
 for life health or welfare (0.7)
 but there was (0.6)
 also ((THE SCHOOL REPRESENTATIVE)) took it quite seriously
 >and before we knew it< (0.3)
 the police had arrived (0.9)
 can one then (.) do away with all this by writing L M A O (0.8)
 can one be TOTALLY sure that everyone is thinking
 this is just for fun? (0.3)
 yes there ARE some (0.6) uh: who have <another> (0.6)
 insider uh: immediate understanding of THIS (.)
 when they write something to each other
 but she hadn't limited it to them (0.9)
 when all that's (.) said
 i'm really trying to sum this UP and recommend
 that the pro- that the court (0.7) establishes
 that she is to be convicted in accordance with the indictment (.)
 and found guilty of violation of section twohundredandsixtysix
 (1.3)
 this is really something that AFFECTS a lot of people (.)
 it is something that has had a BIG effect on the SCHOOL (0.5)
 and it is also somethi:ng that has a SIGnalling effect (0.3)
 uh that which (0.4) uh becomes the reACTION from the court toDAY
 (1.4)
 in light of THESE facts (.) then it is my recommenDation to the court
 >that this must be< settled wi:th a
 shorter term of imprisonment, of twenty to thirty days

((the journalist sums up the prosecutor's closing statement and explains that the time has come for the defense lawyer's closing statement))

Time interval: 22.378-24.448 minutes

136 L: >the a- the defendant still pleads< not guilty (0 .6)
 and i umm am ((Danish: jo)) in a NUMBER of (0.3)
 esSEntial points in DISAGREEMENT with the prosecutor
 uh- one has to interpret <the formulation itself> (.)
 AND (.) >one has to< take into considerATIOn (.) in what CONTEXT (.)
 it is stated. (0.4)
 it is STATED in- on social media (.) on instagram (0.4)
 it ISn't STATED (0.9) towards the school, (0.3)
 it is stated (.) uh in a forum (0.3)
 where they SPEAK to each other and use a:n- a (0.3)
 >yeah well almost< for: those who DON'T know it (0.7)
 which apparently is ((THE SCHOOL REPRESENTATIVE)) (0.5)
 they then almost SPEAK in a CODED LANGUAGE, (0.7)
 and if you're NOT <CAPABLE> (0.3)
 of uh- of reading and understanding that <conversational form, (0.6)
 then of course ((Danish: jo)) you're not really capable of ASSESSING
 whether we're talking, about a THREAT, or a joke. (0.3)
 >and then you can say< it can be a: STupid joke
 (1.3)
 it can be a <RECKLESS> joke (0.6)
 >you can't convict that after twosixtysix< (0.6)
 BUT but you have to (0.3) understand (0.4)
 what L M A O MEans in this <CONTEXT> (0.4) it is s- uh stated, within
 to the (0.4) uh young people to the (.) TEENAGERS
 who ALL, communicate in the same way (0.3)
 and when they use that EXPRESSION (.)
 and it's ((Danish: jo)) important that the court UNDERSTANDS this (0.6)
 then it MEANS (0.6)
 i'll be the next <SCHOOL shooter> (0.8) just for fun (0.8)
 and it's clear (.) >that if you don't< UNDERSTAND this. (0.7)

then of course ((Danish: jo)) it can be a PROblem↑ (0.5)
 but that is of course ((Danish: jo)) not (0.5) your honors↓ (.)
 in (.) the criminal code's section tohundredsixtySIX' san-sense (.)
 it's not supposed to be <held against> the DEFENDANT
 (1.0)
 because the SECOND article (0.4) in TWOSixtySIX is (0.7)
 that the DEFENDANT has had inTENT (0.6) to PRODUCE a threat
 (1.1)
 a-and that's not an article that can be inflected (0.6)
 one can't REALIZE AFTERwards that this was too STUPID (.)
 or (0.3) <this could have been> something or other
 or SOMEONE could PERHAPS understand it-
 (1.1)
 with regards to (0.4) if your honors it must come to
 there being a CONVICTION↓ in the present case (0.3)
 then-then it must ABSOLUTELY be the court's most lenient SENTENCE

((the journalist explains that the defendant gets the final word in front of the judges))

Time interval: 24.509-25.058 minutes

137 **J**: is there anything you wish to say before we go outside (.)
 and discuss the case?
 and come back and make a decision?
 138 **D**: uh (0.7) a little (0.7)
 139 **J**: here you go
 140 **D**: umm uh i would just say that i never meant it as a THREAT
 (0.5) and that i'm of course SORRY about ^owhat has happened^o

((imellem 140-141 forklarer journalist at dommer og domsmænd voterer og kommer ind igen))

Time interval: 25.236-25.474 minutes

141 **J**: it is adjudged that (.) the defendant ((NAME)) is to be acquitted. (0.6)
 the state treasury must pay the costs of the case (0.5)
 take your seats
 (3.1)
 and i can say to you that uh we agree that you are to be acquitted
 (0.9) no:t because (0.3) we're not talking about an (0.8)
 objectively seen illegal threat but because we believe
 that uh you didn't have the required intent. uh: (0.5)

Disguise and imitation of language style in WhatsApp messages

Meike M. de Boer ¹, Susan A. Blackwell², and Fleur van der Houwen³

¹Vrije Universiteit Amsterdam, m.m.de.boer2@vu.nl

²Utrecht University, s.a.blackwell@uu.nl

³Vrije Universiteit Amsterdam, f.vander.houwen@vu.nl

Abstract

The deliberate manipulation of evidence is routinely considered in most branches of forensic science, but in the field of forensic authorship analysis it is a relatively new concern. This paper reports the findings from an experiment which explored what happens when authors are asked to manipulate their language style. We asked 120 Dutch students to write WhatsApp messages in their normal style and in two manipulated conditions: disguise and imitation. Results show that authors are able to introduce major changes in their language use, which are most substantial when they have access to messages written by the person they are trying to impersonate. These findings imply that the possibility of disguise and imitation should be considered when conducting forensic authorship analysis, especially when more than one candidate for authorship could have had access to the same device or account.

Keywords: Authorship analysis, WhatsApp messages, forensic stylistics, disguise, imitation.

Resumo

A manipulação deliberada de provas é uma prática corrente na maioria dos ramos das ciências forenses, mas no domínio da análise de autoria é uma preocupação relativamente recente. Este artigo apresenta os resultados de uma experiência que explorou o que acontece quando se pede aos autores que manipulem o seu estilo linguístico. Pedimos a 120 estudantes holandeses que escrevessem mensagens no WhatsApp no seu estilo normal e em duas condições manipuladas: disfarce e imitação. Os resultados mostram que os autores são capazes de introduzir grandes mudanças no seu uso da linguagem, que são mais proeminentes quando têm acesso a mensagens escritas pela pessoa que estão a tentar imitar. Estes resultados implicam que a possibilidade de disfarce e imitação deve ser considerada aquando da realização de análises forenses de autoria, especialmente quando mais do que um candidato à autoria pode ter tido acesso ao mesmo dispositivo ou conta.

Palavras-chave: Análise de autoria, mensagens do WhatsApp, estilística forense, disfarce, imitação.

1. Introduction

Trying to avoid detection is inherent in criminal behaviour: manipulation of the evidence is therefore a reasonable possibility in virtually any forensic investigation (Ekblom, 1999). Within the context of linguistic evidence, manipulation has so far been mostly considered in the analysis of speech (voice disguise, see e.g. Cambier-Langeveld, 2016; Eriksson, 2005) and threat assessment (Simons & Tunkel, 2021). In the context of authorship analysis, although the possibility of deception has been acknowledged in the literature (Dern, 2008; Marko, 2017), it is mentioned only briefly, if at all, in discussions of cases or procedures. Over the past few years, however, the present authors have been consulted in authorship cases where manipulation could have played a role. Messages had been sent from an alleged murder victim's phone after their disappearance, and the police wanted to know whether the disappeared or deceased persons had sent the messages themselves or whether someone else might have written them¹. According to Grant (2013), in such cases it is possible to indicate the most likely author by comparing the disputed message(s) to the language style of both candidate authors. We were provided with messages written by the victim together with messages written by one or more suspects and compared these to the disputed message(s). In some cases we found no linguistic evidence that the suspect had written the message(s), which were consistent with the language style of the victim. However, we had to insert the caveat that if anyone had imitated the victim's language style, our conclusions would be invalid.

The current study investigates how lay people change their language style in a WhatsApp message when they are asked to disguise their own language style or to imitate that of another person. In order to investigate this, we carried out an experiment in which participants acted as their own controls, producing WhatsApp texts firstly in their own personal style and subsequently under two manipulated conditions (i.e. disguise and imitation).

2. Authorship analysis

For short texts such as WhatsApp messages, a stylistic or qualitative approach is preferred over a stylometric or quantitative approach (Grant, 2013). While stylistic approaches to authorship analysis have been criticized for being subjective or prone to 'cherry picking', since the selection of features tends to be dependent on the analyst (for a discussion see Solan, 2013), the number of WhatsApp messages in forensic cases is often insufficient to adopt a stylometric approach (see e.g. Brennan & Greenstadt, 2009; Juola, 2012). The stylistic approach focuses on the choices made by the author, either deliberately or unconsciously. These have been termed 'style-markers' (McMenamin, 2002: §11.4) or 'possible textual dualities' (Olsson, 2009, p.57) and are features that can be written in more than one way, e.g. abbreviations, symbols, numbers, and (mis)spellings. Together, these are said to form an author's idiolect (Coulthard & Johnson, 2007).

Grant (2013) tried to adopt a more systematic approach, focusing specifically on mobile phone text messages. His method is data-driven, based on a word list contain-

¹For the verdicts of the court and appellate court in one of these cases, see: *Rechtbank Midden-Nederland* (15 December 2015). ECLI:NL:RBMNE:2015:8972
Gerechtshof Arnhem-Leeuwarden (24 February 2017). ECLI:NL:GHARL:2017:1544.

ing all word forms present in the messages of two candidate authors. Word forms are selected as a feature for the analysis if they show a sufficient degree of *consistency* and *distinctiveness*. In order to be consistent, a feature must be present in the messages at least ten times. To be distinctive, the feature must be twice as prevalent in the messages of one candidate author than in the messages of the other. These criteria make it possible to work with relevant features – the word forms that are selected for the analysis are similar to the style-markers that are found without a word list – while objectivity is preserved: the selection does not depend on the analyst and would be the same in replications. However, the approach still requires a decent amount of text per author in order to find enough occurrences of linguistic features, which is often unavailable in forensic casework.

The comparison of style-markers is based on the idea that the choices authors make in their texts are distinctive from other authors' choices. However, it is not possible to distinguish one author from all other possible authors in the world. Grant (2013) instead proposes the notion of 'pairwise distinctiveness', which holds that it is possible to identify the most likely author from a small number of candidates, e.g. the victim and one or more suspects. This, in addition to other evidence, can help the police or judge to discriminate between scenarios (Coulthard & Johnson, 2007). This means that the notion of the 'linguistic fingerprint' or 'stylistic genome', as idiosyncrasy in language has been dubbed (see e.g. Bernhardsson, Correa da Rocha, & Minnhagen, 2009; Luyckx, Daelemans, & Vanhoutte, 2006), is flawed and misleading (cf. Coulthard, 2014).

The implicit assumption of authorship analysis is that the variability of texts written by the same author (within- or *intra*-author variability) is less than the variability of texts written by different authors (between- or *inter*-author variability). There are, however, factors that increase within-author variability: language is continually adapted to fit the circumstances (Nolan, 2001). For example, authors adapt their language style depending on the recipient, topic, context, medium, and genre of the text (e.g. Grant, 2013; Olsson, 2008, 2009; Shapero & Blackwell, 2012). One way in which within-author variability could be artificially increased is by changing features to disguise one's language style (Olsson, 2008), which could lead the analyst to conclude incorrectly that the text is not written by the same author as the reference texts. Authors might also be capable of consciously decreasing between-author variability by imitating someone else's language style (De Vel, Anderson, Corney, & Mohay, 2001), which could lead the analyst to conclude that the text is written by the imitated person instead of the imitator.

3. Disguise and imitation of language style

Scholars within the field of forensic phonetics have long been considering disguise (e.g. Amino, Makinae, & Kamada, 2018; Künzel, 2000; Masthoff, 1996; Perrot, Preteux, Vasseur, & Chollet, 2007). While Eriksson (2005, p.10) describes most voice disguise techniques as rather unsophisticated, he also acknowledges that they can have a 'considerable detrimental effect on speaker identification'. Disguise is estimated to play a role in 15 (Künzel, 2000) to 40 (Clark & Foulkes, 2007) percent of forensic speech investigations, and in the view of Schilling and Marsters (2015) should always be considered as a possibility.

The same would appear to hold true for manipulation of language style in written texts. The literature discussing this issue is very scarce, especially from a stylistic viewpoint (but see Dern, 2008; Marko, 2017). Moreover, each publication seems to create its own terminology to describe the phenomena (e.g. disguise, circumvention, obfuscation, masking, imitation, mimicking), showing a lack of uniformity in the scientific debate. In the current paper, we use the terms disguise and imitation to refer to the two experimental conditions, which are approaches in which an author can attempt to manipulate their language style to deceive the addressee and/or the police – thereby trying to avoid apprehension (cf. Simons and Tunkel 2021, p. 245). The remainder of this section describes some of the few works available on this topic.

Most of the work on disguise and imitation of language style has been done in the stylometric tradition. Computational systems perform at around chance level when attributing disguised texts to authors, and attribute imitated texts mostly to the person who is imitated (see e.g. Brennan, Afroz, & Greenstadt, 2012; Brennan & Greenstadt, 2009). However, although authors might not be individually identified when they manipulate their language style, the deceptive intent may be recognized as a separate style by stylometric systems (Juola, 2012).

Dern (2008, pp. 243-244), discussing manipulation from a stylistic viewpoint, distinguishes four forms, which she based on “common-sense”: 1) arbitrary changes, without a particular pattern or apparent strategy; 2) ‘dumbing down’, i.e., feigning a lower language proficiency, level of education, or social status; 3) feigning non-nativeness; and 4) imitation of the language style of a specific person. According to Dern, arbitrary changes are the easiest way to disguise one’s language style since it does not require much consistency and is not used to evoke a certain image in the reader. Imitation, on the other hand, is described as the most difficult for authors to perform accurately, since it requires a high level of metalinguistic awareness both of one’s own language style and of the style of the person one is imitating. This is corroborated by MacLeod and Grant (2017), who show that for police officers posing as children in undercover operations, imitation is difficult – even after extensive training. A further study by Grant and MacLeod (2020) found that such undercover officers have more success with imitation at the level of lexis, spelling and punctuation than with the pragmatic layer of online messages.

Most of the research so far has focussed on Dern (2008)’s third category, feigned non-nativeness, prompting the participants to adopt a specific disguise strategy. Overall, although authors may be successful in disguising their own language style using this strategy, the act of manipulation has been described as detectable. According to Fobbe (2014), authors who are asked to simulate a non-native speaker identity introduce similar patterns of morphological and syntactic errors, suggesting a recognizable language style in deliberate error production. Dern (2008) states that feigned non-nativeness can be distinguished from genuine low language proficiency, since authors unconsciously show their language proficiency and native skills by their correct use of idioms and choice of words. Moreover, authors who consciously introduce errors into their texts tend to be inconsistent in the type and number of errors they use. Thus, a high occurrence of errors does not necessarily mean that the author cannot be a native speaker, and rather than focussing on individual errors, authorship analysts should examine the pattern of errors throughout the text. Marko (2017) found that whereas people expect

others to use a simulated non-native language style as a disguise strategy, her German and English-speaking participants rarely did so, instead making arbitrary changes such as deliberate spelling errors, changes in register, and changes in the purported identity of the author, such as using 'we' instead of 'I'. Indeed, when not prompted to feign a non-native identity, the changes made by Dern (2008)'s participants were superficial, mostly consisting of random deliberate spelling errors.

A recurring remark in the literature is that there is a lack of knowledge about the manipulation strategies that are used by authors in real life (e.g. Bredthauer, 2013; Dern, 2008; Marko, 2017; Seifert, 2010). Therefore, Bredthauer (2013) studied such strategies in forensic texts investigated by the Bundeskriminalamt (BKA, the German Federal Police Office), which led to a few findings regarding manipulated language styles. First, she found that manipulated texts are typically shorter than authentic texts, possibly because of the higher mental concentration required. Second, 76 per cent of the manipulated texts showed parts written entirely in capitals or entirely without capitals: Bredthauer believed this could be used by the authors to hide their true language proficiency². Thirdly, only 23 per cent of the texts contained punctuation that was 'predominantly correct', whereas 36 per cent of the authors used punctuation in a way that was 'totally incorrect', e.g. by not using any commas. Although the many punctuation errors may indicate that they are easy to manipulate due to high metalinguistic awareness, Bredthauer also suggests that they could equally be a feature of the authors' genuine language styles. Finally, errors in orthography (spelling), morphology, syntax, and lexis were inconsistent, and no pragmatic errors were observed. Overall, Bredthauer (2013) concluded that all manipulation strategies used in the forensic texts have the potential to be deceptive, although the manipulation may be indicated by inconsistencies. It is especially difficult to manipulate several different linguistic parameters simultaneously in a consistent manner over time (Kniffka, 1993): authors tend to revert to their usual style, especially at the end of longer texts (Bredthauer, 2013).

This paper attempts to supplement the meagre research on the manipulation of language style. It describes an experimental study of Dutch messages sent via WhatsApp, which is currently the most popular social media platform in the Netherlands, being used by 80 per cent of the population above the age of 15 (Van der Veer, Boeke, & Peters, 2017). The aim of this study is to compare participants' own, non-manipulated language styles as well as their manipulated styles, rather than merely describing features of manipulated language. In addition, this is the first study on manipulation of language style in Dutch, or as far as we know in any language other than German or English.

²The possible relationship between correct use of capitalization and higher language proficiency might be typical for German, since in German, all nouns start with a capital letter and other word classes do not.

4. Method³

4.1. Participants

The experiment in this study was conducted among 120 students of an introductory course in forensic linguistics at Vrije Universiteit Amsterdam in September 2017. The participants, who came from a variety of academic disciplines, had just commenced the third year of their bachelor's degree and could be regarded as lay people in the field of authorship analysis. Most participants (88.2%) indicated that they had Dutch as their sole native language; eleven (9.2%) were raised bilingually. The participants used their own smartphones to participate in the experiment, using their own preferred keyboard settings. 55 (46.2%) indicated that they used auto-correct (automatic correction of spelling) and 29 (24.4%) used predictive text (automatic suggestion of the next word). Participants had a free choice as to whether to attend the tutorial in which the experiment was conducted and could drop out of the experiment at any time. This led to the exclusion of one participant. There were six different tutorial groups, which were placed in different experimental conditions.

4.2. Materials and procedure

The experiment consisted of two scenarios with short writing tasks. After signing the consent form, the participants received written instructions for the first scenario: they were told to imagine that they had just changed their weekend plans and were asked to communicate this to their mother and a friend via WhatsApp (for each of which we provided a phone number). Scenario 1 was the same for all groups and functioned as a control task, intended to elicit the participants' non-manipulated language styles. After completing this task, the participants received their instructions for scenario 2, in which they had to imagine that a friend (given the pseudonym "Lieke", which we will use henceforth to designate this fictitious person) had died in dubious circumstances while staying with them. They were instructed to send WhatsApp messages from Lieke's phone to her mother and/or friend to tell them that she was going on a camping trip and to cancel their appointments, making it look as if she was still alive. Scenario 2 functioned as the experimental phase and was intended to elicit WhatsApp messages in which the participants manipulated their language style.

While scenario 2 started in the same way for each group, the instructions differed on two parameters: (1) the overall manipulation strategy and (2) the recipient. Participants were instructed either merely to disguise their own language style, or to imitate the language style of Lieke. Some groups were instructed to send messages in both conditions. Regarding the recipient, participants who were placed in both the disguise and the imitation condition were asked to send messages either to Lieke's mother or to her friend "Romy". Participants who were only placed in one condition (either disguise or imitation) were asked to send messages to both recipients, in alternate orders. This design had the purpose both of restricting the number of messages each participant had to send, and of counterbalancing any priming effects due to the sequencing of the tasks (see e.g. Gaito, 1961). Table 1 provides an overview of the different tasks performed by each of the groups.

³A full account of the methodology, along with a second experiment to determine how convincing the manipulations were, is described in De Boer (2018).

group	<i>First manipulation</i>		<i>Second manipulation</i>	
	Strategy	Recipient	Strategy	Recipient
A	disguise	mother	imitation	mother
B	disguise	friend	imitation	friend
C	imitation	mother	imitation	friend
D	disguise	mother	disguise	friend
E	imitation	friend	imitation	mother
F	disguise	friend	disguise	mother

Table 1. Tasks of the experimental phase per group.

Thus, each of the participants sent four messages in total: two in scenario 1 to provide their genuine language style (one to their mother and one to their friend), and two in scenario 2, under different conditions.

Participants who were asked to imitate Lieke's language style received three pages of a WhatsApp conversation between Lieke and her mother and/or 2.5 pages of a WhatsApp conversation between Lieke and Romy. The WhatsApp conversations of Lieke were genuine texts between a real student and her actual mother and friend, although the names had been changed. The participants had 15 minutes to study the conversation before they had to send a message to Lieke's interlocutor in that conversation. They could take notes and keep the conversation and notes with them while sending the imitated message.

After completion of the two writing tasks, the participants were asked to fill out a survey about their demographic details and the strategies they had used in the manipulated messages. These self-reported strategies were used to determine the features for the current analysis, but will not be further discussed in the current paper (but see De Boer, Van der Houwen, & Blackwell, in prep.).

4.3. Analysis of the messages

There were 113 participants who mentioned at least one strategy after completing the disguise and/or imitation task. A coding scheme for possible strategies was drawn up using open coding. In the end, all strategies were coded as present (1) or not present (0) in the participants' answers; the answer of each participant could include multiple strategies. Frequency tables were made to show which strategies were mentioned most often. For a thorough discussion of the self-reports, see De Boer et al. (in prep.). For the ten most frequently reported strategies, excluding those that were considered too vague or subjective to analyse systematically, at least one feature was included in the current analysis of the messages (see Table 2). Each variable was recorded and analysed using SPSS and will be discussed in the remainder of this section, sorted by variable type.

Strategy	Freq.	Feature in the analysis	Type of feature
Emoticons/emojis ^a	62	Number of emojis ^b	scale
Opening	47	Type of first emoji ^b	categorical ^c
		Full opening	qualitative (string)
		Greeting	qualitative (string)
		Addressee	qualitative (string)
Punctuation	45	Punctuation	qualitative (string)
		Number of full stops	scale
		Number of commas	scale
Word use ^d	43	-	-
Writing style ^d	30	-	-
Sentence/message length	21	Number of separate text messages sent	scale
		Number of sentences	scale
		Number of words	scale
		Average sentence length	scale
Formality ^d	21	-	-
Sentence structure	18	Subject deletion	binary
		Article deletion	binary
Exclamation marks ^e	17	Number of excl. marks	scale
Closing	15	Full closing	qualitative (string)

Table 2. Top 10 strategies for manipulating the language style (N = 113).

^aAn emoticon is composed of typed keyboard characters, e.g. :-), whereas an emoji is an image, e.g. 😊 (Pavalanathan & Eisenstein, 2015).

^b This variable includes emoticons; however, most messages only contained emojis.

^c Values: 0 = N.A.; 1 = Facial emoji; 2 = Other emoji; 3 = Emoticon.

^d These features were excluded from the current analysis.

^e ‘Changing the number of exclamation marks’ was often mentioned as an additional strategy besides ‘changing the punctuation’, even in the same responses. Therefore, both were counted separately. In the discussion of the results, however, the features are grouped thematically and the exclamation marks will fall under punctuation.

4.4. Scale variables

The following quantitative variables were counted or calculated in each of the messages:

- number of WhatsApp text messages sent to complete the assignment;
- number of sentences, i.e. strings of words separated from another string of words by a full stop, exclamation mark, question mark, emoji/emoticon, line break, empty line, or the end of the text message;
- number of words, calculated using the ‘word count’ function of Microsoft Word;
- average sentence length, calculated as the total number of words divided by the number of sentences;
- punctuation features: number of full stops; number of commas; number of exclamation marks;
- number of emojis and emoticons.

For each of these variables, means of the different conditions (i.e. disguise and imitation) were compared to the control condition using the non-parametric Wilcoxon Matched-pairs test (due to the data not being normally distributed). In addition to testing whether two means differ significantly, the output of the Wilcoxon Matched-pairs

test shows for how many participants the score changed and in which direction, making it possible to see the adaptations made by the authors in their manipulated messages compared to their non-manipulated language style.

To complement the information regarding the use of emojis and emoticons, the type of the first one used in the messages was categorized as 1) emoticon; 2) facial emoji; or 3) other emoji. Whereas different types of emojis/emoticons were sometimes present within one message, the analysis of the first one was considered the most representative way of systematically analysing all messages.

4.5. Binary variables

To analyse subject deletion and article deletion, these features were coded as present (1) or absent (0). For each of the binary variables, the three conditions were compared with each other one-by-one using a chi-square (χ^2) test. For the significant results, their strength was determined using a post-hoc test which calculated the phi coefficient (ϕ). This coefficient was interpreted according to the recommendations of Huizingh (2010): 0 – 0.3 = weak; 0.3 – 0.7 = moderate; 0.7 – 0.9 = strong; 0.9 – 1.0 = exceptionally strong.

4.6. Qualitative variables

Qualitative variables, i.e., the exact formulations of the openings and closings, were treated as nominal string variables with an unlimited number of options and analysed using frequency tables demonstrating the types and number of formulations used by the participants in each of the conditions (i.e., between-author variability). The openings often consisted of different components that were also analysed separately: a greeting (e.g. 'Hi'), addressee (e.g. 'mom'), and a punctuation mark (e.g. ';'). The majority of the closings only consisted of one component (e.g., an emoji or a word), so they were not divided into subcomponents.

5. Results

5.1. Qualitative features

Of the features that were counted in the messages, only the messages in the imitation condition showed significant differences to the control messages. None of the differences between the control and disguise conditions were significant. Table 3 gives an overview of the quantitative features in the three conditions. For reference, the features of the messages of the "victim" Lieke are presented in the right column. Note, however, that the conversations between Lieke and her mother and friend consisted of semi-live chat sessions in which the interlocutors were involved in turn-taking, whereas the messages created in the experiment consisted of a single message or short unilateral stream of messages with no intention of starting a conversation.

Variable	Control		Disguise		Imitation		Lieke	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
RECIPIENT: MOTHER	(N=87)		(N=53)		(N=55)		(N=49)	
# messages	1.14	0.55	1.04	0.28	1.80	1.08	-	-
# sentences	3.54	1.47	3.66	1.52	4.11	1.75	1.29	0.65
# words	32.60	13.67	32.40	10.87	33.09	16.95	8.14	6.36
Mean sentence length	10.09	4.48	10.53	6.10	8.35	2.97	6.12	3.72
# full stops	1.51	1.31	1.45	1.23	1.18	1.36	0.20	0.46
# commas	1.20	0.86	1.23	0.93	0.87	0.92	0.08	0.28
# excl. marks	1.05	1.03	1.28	1.26	1.11	0.94	0.10	0.31
# emojis/emoticons	0.71	0.82	0.75	1.16	1.96	1.47	0.31	0.55
RECIPIENT: FRIEND	(N=94)		(N=60)		(N=61)		(N=35)	
# messages	1.30	1.17	1.11	0.61	1.54	1.06	-	-
# sentences	4.22	1.72	4.45	1.72	5.08	1.48	1.82	1.12
# words	34.02	12.89	36.33	15.31	37.77	14.49	16.60	13.07
Mean sentence length	8.81	3.86	8.83	3.95	7.62	2.70	9.08	4.90
# full stops	1.94	1.70	1.95	1.75	1.25	1.03	0.34	0.68
# commas	1.19	0.88	1.02	0.95	0.77	0.86	0.26	0.51
# excl. marks	1.62	1.43	1.42	1.23	1.89	1.08	0.37	0.60
# emojis/emoticons	1.05	1.27	0.98	1.64	1.90	0.72	0.57	0.70

Table 3. Means and standard deviations of the quantitative features of the messages. Means in bold were significantly different ($p < .05$) from the means in the control condition (see Table 4 for the statistical results).

As can be seen in table 3, there are a few significant differences between the control and imitation condition. Whereas in the control task, most participants sent one text message to complete the assignment, when imitating the victim, the number of texts increased significantly in the messages to the mother. In addition, more but shorter sentences were used in these messages. In the messages to both recipients, the number of words did not change significantly. In the imitation condition, the use of commas and full stops in the messages to the friend decreased, whereas the use of exclamation marks increased. Table 4 gives an overview of the significant results of the Wilcoxon Matched-Pairs test.

Variable	Messages to the mother				Messages to the friend			
	Z score	↑	↓	=	Z score	↑	↓	=
Number of messages	-4.267***	24	1	30				
Number of sentences	-3.217**	30	12	13	-2.441*	34	15	12
Number of words								
Sentence length	-3.727***	16	37	2				
Number of full stops					-3.248**	14	36	11
Number of commas	-2.462*	13	28	14	-2.802**	15	31	15
Number of excl. marks					-1.997*	28	13	20
Number of emojis/emoticons	-4.832***	34	2	19	-4.306***	42	8	11

Table 4. Significant results of the Wilcoxon Matched-Pairs tests, including details on the number of authors for which the feature increased, decreased, and stayed equal in the imitation condition when compared to the control condition. Note that there were only significant results in the imitation condition. * < .001; ** < .01; * < .05**

Overall, the quantitative features studied were present in the majority of the messages written by the participants (see Table 5). The largest difference between the conditions is the number of emojis: almost all (95 to 98%) of the imitated messages contained at least one emoji or emoticon, whereas in the control condition, this was only 53% (to the mother) and 64% (to the friend). The increased use of emojis in the imitation condition may have been prompted by Lieke's messages, in which emojis were used relatively often (in approximately 25% and 50% of the messages to the mother and friend, respectively).

Grapheme	Messages to the mother				Messages to the friend			
	Cntrl N = 87	Disg. N = 53	Imit. N = 55	Lieke N = 49	Cntrl N = 94	Disg. N = 60	Imit. N = 61	Lieke N = 35
Full stop	74	72	60	18	75	73	74	26
Comma	79	74	55	8	78	63	53	23
Exclamation mark	63	66	76	10	75	75	92	31
Emoji/emoticon	53	42	95	27	64	42	98	46

Table 5. Percentage of messages that included the grapheme listed in the left column (per condition and for the "victim" Lieke).

When looking at the type of symbol (i.e., emojis or emoticons) used, only the type of the first in each message was registered, to allow easy comparison between the messages. This was rarely an emoticon (in a maximum of 10 messages per condition and recipient); in the imitated messages to the mother, no emoticons were found. Of the emojis used, most were facial emojis, e.g. 😊 or 😄. These occurred particularly frequently in the imitation condition (73% to the mother and 87% to the friend). Overall, the emoji use in the imitated messages was very similar to that of Lieke, who used facial emojis in 80-81% of the cases, and 'other' emojis (e.g. a spaghetti emoji) in all other cases.

5.2. Sentence structure

Regarding sentence structure, two strategies were explicitly mentioned by the participants and thus analysed: first-person subject deletion and article deletion. Table 6 gives an overview of the percentage of messages in each condition that contained these strategies at least once.

Feature	Messages to the mother				Messages to the friend			
	Cntrl N = 87	Disg. N = 53	Imit. N = 55	Lieke N = 49	Cntrl N = 94	Disg. N = 60	Imit. N = 61	Lieke N = 35
Subject deletion	47	41	49	16	47	58	64	31
Article deletion	10	9	7	2	13	3	13	9

Table 6. Percentage of the messages in the three conditions that contained first-person subject deletion or article deletion. Results that were significantly different from the control condition are presented in bold. For reference, the percentages are also presented for the victim's messages.

In the control condition, almost half of the participants deleted the first-person subject at least once (see Table 6). In the imitated messages to the friend, this percentage

increased significantly ($\chi^2 = 4.362$, $p < .05$). Lieke used subject deletion nearly twice as often in the messages to her friend than to her mother – which may have caused the feature to have been picked up by the imitators in the messages to the friend but not the mother. The relationship between condition and subject deletion was weak ($\phi = 0.168$). Article deletion was less common in the messages: 10 and 13% in the control condition. The disguised messages to the friend showed a significant decrease in the presence of article deletion ($\chi^2 = 3.943$, $p < .05$). The relationship between condition and article deletion was weak ($\phi = -0.160$).

5.3. Openings

The participants show a high degree of between-author variation in the openings of the manipulated messages. Depending on the condition and recipient, the participants used 38 to 73 different openings. Table 7 gives an overview of the openings that were used by more than one author in the control or disguise condition.

Messages to the mother				Messages to the friend			
Control	Freq.	Disguise	Freq.	Control	Freq.	Disguise	Freq.
Hee mam!	9	Hoi mam,	7	Hee Romy!	6	Hoi Romy,	4
Hoi mam,	7	Hoi mam!	3	Hey Romy,	4	Hee Romy!	3
Hee mam,	6	Hoi mama,	3	Hey Romy!	3	Hee room!	2
[none]	5	Heee mam	2	Hi Room!	3	Hoi Room!	2
Hoi mam!	5	Hey mam!	2	Hee Romy,	2	Hoihoi,	2
Hi mam,	3	Hoi mama!	2	Hee Room!	2	[none]	2
Hey mam,	3	Hoi mama.	2	Heee!	2		
Hoi mam.	2	Mam,	2	Hoi Romy,	2		
Hee mams!	2			[none]	2		
Ma,	2			Romy!	2		
				Room!	2		
				Yo	2		
other	43	other	30	other	58	other	45
<i>Total</i>	<i>87</i>	<i>Total</i>	<i>53</i>	<i>Total</i>	<i>90</i>	<i>Total</i>	<i>60</i>

Table 7. Openings to the mother and friend in the control and disguise condition (sorted on frequency per condition per recipient; openings that occurred only once are not presented). N.B. Other openings were each used by one participant only (per condition and recipient).

Apart from the openings presented in table 7, there were openings that were used by only one author in a condition. In the control messages, 49% (to the mother) and 64% (to the friend) of the messages had such a unique opening; in the disguised messages, this was 57% and 75%, respectively.

In the imitation condition, on the other hand, only 3 of the total number of messages contained a unique opening. Most authors had imitated Lieke's openings: only 4 participants used an opening that was not used by Lieke in the messages provided to them (see Table 8). One participant, who used 'Hey lief!', could be detected in her attempt at imitation, since she imitated the wrong person: Lieke's mother calls Lieke 'lief' ('sweetheart') in all four opening sequences, whereas Lieke did not use it once.

Opening	Mother		Friend	
	Imitation	Lieke	Imitation	Lieke
[none]	1	1	0	2
Hey,	25	2	14	2
Hey!	26	1	45	2
hey!	0	0	1	0
Hey mam,	2	0	0	0
Hey lief!	1	0	0	0
<i>Total</i>	55	4	60	6

Table 8. Openings in the imitation condition compared to the openings used by Lieke in the materials provided to the participants.

A closer inspection of the three possible elements of the openings, i.e., greeting (e.g., hey, hi), naming the addressee (e.g. Romy, Room), and punctuations (e.g. , or !), still shows considerable between-author variation in the control and disguise conditions. Per condition and recipient, there are 15–33 varieties of greetings and 12–20 ways of naming the addressee (see Table 9 for an overview). As there was only minimal variation in the imitation condition, where only two types of greetings were used and the addressee was only rarely mentioned, the verbal part of the openings of imitated messages are not further discussed.

	Opening part 1: Greeting				Opening part 2: Addressee					
	Mother		Friend		Mother			Friend		
	Cont.	Disg.	Cont.	Disg.	Cont.	Disg.		Cont.	Disg.	
[none]	9	9	14	10	[none]	15	3	[none]	25	13
Hallo	0	2	0	0	Ma	2	0	lieverd	2	1
Hee	23	4	16	9	mam	47	27	mop	3	0
Heee	0	2	12	1	Mam	3	4	Romy	29	17
Heeee	1	0	2	2	Mama	3	2	romy	0	2
Heeeee	0	0	2	0	mama	3	8	Room	15	10
Heeey	1	1	3	0	mams	6	3	room	5	2
Heey	2	2	2	0	mamsie	2	0	Rooms	0	2
Hey	13	6	14	7				schat	3	1
Hi	6	2	6	3						
Hoi	21	20	6	12						
Hoihoi	0	0	1	2						
Yo	0	0	2	0						
other	11	5	13	14		6	6		11	12
<i>Total</i>	87	53	93	60		87	53		93	60

Table 9. Overview of the verbal parts of the openings (i.e., greetings and ways to address the addressee) per condition (control, disguise) and recipient. ‘Other’ greetings/addressees were present for only one author.

Despite the overall variation in the verbal parts of the openings, some greetings and ways to address the addressees were clearly most popular: at least half of the messages contained the same few greetings or addressee parts. For instance, in the control condition, 66% of the authors opened the messages to their mother with either ‘Hee’, ‘Hoi’, or ‘Hey’, and 54% addressed their mother with ‘mam’.

When looking at the third part of the openings, the punctuation, there is less variation. 6 different (combinations of) punctuation marks were used in the control con-

dition, 9 types (incl. emojis) in the disguise condition, and 3 types in the imitation condition (see table 10). The majority of the openings (64–84% in the control and disguise condition, and 98–100% in the imitation condition) had either a comma or an exclamation mark as punctuation.

Opening 3: Punctuation	Mother				Friend			
	Cont.	Disg.	Imi.	Lieke	Cont.	Disg.	Imi.	Lieke
!	26	13	27	1	39	24	46	2
,	47	21	27	2	31	18	14	2
no punc.	10	6	1	1	14	13	0	2
.	2	5	0	0	4	4	0	0
!!	0	1	0	0	4	0	0	0
;	1	0	0	0	1	0	0	0
..	0	0	0	0	0	1	0	0
...	1	3	0	0	0	0	0	0
❤️	0	2	0	0	0	0	0	0
!:))	0	1	0	0	0	0	0	0
😊	0	1	0	0	0	0	0	0
<i>Total</i>	<i>87</i>	<i>53</i>	<i>55</i>	<i>4</i>	<i>93</i>	<i>60</i>	<i>60</i>	<i>6</i>

Table 10. Overview of the punctuation used in the openings in the different conditions and in Lieke's messages (provided to the participants in the imitation condition).

5.4. Closings

Whereas almost 98% of the messages in the dataset contained an opening of some kind, only about 70% of the messages contained a closing (60.0 – 80.5% per condition and recipient). In fact, the absence of a closing was the most popular ending across conditions and recipients. For the messages that contained a closing of some kind, emojis were quite common. The closings used by at least two participants per condition are presented in table 11.

Closing	Mother				Friend			
	Cont.	Disg.	Imi.	Lieke	Cont.	Disg.	Imi.	Lieke
[none]	17	12	11	2	37	24	24	4
😘 / 😊 *	7	5	17	1	11	3	7	0
X	4	2	0	0	6	4	0	0
❤️ / 🍷 / ❤️ *	4	2	0	0	3	1	2	0
:)	5	2	0	0	2	1	0	0
😊 / 😊 *	4	1	6	0	5	0	16	2
x	3	0	0	0	1	0	0	0
xx	3	0	0	0	2	0	0	0
xxx	2	1	0	0	2	0	0	0
Xxx	2	2	0	0	0	1	0	0
❤️❤️ *	2	0	0	0	1	0	0	0
😘😘 *	2	0	1	0	1	0	0	0
Kus!	2	0	0	0	0	0	0	0
Liefs	1	4	0	0	1	0	0	0
👉	0	0	0	0	3	0	0	0
Miss you! 😘 *	0	0	0	0	0	0	2	1
😊😘 *	0	0	5	1	0	0	0	0
🌲😘	0	0	3	0	0	0	0	0
😁 / 😁 *	0	0	3	0	0	0	1	0
👦	0	0	2	0	0	0	0	0
😊😊	0	0	2	0	0	0	1	0
other	29	22	5	0	22	26	16	0
<i>Total</i>	<i>87</i>	<i>53</i>	<i>55</i>	<i>4</i>	<i>93</i>	<i>60</i>	<i>60</i>	<i>7</i>

Table 11. Closings to the mother and friend in the condition (sorted on frequency per condition per recipient). ‘Other’ closings were used by only one participant in the dataset (per condition and recipient). Closings in the messages by the victim Lieke are provided for reference. *For these emojis, similar variants were treated as the same symbol.

In the control messages, about 30% of the participants used a unique closing (not presented in table 11), and in the disguised messages, about 47% did so. When the participants were instructed to imitate Lieke’s language style, 60% and 70% of the messages to the mother and the friend respectively contained a closing that was also used by Lieke.

As can be seen in Table 11, most closings consisted of one component only (e.g., only an emoji or a word). Hence, the different components of the closings were not analysed separately.

6. Discussion

In this study, we investigated the manipulation of language style in WhatsApp messages. Whereas prior research mainly focussed on the effect of feigned non-nativeness (Dern, 2006, 2008; Fobbe, 2014), in line with the findings of Marko (2017), the participants in the current study did not report to have feigned a non-native identity (cf. De Boer et al., in prep.). Neither did our participants, on the whole, report that they had introduced errors into their manipulated messages, whereas earlier publications often

focused on deliberate error production⁴. This implies that research investigating manipulation of language style should rather focus on strategies adopted spontaneously by authors, as was done in the current paper. The strategies investigated in this study overlapped strongly with features of language use that are described in the literature as ‘textual dualities’ (e.g. Olsson, 2009), showing that stylistic authorship analysis may be vulnerable to manipulation.

The exploitation of textual dualities was most striking in the imitation condition. Whereas people in both experimental conditions introduced changes into their language use, without access to the victim’s messages the changes seem less substantial. For example, most quantitative features of the messages in the disguise condition did not differ significantly from the control messages. Note, however, that this may be partly explained by the participants changing their language use in different directions: some may over-use certain features when compared to their control messages, whereas others may under-use them. This study looked at aggregated features and averages, rather than changes made in the experimental conditions when compared to individual authors’ control messages. It is possible that some individuals were more successful at either disguise or imitation than others: this could be the topic for further research based on our data.

A qualitative investigation of changes across conditions per author could shed more light on the nature of the changes and the extent to which certain linguistic features were manipulated in both experimental conditions. What the current analysis did show is that the messages in the disguise condition showed as much variation as the messages in the control condition, whereas the imitated messages were more similar to each other and were also similar to the victim’s language use. This suggests that it might be easier for lay people to change their language use when they have access to messages written by the person they are impersonating. This finding contradicts Dern (2008)’s assertion that imitation requires the highest level of metalinguistic awareness and is therefore the most difficult task for authors who are trying to manipulate their language use. We suggest that metalinguistic awareness is in fact heightened when comparison material is available and the author is confronted with alternatives to their own language use. Whereas most choices in writing are usually made unconsciously (Bayer, Dal Cin, Campbell, & Panek, 2016; Kniffka, 1993), when other people make different choices, this may draw the writer’s attention. This may be especially true for WhatsApp messages, where authors are less constrained by genre-specific conventions compared to e.g. letters (as used in Dern’s experiment) and show high between-author variability. In addition, WhatsApp messages are typically short and may be largely built up by elements (e.g. opening, closing) that can covertly be copied from earlier messages.

According to prior literature, manipulation of language style can be detected because authors are not able to maintain a consistent language style (as suggested by Bredthauer, 2013; Dern, 2006; Grant & MacLeod, 2018; Juola, 2012; MacLeod & Grant, 2017). A conventional way to detect manipulated language is to consider the consistency of features throughout a text. Since manipulation takes up a lot of effort and energy, it is claimed to be difficult to sustain (Bredthauer, 2013; Fobbe, 2014; Kniffka, 1993). Our data make it hard to assess this claim, since the texts concerned were ex-

⁴Errors were not one of the ten most frequently mentioned manipulation strategies, meaning that it was mentioned in <15 responses in the entire database, and were hence not analysed for this paper.

tremely short: the messages were on average 30 to 40 words in length. Although the beginnings (i.e. openings) of the messages were indeed more often imitated than the endings, the short messages produced by the participants may not be long enough to lose their concentration as found in previous studies. Still, this situation is reflective of a certain type of forensic casework that we have encountered, where there may be only one or two final messages sent from the victim's phone. This implies that imitation in such cases may be easier than so far established from experimental studies.

Overall, we can say that the manipulated messages did not share similarities that could make them stand out as deceptive – at least not on the linguistic features analysed in this study. In the disguise condition, most quantitative features did not display significant differences from the control messages; unlike Bredthauer (2013)'s findings on actual manipulated forensic texts, in our study disguised messages were not shorter (see also Marko, 2018). Of the features that changed, such as the openings and closings, there was still as much between-author variation in the disguised messages as in the messages that were not manipulated. In the imitation condition, the authors did become more similar to each other, but this was always in the direction of the victim's language style. Hence, the shared features between the imitated messages are not signs that the manipulation could be easily detected, but merely that the imitations could have been effective.

Whether the imitated messages are actually effective in misleading a recipient or even forensic linguist is a question that remains unanswered. The comparison between the messages in the imitation condition and the messages of the victim cannot be made one-to-one, as the conversation handed to the participants to study the victim's language consisted of multiple exchanges of messages in which the interlocutors were both involved in the conversation simultaneously. This means that for some features, such as emojis and subject deletion, it appears that the participants may have 'over-done' it – a feature known in the sociolinguistic literature as hypercorrection (e.g. Decamp, 1972), but it may in fact be proportional considering the high information-density of the messages constructed by the participants for the experiment. In addition, the knowledge that the victim may use a certain feature at least once in their authentic messages may be enough for both imitators and analysts to conclude that the feature fits their linguistic repertoire.

The current study should be considered as an initial exploration of the field, with limitations concerning the external validity of the results. Firstly, the diversity and representativeness of the participants in the experiment was limited. The participant group was very homogeneous on demographics such as gender, age, and education level. In addition, they did not represent the relevant population, namely suspects in police investigations for abduction and murder, who are most often men aged between 25 and 44 (Centraal Bureau voor de Statistiek, 2018) with vocational rather than academic education. Metalinguistic awareness may be higher for our participants and it may be easier for the young women in the current study to imitate another woman of a similar age. In fact, the control messages of the participants may already show a high degree of similarity to the messages of our victim. In actual cases, the victims may not be of similar age, gender and background, so perpetrators are expected to show less awareness of how their victims generally talk, act and relate to others. We would expect imitators who did not know the victim to have particular difficulty with the pragmatic level, as

experienced by the undercover police officers described by Grant and MacLeod (2020), and this would be fruitful material for future studies.

In addition, the tasks were different from real-life situations in a number of ways. Firstly, the participants were asked to imagine that a friend had just died, while in fact they were sitting among their fellow students in a classroom. They were not rewarded for their manipulations, while in real-life situations a lot is at stake for the person trying to manipulate their language style. In real cases, authors probably put more effort into their act of deception and spend more time studying the victim's messages and formulating their manipulated messages than the students in this study. In fact, they may have even known their victim and have exchanged messages with them for years, creating familiarity with their language use. On the other hand, criminals in real-life situations may be faced with stressful and unplanned situations, and may not be considering language use suiting their victim's style at all. What could have further affected the results is that the participants were first asked to send a message in a non-deceptive style, which was fresh in their memory and even visible while formulating the manipulated messages. This gave the participants an example of how they would write the message themselves, which is expected to have led to increased metalinguistic awareness of their own style.

7. Conclusion

The goal of this study was to investigate how lay people change their language use in WhatsApp messages when asked to disguise their own language style or to imitate someone else's. The findings show that most authors introduce major changes in their messages on a whole range of linguistic and stylistic features. Their strategies did not involve introducing deliberate errors or feigning non-nativeness; rather, they paid attention to the use of emojis and emoticons, punctuation, and openings and closings in their messages. Moreover, our participants displayed a reasonably accurate degree of awareness of the style of the person they were trying to imitate and impersonate. This study shows that the manipulation of language style is an important topic that should be investigated further. In particular more research is needed on the different strategies pursued by individuals attempting to disguise their own language or imitate that of someone else.

Based on this study, it is strongly recommended that in any authorship case the possibility of manipulation should be considered. The linguistic expert should take into account factors such as the level of education of any possible suspect(s), how well they knew the missing/deceased person and the extent to which they might have had access to texts produced by that person. In some cases this might lead to a report which states that the findings can be interpreted in more than one way, which could be regarded as unhelpful advice. However, acknowledging the possibility of manipulation should still be preferred over presenting an invalid conclusion. Presenting neutral conclusions might in the short term delay police investigations, but hopefully it will lead to greater acceptance of the field as a solid forensic science.

References

- Amino, K., Makinae, H., & Kamada, T. (2018). Auditory discrimination of natural speech and synthetic speech used as voice disguise. *Acoustical Science and Technology*, 39(1), 48–50.
- Bayer, J. B., Dal Cin, S., Campbell, S. W., & Panek, E. (2016). Consciousness and Self-Regulation in Mobile Communication: Consciousness in Mobile Communication. *Human Communication Research*, 42(1), 71–97.
- Bernhardsson, S., Correa da Rocha, L. E., & Minnhagen, P. (2009). The meta book and size-dependent properties of written language. *New Journal of Physics*, 11(12).
- Bredthauer, S. (2013). *Verstellungen in inkriminierten Schreiben: Eine linguistische Analyse verstellten Sprachverhaltens in Erpresserschreiben und anderen inkriminierten Texten*. Wiesbaden: Springer Fachmedien Wiesbaden.
- Brennan, M., Afroz, S., & Greenstadt, R. (2012). Adversarial stylometry: Circumventing authorship recognition to preserve privacy and anonymity. *ACM Transactions on Information and System Security*, 15(3), 1–22.
- Brennan, M., & Greenstadt, R. (2009). Practical attacks against authorship recognition techniques. In *Proceedings of the Twenty-First Innovative Applications of Artificial Intelligence Conference*. (pp. 60–65). California: AAAI Press.
- Cambier-Langeveld, T. (2016). Language analysis in the asylum procedure: a specification of the task in practice. *International Journal of Speech Language and the Law*, 23(1), 25–41.
- Centraal Bureau voor de Statistiek. (2018). *Verdachten; delictgroep, geslacht, leeftijd en herkomstgroepering* (Tech. Rep.). Centraal Bureau voor de Statistiek. Retrieved from <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/81947NED/table?ts=1529502620833>
- Clark, J., & Foulkes, P. (2007). Identification of voices in electronically disguised speech. *International Journal of Speech Language and the Law*, 14(2), 3820.
- Coulthard, M. (2014). Whose text is it? On the linguistic investigation of authorship. In S. Sarangi & M. Coulthard (Eds.), *Discourse and Social Life* (1st ed., pp. 270–287). Routledge. doi: 10.4324/9781315838502-15
- Coulthard, M., & Johnson, A. (2007). *An introduction to forensic linguistics: Language in evidence*. London: Routledge.
- De Boer, M. M. (2018). *Disguise and imitation of language style in WhatsApp Messages: An analysis of linguistic and stylistic variation in manipulated texts* (Master's thesis, Vrije Universiteit Amsterdam). Retrieved from https://www.ubvu.vu.nl/pub/fulltext/scripties/13_2121611_0.pdf
- De Boer, M. M., Van der Houwen, F., & Blackwell, S. (in prep.). Metalinguistic awareness of linguistic disguise and imitation in WhatsApp messages.
- Decamp, D. (1972). Hypercorrection and rule generalization. *Language in Society*, 1(1), 87–90. doi: 10.1017/S0047404500006552
- Dern, C. (2006). Bewertung inkriminierter Schreiben: Zum Problem der Verwischung von Spurendurch Verstellung. *Kriminalistik*, 5, 323–327.
- Dern, C. (2008). “Wenzahl nix dan geht dir schlecht”: Ein Experiment zur sprachlichen Verstellungsstrategien in Erpresserbriefen. *Zeitschrift für Germanistische Linguistik*, 36(2), 240–265.
- De Vel, O., Anderson, A., Corney, M., & Mohay, G. (2001). Mining e-mail content for author identification forensics. *ACM SIGMOD Record*, 30(4), 55–64.

- Ekblom, P. (1999). Can we make crime prevention adaptive by learning from other evolutionary struggles? *Studies on Crime and Crime Prevention*, 8, 27–51.
- Eriksson, A. (2005). Tutorial on forensic speech science. Part I: Forensic phonetics. In *Proceedings of the 9th European conference on speech communication and technology*. Lisbon.
- Fobbe, E. (2014). Fingierte Lernersprachen: Strategien der muttersprachlichen Fehlerproduktion im Dienste der Verstellung. *Zeitschrift für germanistische Linguistik*, 42(2), 196–222.
- Gaito, J. (1961). Repeated measurements designs and counterbalancing. *Psychological Bulletin*, 58(1), 46–54.
- Grant, T. (2013). TXT 4N6: Method, consistency, and distinctiveness in the analysis of SMS text messages. *Journal of Law and Policy*, 58(1), 467–494.
- Grant, T., & MacLeod, N. (2018). Resources and constraints in linguistic identity performance. *Language and Law / Linguagem e Direito*, 5(1), 80–96.
- Grant, T., & MacLeod, N. (2020). *Language and Online Identities: The Undercover Policing of Internet Sexual Crime* (1st ed.). Cambridge University Press.
- Juola, P. (2012). Detecting stylistic deception. In *Proceedings of the EACL Workshop on Computational Approaches to Deception Detection* (pp. 91–96). Avignon, France: European Association for Computational Linguistics.
- Kniffka, H. (1993). Forschungsethik. In H. Goebel, P. H. Nelde, Z. Starý, & W. Wölck (Eds.), *Kontaktlinguistik* (pp. 819–825). Walter de Gruyter.
- Künzel, H. J. (2000). Effects of voice disguise on speaking fundamental frequency. *Forensic Linguistics*, 7(2), 149–179.
- Luyckx, K., Daelemans, W., & Vanhoutte, E. (2006). Stylogenetics: Clustering-based stylistic analysis of literary corpora. In *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC'06)*. Genoa, Italy: European Language Resources Association.
- MacLeod, N., & Grant, T. (2017). “go on cam but dnt be dirty”: Linguistic levels of identity assumption in undercover online operations against child sex abusers. *Language and Law / Linguagem e Direito*, 4(2), 157–175.
- Marko, K. (2017). Strategies for disguise in written threatening communications and ransom demands: an analysis of English and German texts. *International Journal of Speech Language and the Law*, 24(2), 243–247.
- Marko, K. (2018). Underlying motivations for the use of linguistic disguise in written English and German threats and ransom demands in an experimental corpus. *Nordic Journal of Linguistics*, 41(2), 205–226.
- Masthoff, H. (1996). A report on a voice disguise experiment. *International Journal of Speech Language and the Law*, 3(1), 160–167.
- Nolan, F. (2001). Speaker identification evidence: Its forms, limitations, and roles. In *Proceedings of the conference Law and language: Prospect and retrospect* (pp. 1–19). Levi, Finnish Lapland.
- Olsson, J. (2008). *Forensic linguistics* (2nd ed.). New York: Continuum.
- Olsson, J. (2009). *Word crime: Solving crimes through forensic linguistics*. London: Bloomsbury.
- Pavalanathan, U., & Eisenstein, J. (2015). Emoticons vs. Emojis on Twitter: A Causal Inference Approach.

- Perrot, P., Preteux, C., Vasseur, S., & Chollet, G. (2007). Detection and recognition of voice disguise. In *Proceedings, IAFPA*. Plymouth, UK: International Association for Forensic Phonetics and Acoustics.
- Schilling, N., & Marsters, A. (2015). Unmasking Identity: Speaker Profiling for Forensic Linguistic Purposes. *Annual Review of Applied Linguistics*, 35, 195–214.
- Seifert, J. (2010, January). Verstellungs- und Imitationsstrategien in Erpresserschreiben: Empirische Studien zu einem Desiderat der forensisch-linguistischen Textanalyse. *Zeitschrift für angewandte Linguistik*, 2010(52).
- Shapero, J., & Blackwell, S. A. (2012). ‘There are letters for you all on the side-board’: What can linguists learn from multiple suicide-note writers. In S. Tomblin, R. Sousa-Silva, & M. Coulthard (Eds.), *Proceedings of the international association of forensic linguists’ tenth biennial conference* (pp. 225–244). Birmingham: Centre for Forensic Linguistics.
- Simons, A., & Tunkel, R. F. (2021). The Assessment of Anonymous Threatening Communications. In J. R. Meloy & J. Hoffmann (Eds.), *International Handbook of Threat Assessment* (pp. 235–256). Oxford University Press.
- Solan, L. M. (2013). Intuition versus algorithm: The case of forensic authorship attribution. *Journal of Law and Policy*, 21(2), 551–576.
- Van der Veer, N., Boekee, S., & Peters, O. (2017). *Nationale social media onderzoek 2017: Het grootste trendonderzoek van Nederland naar het gebruik en verwachtingen van social media #NSMO*. (Tech. Rep.). Enschede: Newcom Research & Consultancy.

Developing a Resource Model of Power and Authority in Anonymous Online Criminal Interactions

Helen Newsome-Chandler ¹ and Tim Grant ²

¹University College Dublin, Ireland, helen.newsome-chandler@ucd.ie

²Aston University, UK, t.d.grant@aston.ac.uk

Abstract

The purpose of this paper is to provide both a theoretical foundation and a practical framework for analysing power and authority in online interactions. This is to assist forensic linguists and law enforcement in their understanding of anonymous online criminal networks, and the roles of individuals in these online communities. The lack of contextual knowledge present in anonymous online fora creates a challenge for the analyst in finding a framework to theorise, explore, and describe different types of power performance and thus the different roles of interactants in these fora. In this paper, we provide a framework to describe the basis on which individuals make claims to power and use this framework to explore the nature and distribution of power across different fora of both criminal and benign intent. This is developed through an analysis of three online discussion fora, of approximately 160,000 words, resulting in a framework of nine main categories of power resource. This allows us to contrast the three fora, showing differences in the nature and distribution of power resource, and also enables description of individuals as high-resource or low-resource with regards to their claims to power. This theory and framework can also be productive in the analysis of language and power in computer mediated communication (CMC) more widely.

Keywords: *Language and power, Pragmatics, Forensic linguistics, Language and identity, Computer mediated communication (CMC).*

Resumo

O objetivo deste artigo é fornecer uma base teórica e um quadro prático para a análise do poder e da autoridade presentes em interações online. Pretendemos ajudar os linguistas forenses e as autoridades a compreenderem as redes criminosas anónimas e os papéis dos indivíduos neste tipo de comunidades online. A falta de conhecimento sobre o contexto, que se verifica nos fóruns anónimos online, cria um desafio para o analista, que tem de encontrar um quadro para sistematizar, explorar e descrever os diferentes tipos de manifestação de poder e, conseqüentemente, os diferentes papéis dos interlocutores nesses fóruns.

Neste artigo, apresentamos um quadro para descrever a base sobre a qual os indivíduos reivindicam poder e utilizamos esse mesmo quadro para explorar a natureza e a distribuição de poder em diferentes fóruns, tanto com intenções criminosas "malignas" como "benignas". Este processo foi desenvolvido através de uma análise de três fóruns de discussão online, com aproximadamente 160000 palavras, resultando num quadro de nove categorias principais de recursos de poder. Isto permite-nos comparar os três fóruns, mostrando diferenças na natureza e distribuição dos recursos de poder, permitindo ainda descrever os indivíduos como tendo recursos elevados ou baixos no que diz respeito às suas reivindicações de poder. Esta teoria e quadro podem também ser produtivos na análise da linguagem e do poder na comunicação mediada por computador (CMC) de forma mais alargada.

Palavras-chave: *Linguagem e poder, Pragmática, Linguística Forense, Linguagem e identidade, Comunicação mediada por computador (CMC).*

1. Introduction

One important and growing problem for law enforcement is the need for better understanding of criminal online spaces, including the highly-anonymous spaces of the dark web. One area where linguistic analysis might assist is understanding how powerful or influential individuals behave and interact within online communities. The aim of this paper is to provide both a theoretical foundation and a practical framework for analysing power and authority in online interactions, for application in forensic linguistics analysis. Specifically, our objective is to develop a framework that can be applied to the analysis of anonymous interactions of criminal online fora. However, this theory and framework can also be productive in the analysis of language and power in other domains of computer mediated communication (CMC).

We explore this by considering three online fora. The first is a non-criminal, parenting forum, which frequently discusses sensitive topics, but is freely-accessible on the internet. There is no specific provision for individuals to contribute anonymously, over and above the felt anonymity of any online social interaction. The second forum is a white nationalist discussion forum, also freely-accessible on the internet. As discussions on the site often focus on socially unacceptable topics, and can also include the discussion of criminal activities, this creates a somewhat different focus for concerns around anonymity. The final forum is a Child Sexual Exploitation and Abuse (CSEA) forum that is hosted on the 'dark web' - that portion of the internet that is heavily encrypted and requires specialist knowledge and software to access. Participants in this forum make considerable efforts to ensure their anonymity and there are discussions on the forum as to how this can be best achieved. By considering these three fora, with their similarities and differences, we hope that the theory and framework of power that we are developing will generalise to the analysis of other online contexts.

2. Theories of Language and Power

The study of language and power stretches across many social science disciplines and has a long history including in the fields of psychology (Conley, O'Barr, & Lind, 1978), sociology (Bourdieu, 1991), and linguistics (Fairclough, 1984), etc. It is, however, virtually impossible to come to a universal consensus to define exactly what power is, and this is reflected in the literature with many studies approaching the notion of power from very different perspectives. Some studies, for example, have chosen to examine how clear institutional power differences can be mapped on a linguistic level, for example, in doctor-patient interactions (Ainsworth-Vaughn, 1998), in workplace communication (Kacewicz, Pennebaker, Davis, Jeon, & Graesser, 2014), and in police interviews (Haworth, 2010). Other studies have instead focused on the relationship between language and traditional social power categories such as class (e.g. Conley et al., 1978), gender (e.g., Lakoff, 1973), and age (e.g., Blum-Kulka, 1990). This literature often sees power as a binary distinction, with one individual being more powerful in an interaction than another either because of their institutional position or because of their demographic characteristics.

These diverse literatures show that power is a notoriously slippery concept to define across linguistics and the social sciences. However, for the purposes of this paper we are interested in the basis on which individuals make claims to authority in interactions, including making claims to be heard by other participants. This situates our working definition of power as being interactional, rather than as a performance of institutional or social power. Further to this, our model of power proposes that power it is not a binary or monolithic concept, rather that in any interaction it is a resource that can be drawn upon by some individuals and not by others and that this access to power varies according to context. Thus, in the anonymous online spaces that we are most interested in, power may operate very differently, and be characterised differently from offline contexts or from contexts where individuals, or their institutional roles and power bases are well known.

The analysis of power has been the focus of much research in the field of pragmatics (including Bousfield & Locher, 2008; Brown & Levinson, 1978, 1987; Leech, 1983; Locher, 2004; Watts, 2003). In their seminal work on politeness theory, Brown and Levinson define power as an "asymmetric social dimension of relative power" and "the degree to which H [the hearer] can impose his own plans and self-evaluation (face) at the expense of it S's [the speaker's] plans and self-evaluation." (Brown & Levinson, 1987, 77 quoted in Culpeper, 2008, 33). Brown and Levinson's power variable has, however, been subject to some criticism. For example, Culpeper has highlighted that Brown and Levinson "say little about the type of power that might be involved in interaction" (of which there are many different types) and their study does not account for "the complexity of how power works in interaction" (Culpeper, 2008, 34). Subsequent research in pragmatics has sought to address this research gap (e.g. Locher, 2004; Watts, 2003) and the present paper seeks to contribute to this work through an investigation of the nature and linguistic performance of power in anonymous online discourse.

As such, this paper sits within the domain of interpersonal pragmatics, which explores the 'ways social actors use language to shape and form relationships in situ' (Locher & Graham, 2010, 1). In this paper, we focus exclusively on the claims to power made by interactants in their answers to requests for advice posed by Original Posters

(OP) in three online discussion fora. By making claims to power, the interactants attempt to persuade the Original Poster (and other interactants present in the fora) to listen to their point of view and potentially to act in a particular fashion. Whilst these interactional moves seek to influence the actions of the Original Poster, this is often achieved without engaging in directly face-threatening assertions.

Online discussion fora provide an interactional context where differences in institutional or social power are often obscure or indeterminate, and posters may be partially or fully anonymous. This lack of knowledge of the institutional and social power structures in these contexts can create a challenge for the analyst in finding a framework to theorise, explore, and describe these sorts of power performance. However, in these contexts, it is clear that interactants still shape and form relationships by drawing on linguistic resources that enable them to perform more or less powerfully in those interactions. One aspect of relational work that becomes particularly apparent in these performances of power is the interactants' implicit or explicit efforts to establish their legitimacy to contribute to and be heard in the interaction. The aim of this paper is to develop a set of analytical techniques that may be productive in helping to categorise individuals in positions of power in online interactions. We do this by providing a framework to describe the basis on which individuals rest their claims to be listened to. Further to this, we use this framework to explore the nature and distribution of power across different online fora of both criminal and benign intent. This is demonstrated through an analysis of three different fora, as outlined above.

3. Data

To address the aims of this paper we collected three datasets. To address ethical issues of researcher wellbeing (BAAL, 2021), we first developed and tested frameworks for analysis on a benign clear web parenting discussion forum (Forum 1). By clear web, we refer to publicly-accessible internet sites, indexed by common search engines. We then moved on to analyse two forensically interesting and potentially distressing datasets. The second dataset was a clear web, white nationalist discussion forum (Forum 2), which included talk of potentially illegal activities and extreme ideologies. As such, it is blocked by some internet providers and in some jurisdictions. The third data set was a dark web CSEA discussion forum (Forum 3) in which apparent offenders discuss their offending activities and arrange to exchange abuse images and media.

This approach enabled us to spend less time with the disturbing data, whilst ensuring the generalisability and transferability of our results to these contexts. It also allows us to contrast communities with criminal and non-criminal intent, in the clear web and dark web (with different expectations around anonymity and security), and with communities with different topics of discussions and reasons for interacting.

3.1. Forum 1

Forum 1 is a well-established discussion forum, which focuses on a range of parenting and lifestyle topics. As such it has no criminal purpose and no specific ideological purpose. It contains a 'Talk' section, freely accessible to visitors without a profile but one is required to post on the site. To register, users must supply an email address and select a username but otherwise their profiles are anonymous. In this sense it is a clear web forum, but with some generic sense of anonymity for the posters. All sub-fora and threads

are open both to registered users and ‘lurkers’. Lurkers are those who visit the fora, but do not themselves post. It does have a collection of paid moderators who through their usernames are identifiable as moderators. Moderators occasionally post on threads, for example referring a user to formal sources of support and deleting unsuitable posts, but overall, they do not appear to have a particularly dominating presence on the site.

3.2. Forum 2

Forum 2 focuses on propagating white nationalist, racist ideologies. These ideologies are often disapproved in broader society, and activity on the site can include discussions of criminal activities. As with Forum 1, this site is a clear web forum and users can access some areas of the site without a registered profile. If they want to post on the site, they need to register with an email address. Unlike Forum 1, Forum 2 features specific sections reserved for higher level users. These users are often moderators (who are unpaid) or individuals who have made financial donations to the site.

3.3. Forum 3

The third dataset is a CSEA forum found on the dark web. It features extreme illegal activity, and as a result its users go to significant technological efforts (via advanced encryption software) to avoid detection. Users with sufficient technical expertise can view or ‘lurk’ on some discussion boards in the forum, but the forum requires users to have a registered username if they want to post on the discussion board. Given the high-risk nature of the site users are constantly vigilant of being infiltrated by law enforcement and take great efforts to ensure their anonymity and protection. Hierarchies of power and exclusivity are present in the forum. For example, to gain access to more exclusive and higher-level discussion boards users may be required to provide novel or extreme child abuse media. Possession and ability to distribute such material is therefore an extralinguistic marker of power in this forum. Moderators act as gatekeepers to these areas and unlike Forum 1 and Forum 2 will also advise users and delete posts if their contributions may potentially leak unnecessary information about themselves and their activities or contribute to a breach in anonymity.

4. Developing a resource model of authority and power in interaction

4.1. Sampling Strategy

From prior data collection, we had access to multi-million word corpora for each of the three fora. As this project initially focused on the markup of pragmatic features, it was necessary to sample a subset of each forum for qualitative analysis and manual coding. Sampling criteria were developed which could be applied across the three datasets. We identified a series of eight prompting phrases present in titles or opening posts of threads from across the three datasets. Preliminary corpus analysis revealed that prompts such as ‘how do you’, ‘do you know’, ‘why did’, ‘what are the’, and ‘how do I’ elicit responses from users which contain claims of power. Users make claim to different sources of power to present themselves as being sufficiently authoritative in a given discussion and to give them the legitimacy to offer advice to a poster and/or contribute their perspective to a thread and thus to make a claim to be listened to in the given interaction.

To further investigate these performances of power from the identified threads, we selected a smaller dataset of 160 threads by taking the 20 most recent threads of each prompt, with each thread being initiated by a different author. As some threads were 1000s of posts long and because we coded the threads manually, it was necessary to reduce the data further. In the absence of any protocol established by the literature, we did this by taking the first 3 threads for each prompt which fulfilled the following criteria:

- threads which were 5 to 9 posts in length;
- threads which were 10 to 19 posts in length;
- threads which were 30 to 49 posts in length.

This final dataset was used to identify the different ways in which claims to power were performed across the three fora. This amounted to a corpus of approximately 160,000 words (72 threads in total) which was explored and manually coded, with roughly equal representation across the three fora. The process led to the identification of a set of power resources which are exemplified and discussed in detail below, and allows for analysis of both individual power performances and the analysis of how power is performed differentially across the three speech communities.

4.2. The resource model of authority and power in interaction

Our model of power in online interactions includes nine main categories of power and authority, some of which include two or three subcategories. These power types are described below and are exemplified from the three datasets (where such examples exist). As the power resources arise from the data we include resources that occur in at least one forum but we make no claim that they make an exhaustive list. From our exploration of the broader corpora we believe that all of these power categories are likely to occur across the three datasets even if not observed in the 72 threads analysed in the data subsample. Whilst we also believe that the resources will be found in online fora beyond the three studied in this paper, we have no expectation that they will be present across all internet discussion fora. We also expect that different fora would demonstrate different distributions of power resource. This is explored below in the community level comparison.

In the examples below we provide explanations of community specific terms or initialisms in square brackets.

4.2.1. Community Expertise

By making a claim of community expertise, users demonstrate experience of the norms of this speech community or the community as a whole. For example, this can include referral to different threads and acceptable/unacceptable practices on the site, as well as the general views and activities of the online community. Users may make a claim of **direct community expertise** through their own experience of a forum, or they may make claim of **indirect community expertise** through reference to the experience of someone they know.

Examples of direct community expertise:

- **Forum 1:** if you need advice it would be better to start a new thread (there is a start new thread link under the adoption topics banner) then people will be able to post specifically to you.

- **Forum 2:** When you go to post a picture on [Forum 2]. You have to right click and copy the image url. Then on your post up at the top is a series of small boxes. One of them has a mountain on it. You then click it. Then you paste the image url into the given spot and click on. That will post your picture.
- **Forum 3:** I've posted a lot elsewhere in the forum on why I believe pedophilia is normal ...

Examples of indirect community expertise:

No examples of indirect community expertise were found in the current data samples analysed in this stage of the project.

4.2.1.1. Community specific initialism One form of community expertise can be seen through the use of acronyms or initialisms specific to the online forum or community.

Examples of community specific initialisms:

- **Forum 1:** I informed SS [social services] as SW [social worker] said that their legal term would take weeks to update them.
- **Forum 2:** I've known people who were not WN [white nationalist] / NS [national socialists] but had issues with black crime, drug dealers and other issues.
- **Forum 3:** Fiendishly clever and elusive guys who make, share and collect terabytes of CP [child pornography] ...

4.2.2. Technological expertise

By making a claim of technological expertise, users draw on their experience or competence in a particular technical area. This is of particular relevance to the dark web forum where expertise in encryption and digital technologies is valued, but the power type can also occur more broadly to include other areas of technology. Users may make a claim of **direct technological expertise** through their own expertise of using a given technology, or they may make claim of **indirect technological expertise** through reference to the technological expertise of someone they know.

Examples of direct technological expertise:

- **Forum 1:** No examples of this power type were found in Forum 1 data sample analysed for this stage of the project.
- **Forum 2:** I saw pictures of Lake Mead and the water behind the dam almost reached the bottom. The inflow pipes were too high for the water to go in. I think engineers were going to devise additional intakes lower on the towers.
- **Forum 3:** Thankfully I've found TOR ['The Onion Router' - software which enables anonymous communication online] and truecrypt [encryption software]

Examples of indirect technological expertise:

No examples of indirect technological expertise were found in the any of the data samples from Forum 1, 2, or 3 analysed in this phase of the project.

4.2.3. Veteran power

By making a claim to veteran power, a user shows long-term membership of a forum and/or of the wider community or subject area. Users may make a claim of **direct veteran power** through their own experience of a fora or topic, or they may make

claim of **indirect veteran power** through reference to the experience of someone they know. In our coding scheme we have no specific criteria of the time frame in which users can make a claim of veteran power. Instead, we simply record when users make claim to this kind of experience.

Examples of direct veteran power:

- **Forum 1:** We have looked after quite a few [...] children over our 33 years of fostering ...
- **Forum 2:** That is really an excellent quote from [white nationalist figure], because it is so true. I have been around here at [forum name] for some time, but I don't recall hearing or reading that before.
- **Forum 3:** These chats have been around for many many years with a lot of long term members who helped sites like this one exist.

Examples of indirect veteran power:

No examples of indirect veteran power were found in the data samples from Forum 1, 2, or 3 analysed in this phase of the project.

4.2.4. *Accredited expertise*

By making a claim to accredited expertise, a user demonstrates or claims that they have official or formally recognised expertise on a given topic, e.g. 'I am a social worker', 'I am a teacher', etc. This category also includes making reference to being an administrator or moderator of a site. Users may make a claim of **direct accredited expertise** through their own accredited expertise of a given topic, or they may make claim of **indirect accredited expertise** through reference to the accredited expertise of someone they know.

Examples of direct accredited expertise:

- **Forum 1:** I work in social care and in this area, each individual assessment is unique and each family different ...
- **Forum 2:** I worked for a major law enforcement agency and we had a very generous budget for in- service training. We qualified once a month with service pistols, patrol carbines and riot shotguns.
- **Forum 3:** Neither myself, nor any of the staff, can prove beyond doubt that we're for real ...

Examples of indirect accredited expertise:

- **Forum 1:** I know two foster carers who take newborn babies. The last three children they got from birth and kept for about 12 months.
- **Forum 2:** Many soldiers I knew back in the day despised being in the military at times, comparing it to more than a life sentence!
- **Forum 3:** No examples of this power type were found in the Forum 3 data sample analysed in this phase of the project.

4.2.5. *Topic expertise through personal experience*

By making a claim to topic expertise through personal experience, a user claims personal experience of a given situation or topic. Users may make a claim of **direct topic expertise** through personal experience, or they may make a claim to having **indirect topic expertise** through personal experience if they draw on the personal experience of someone they know or are close to, for example, a work colleague or friend.

Examples of direct topic expertise through personal experience:

- **Forum 1:** We are just two and a half years in [to the adoption] and it is brilliant to see our son developing close relationships with relatives.
- **Forum 2:** I was involved with Bulgarian Nationalism for a long time. Unfortunately Bulgaria was snookered into joining the EU so I felt I had to change my focus.
- **Forum 3:** I had many dangerous situations where someone almost discovered my secret ...

Examples of indirect topic expertise through personal experience:

- **Forum 1:** Close friends of mine have just mentioned that they gave up the idea of adoption after watching us go through [adoption].
- **Forum 2:** I know a number of people who bought the military surplus in the early to mid 90's and have heard of zero exploding rounds.
- **Forum 3:** You remind me of a friend of mine who is from the UK but lives elsewhere. A very many similarities.

4.2.6. Broad topic expertise

By making a claim to broad topic expertise, a user makes no specific claim of personal experience but does however give the impression of general expertise on a given topic.

Examples of broad topic expertise:

- **Forum 1:** Adopted children need a lot of support. Social workers will ask you about every aspect of your lives and will want to speak to friends and family too.
- **Forum 2:** WWII was born in Versailles in 1919 when the "Allies" confiscated large parts of German industry and otherwise levied sanctions that would have caused mass starvation and the destruction of Germany.
- **Forum 3:** Morality is subjective and changes from culture to culture. Pedophilia is openly practiced in many cultures around the world ...

4.2.7. Private knowledge

By making a claim to private knowledge, a user uses private knowledge of a particular topic or user to make a claim to authority.

Examples of private knowledge:

- **Forum 1:** No examples of this power type were found in the Forum 1 data samples analysed in this stage of the project.
- **Forum 2:** No examples of this power type were found in the Forum 2 data samples analysed in this stage of the project.
- **Forum 3:** Just so nobody gets worried, [User X] will be offline for 6 or 7 days. She got grounded a couple of days ago and her laptop was confiscated.

4.2.8. Citing a secondary source for authority

Citing a secondary source for authority on a given topic can include direct quotations from canonical texts, secondary sources, or through quotations from authority figures.

Examples of citing a source for authority:

- **Forum 1:** No examples of this power type were found in the Forum 1 data samples analysed in this phase of the project.
- **Forum 2:** As Voltaire said it is difficult to free fools from the chains they revere.
- **Forum 3:** I read a report from the 60's which showed that ...

4.2.9. *Subject of law enforcement/investigations*

By making a claim to being a subject of law enforcement or investigations, a user claims that they have been arrested or convicted of a crime. While this power type may occur in more general discussion fora, it is more likely to occur in fora dedicated to criminal or extremist discussions. Users may make a claim of being a **direct subject of law enforcement/investigations**, or they may make a claim of **indirect subject of law enforcement/investigations** where they draw on the personal experiences of someone they know who has been arrested or convicted of a crime.

Examples of direct subject of law enforcement/investigations:

- **Forum 1:** No examples of this power type were found in the Forum 1 data samples analysed in this phase of the project.
- **Forum 2:** As Voltaire said it is difficult to free fools from the chains they revere.
- **Forum 3:** I read a report from the 60's which showed that ...

Examples of indirect subject of law enforcement/investigations:

No examples of indirect subject of law enforcement/investigations were found in the data samples analysed for Forum 1, 2, and 3 in this phase of the project.

5. Applying the power resource model

After developing the resource model of power outlined above, we applied it to the three datasets to answer the following research questions:

1. **Are power resources used differently across the three separate fora?**

In this section, we explore what is generic across the three fora, how power is exercised differently between the fora, and whether the rates of features in each forum are different.

2. **Are some individuals drawing upon power resources differently from other individuals in interactions?**

In this section, we contrast individuals who draw upon multiple power resources with those who draw upon a small number of power resources. Analysis of these features allows us to describe hierarchies of power for individuals in each of the three fora.

With the size of the current three data samples we restrict ourselves to qualitative observations and explanations, guided by some preliminary statistical comparisons.

5.1. Community level comparison

For the three tagged fora, the count of power resources were normalised by number of tokens of the smallest forum (Forum 2) and an initial comparison of use was created using the graph in Figure 1.

A cursory examination of Figure 1 shows clearly an uneven distribution of power resources across the three fora. For example, broad topic expertise is drawn upon more frequently in Forum 2 but direct personal experience appears to be more common, and thus more important, in Forum 1 and Forum 3. This was explored further by producing three hierarchy tree maps to more clearly show the relative distribution of power resources within the three fora (Figure 2).

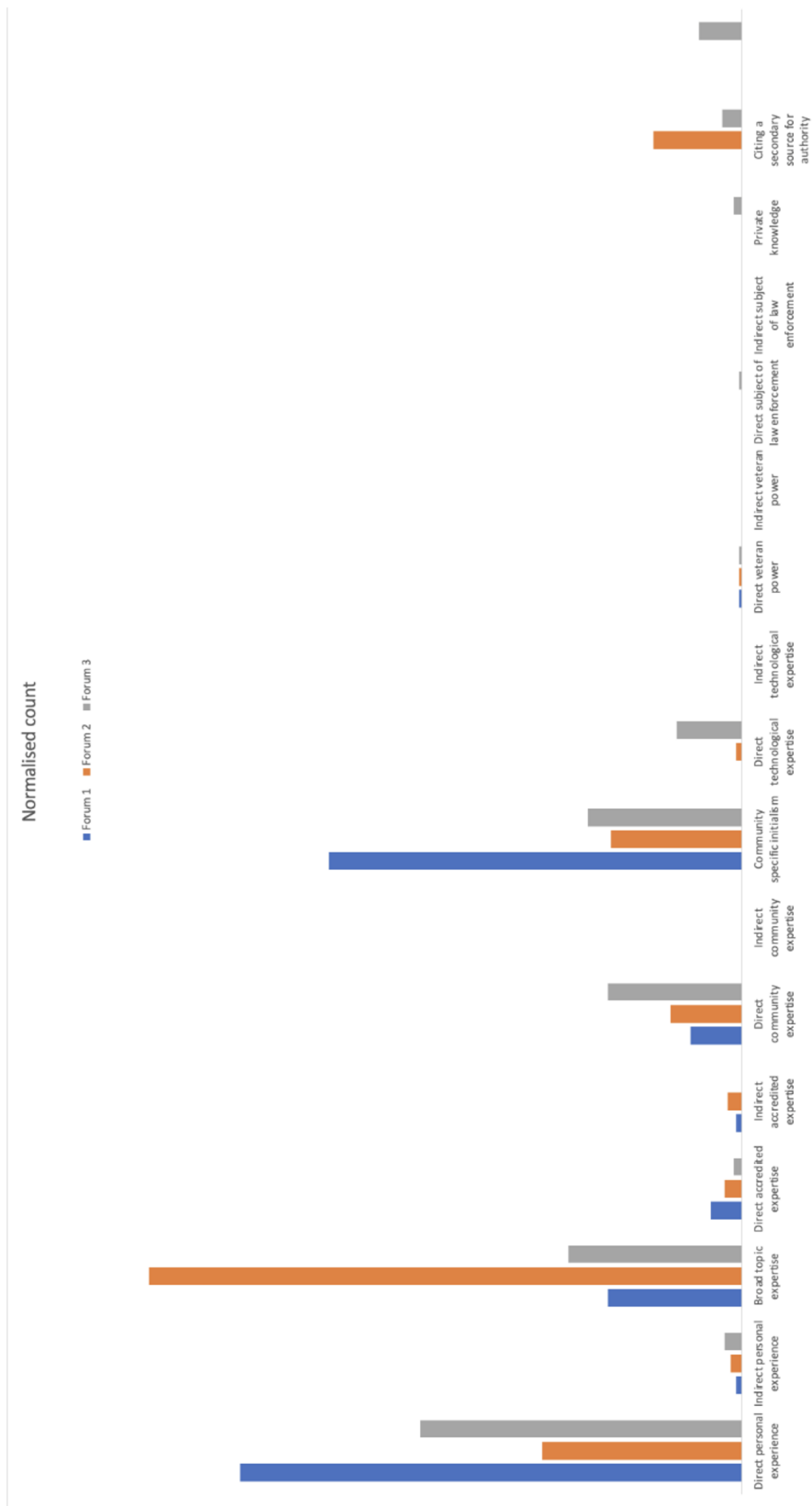
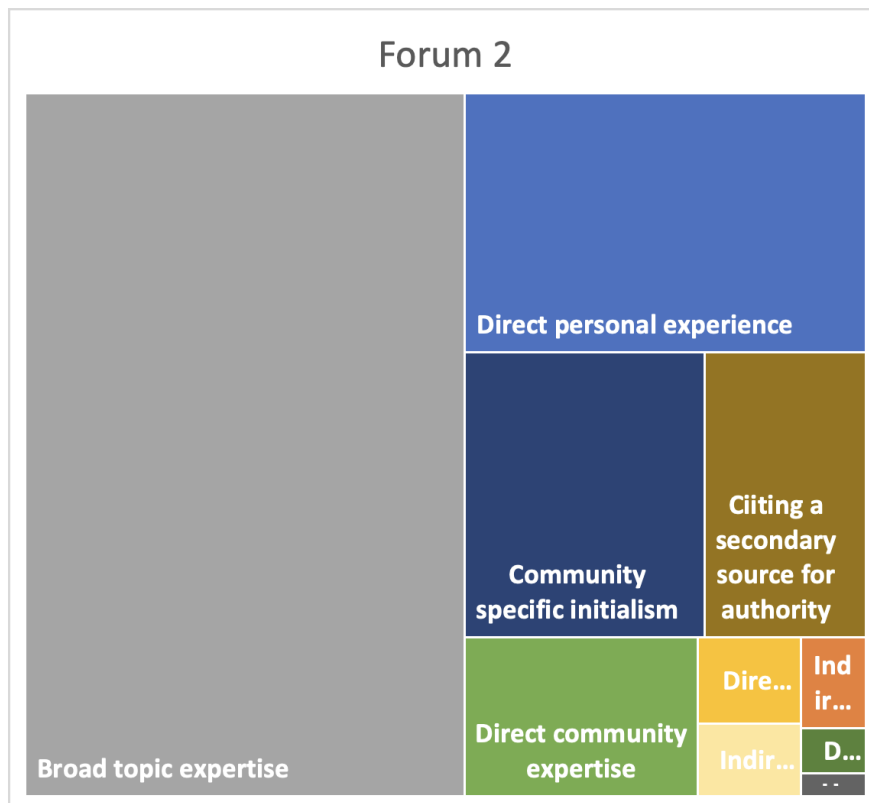
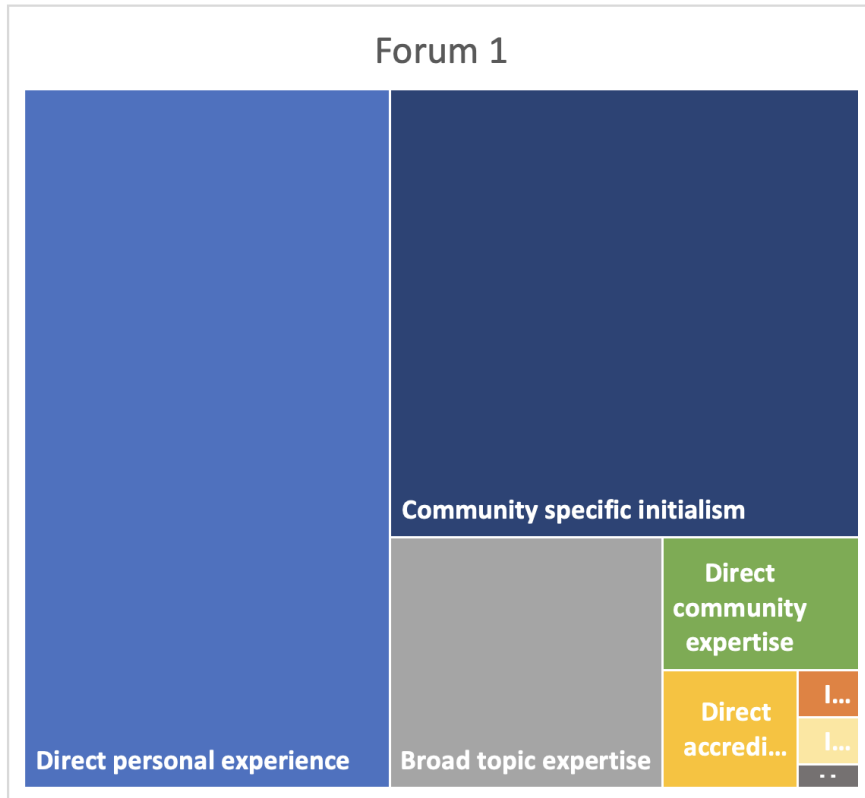
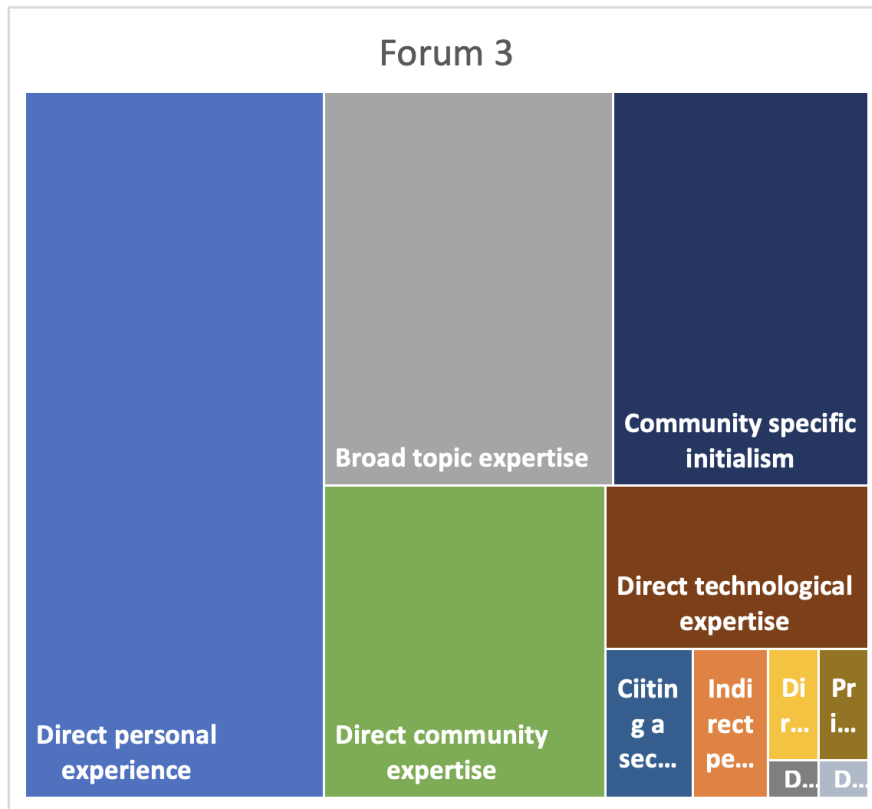


Figure 1. Normalised counts of power resources across the three fora samples.





- Direct personal experience
- Indirect personal experience
- Broad topic expertise
- Direct accredited expertise
- Indirect accredited expertise
- Direct community expertise
- Indirect community expertise
- Community specific initialism
- Direct technological expertise
- Indirect technological expertise
- Direct veteran power
- Indirect veteran power
- Direct subject of law enforcement
- Indirect subject of law enforcement
- Private knowledge
- Citing a secondary source for authority

Figure 2. Hierarchy tree maps of the distribution of power resources across the three fora.

These visualisations show the similarities between Forum 1 and Forum 3 in terms of the relative proportions of direct personal experience, broad topic expertise, and community specific initialisms. In contrast, it can be seen that in Forum 2, citing a secondary source for authority is a frequent power resource. This power resource is less prevalent in Forum 3 and does not occur in Forum 1. These observed similarities and contrasts were tested using a series of χ^2 analyses, as can be seen in Table 1. NB: As the datasets were of roughly equal size, the counts of tags were not normalised between fora.

Power Resources	Forum 1	Forum 2	Forum 3	chi2	chi2 p	phi	Effect
Broad topic expertise	37	208	27	228.3	< 0.001*	0.92	large
Community specific initialism	114	46	24	71.8	< 0.001*	0.62	large
Citing a secondary source for authority	0	31	3	51.6	< 0.001*	1.23	large
Direct personal experience	138	70	50	49.5	< 0.001*	0.44	medium
Direct technological expertise	0	2	10	14.0	< 0.001*	1.08	large
Direct community expertise	14	25	21	3.1	< 0.001*	0.23	small
Direct accredited expertise	9	6	1	6.1	0.028	0.62	large
Indirect accredited expertise	2	5	0	insufficient data for chi2 analysis			
Indirect personal experience	2	4	3	insufficient data for chi2 analysis			
Direct veteran power	1	1	0	insufficient data for chi2 analysis			
Private knowledge	0	0	1	insufficient data for chi2 analysis			
Indirect community expertise	0	0	0	insufficient data for chi2 analysis			
Indirect technological expertise	0	0	0	insufficient data for chi2 analysis			
Indirect veteran power	0	0	0	insufficient data for chi2 analysis			
Direct subject of law enforcement	0	0	0	insufficient data for chi2 analysis			
Indirect subject of law enforcement	0	0	0	insufficient data for chi2 analysis			
Counts normalised by number of tokens in smallest sample (34,587 tokens in Forum 2);							
* sig at p < 0.003 (15 comparisons with Bonferroni correction).							

Table 1. χ^2 comparison of tagged fora

A Bonferroni correction for multiple comparisons was applied meaning that the threshold of $p < 0.003$ was used as a significance threshold and in addition phi was used as a pseudo effect size measure. Against these criteria, Table 1 shows that for seven power resource types there was a non-random distribution across the three fora. In addition to the four non-significant results, there were six contrasts for which there was insufficient coding to carry out the χ^2 analyses. These results, along with further examinations of Figures 1 and 2, were used to guide the qualitative interpretations.

5.1.1. Interpretation of Results:

Broad topic expertise was significantly more prevalent in Forum 2 than in the other two fora. Many of the discussions within this forum are general ideological and political discussions on topics such as neo-nazism, racial superiority, and political responses to national events. In this context, claims to authority centre around political knowledge. Interactants demonstrate their expertise through interactions.

A related code to broad topic expertise is **citing a secondary source**. Again, this is more prevalent in Forum 2. It occurs to a lesser extent in Forum 3 and not at all in Forum 1. The citation of political texts and figures and other sources demonstrates a deeper general knowledge of the white nationalist political movement beyond the bounds of the specific forum and presents an intellectually sophisticated and authoritative persona. In Forum 3, the sources cited differ and vary between popular and scientific sources, which suggest that paedophilia can be normalised as sexuality or behaviour. Demonstrating knowledge of these sources gives the user authority in the interaction.

Direct personal experience was prevalent across all three fora but was used more in Forum 1 and then slightly less in Forum 3 and least of all in Forum 2. Topics in Forum 3 and Forum 1 revolved more around personal problem solving and support seeking. Respondents in Forum 3 and Forum 1 thus often draw upon their own direct personal experience of similar situations to offer advice and reassurance to the Original Posters. This form of authoritative but face-protective answering has been observed in other sensitive advice-giving contexts, for example Sillence, who observes it in half of advice-giving posts in an online breast cancer support group. She notes that “[p]ersonal experience messages offer the advice seeker an insight into the choices and thought processes of a ‘similar’ person.” (Sillence, 2013, 483). It provides a means of claiming authority to be listened to, which can be mitigated by the expectation that the interactant will form their own view. As such it leverages Cialdini (2001) principles of persuasive communications drawing on reciprocity and social proof as well as authority.

Community expertise is prevalent in all fora but manifests itself differently across the three datasets. For example, direct community expertise is significantly more prevalent in Forum 3. Given the highly illegal nature of this forum, users are constantly vigilant to potential infiltration by undercover law enforcement agents. Demonstrating community expertise is thus important to ensure that users are accepted into the forum and regarded without suspicion. Users can make a claim to authority by demonstrating their expertise in community rules and norms. Community expertise in Forum 1 is largely realised through community specific initialisms. This has the pragmatic effect of demonstrating longer term membership and thus engenders trust and authority. Whilst community expertise and initialisms do occur Forum 2, they do so at a combined lower rate than the other two fora. This might be explained through the very high rate of broad topic expertise in Forum 2 which on this site is understood as a more general form of community expertise that might extend beyond the online community.

Technological expertise is very prevalent and important in Forum 3. Because the illicit communication in this forum requires knowledge of advanced security technologies (such as TOR browsers and VPNs), demonstration of such knowledge gives a user considerable authority in this community. Technological expertise does occur in Forum 2 but at a considerably lower rate. Claims of technological expertise in Forum 2 function as a performance of general expertise of a given topic (for example, expertise in weapons systems, water drainage systems, and survivalist technologies). This contrasts with the claims of technological expertise made in Forum 3 which are generally restricted to the discussion of online security technologies.

After Bonferroni correction **direct accredited expertise** was not significant ($p = 0.028$) but as the effect size was large ($\phi = 0.62$) we further explored it qualitatively, noticing that it was highest in Forum 1, with some occurrence on Forum 2 and with

very little use in Forum 3. In Forum 1, this code mostly picked out professionals (such as social workers, foster carers etc.) and this was reflected in Forum 2 where military and ex-military roles were disclosed. No such professional identities were observed in our Forum 3 data, however a form of accreditation that was observed was that of moderator or admins who used these roles to assert authority and to police the posts of other users. Claims of this sort were not found in the data samples consulted from Forum 2 and Forum 1, but they are used elsewhere in both sites.

5.2. Powerful and powerless individuals

In our analysis of the claims of power across the three datasets, we observed that users can make multiple claims of power in a given thread. Through drawing on a greater number of power resources, some users present as being more powerful than others. We can thus see the establishment of hierarchies of power within the three fora. In the below section we exemplify these hierarchies through the qualitative analysis of a sample thread from each forum. The sample thread from Forum 1 focusses on a discussion of adoption practices and is 2970 words in length. The sample thread from Forum 2 focusses on a discussion on the nature of racism and is 3782 word in length. The sample thread from Forum 3 focuses on the discussion of the nature of different sub-fora within the site and is 1130 words in length.

For each forum, we contrast a user who demonstrates a broad repertoire of power resources and makes multiple claims of power with users who seem to draw on limited resources.

Forum 1

	No. of posts in thread	Broad topic	Personal experience (direct)	Personal experience (indirect)	Community expertise	No. of resources drawn on
User 1	2					0
User 2	1	x			x	2
User 3	1		x			1
User 4	1		x			1
User 5	1		x		x	2
User 6	1		x			1
User 7	1		x		x	2
User 8	1		x		x	2
User 9	3	x	x	x	x	4
User 10	1		x		x	2
User 11	1	x				1
User 12	4		x		x	2
User 13	1		x		x	2
User 14	1		x			1
User 15	1	x				1

Table 2. Forum 1 distribution of resources by user

User 9 in this thread draws on the broadest repertoire of claims to power. They make use of both **indirect** and **direct personal experience** on the topic of discussion, for example, they use “I know a couple who decided ...” and they report their own direct expertise through phrases such as “In my experience ...”. User 9 also demonstrates

community expertise through the use of **community specific initialisms** throughout their posts. In one of their posts, User 9 also demonstrates **broad topic expertise** through phrases such as “Yes, it’s not the norm but I have seen it happen.” in relation to a comment made by another user. In contrast to User 9, User 1 makes no attempt at claiming power or authority, their posts are limited to asking further questions of the group. User 3 makes a single claim to authority through demonstrating **direct personal experience** of the topic. They signal this experience through phrases such as “I have a similar [experience] to your husband...”.

In contrast to User 9, User 1 makes no attempt at claiming power or authority, their posts are limited to asking further questions of the group. User 3 makes a single claim to authority through demonstrating **direct personal experience** of the topic. They signal this experience through phrases such as “I have a similar [experience] to your husband...”.

Forum 2

	No. of posts in thread	Broad topic	Personal experience (direct)	Personal experience (indirect)	Community expertise (Direct)	Technological expertise	Citing a secondary source	Veteran Power	No. of resources drawn on
User 1	2	x							1
User 2	1		x		x				2
User 3	1	x							1
User 4	2	x							1
User 5	1	x			x				2
User 6	1	x			x				2
User 7	1	x							1
User 8	2	x	x						2
User 9	1								0
User 10	4	x			x		x	x	4
User 11	2	x							1
User 12	1	x	x		x				3
User 13	1	x			x				2
User 14	3	x					x		2
User 15	1								0
User 16	1		x						1

User 17	1	x						1
User 18	1		x					1
User 19	1		x					1
User 20	1	x						1
User 21	3	x	x		x			3
User 22	2		x		x		x	3
User 23	1							0

Table 3. Forum 2 distribution of resources by user

In the Forum 2 thread, User 10 makes claim to four separate power resources. User 10 makes regular use of **citation of secondary sources**, citing both a prominent speech of a white nationalist leader, and a political philosopher. They then go on to demonstrate other forms of power including **broad topic expertise**, **community expertise** through use of **initialisms**, and **veteran power** by reporting their experience from “the early days of the [community]”. There are also three users who each make claim to three power resources to make claims of authority in this thread. For example, within the same post, User 22 writes ‘That is really an excellent quote from [White Nationalist figure], because it is so true. I have been around here at [Forum 2] for some time, but I don’t recall hearing or reading that before.’. In this post, User 22 demonstrates **direct personal experience** of the topic and they also make claim to **veteran power**. In the same post, User 22 demonstrates **community expertise** with the use of **community specific initialisms**.

In contrast, three users in the thread do not make any claims to power in their posts. They simply offer comments on the posts of others made in the thread. Ten users in the thread make only a single claim to authority, most commonly **broad topic expertise** as highlighted in the between fora comparisons above.

Forum 3

	No. of posts in thread	Broad topic	Personal experience (direct)	Personal experience (indirect)	Community expertise (Direct)	Technological expertise	No. of resources drawn on
User 1	2						0
User 2	2		x		x		2
User 3	2		x		x		2
User 4	1		x		x	x	3
User 5	2		x		x		2
User 6	1						0
User 7	1	x			x		2

Table 4. Forum 1 distribution of resources by user

In this thread, User 4 draws on three separate power resources to claim authority. This includes claims of **direct personal experience** of a topic, as exemplified through phrases such as “I have also been involved in several private forums ...”. They also demonstrate **direct community expertise** through the use of **community specific**

initialisms and through references to their own participation in the online community. User 4 also shows **technological expertise** through the use of technical jargon.

User 1 and User 6 make no claim of expertise. Their contributions to the thread simply include requests for links to online abuse media and requests for details of users' experiences on similar fora. Four users make claim to two sources of power in the thread. For example, User 2 provides **direct personal experience**, using terms such as "I was a member of ...". User 2 also demonstrates **direct community expertise** through use of community specific initialisms and displaying general knowledge of the forum.

5.2.1. Interpretation of Results

Overall this analysis of individuals and how they make claims to power supports the idea that those who can access multiple power resources within an interaction perform generally more powerfully than those who have limited resources to draw on. By making a greater number of claims of power in these fora, these users make a greater claim to be listened to in the online community. Contrasting high resource individuals with low resource individuals within and between communities shows that those who can draw on multiple power resources are rare (see Table 5). It is not obvious in the interactions studied here that this stronger claim is recognised by the other discussants but follow up studies involving bigger datasets will attempt to examine this, through examining how interactants orient to these high resource individuals. What has become apparent through this analysis though is that the assertion of authority is performed more strongly through the use of diverse claims to power.

	Forum 1	Forum 2	Forum 3
Users who make no claims of power	1	3	2
Users who make claim to 1 source of power	5	10	0
Users who make claim to 2 sources of power	7	6	4
Users who make claim to 3 sources of power	0	3	1
Users who make claim to 4 sources of power	1	1	0

Table 5. Forum 1 distribution of resources by user

6. Conclusions

Although this study is preliminary and is based upon the coding of a limited dataset (72 threads, comprising 160,000 words in total) it is already possible to see that this is a fruitful approach to understanding the operation of power in and across different fora and by the interactants within those communities.

By developing a resource model of performed power in interaction we are able to observe the similarities and differences between these online communities. In response to our first research question - *Are power resources used differently across the three separate fora?* - our analysis demonstrates that cultures of power manifest very differently across the three fora studied in this paper and this thus highlights that power is not a monolithic construct.

To directly address our second research question - *Are some individuals drawing upon power resources differently from other individuals in interactions?* - our analysis shows that we need to move away from the idea of powerful and powerless individuals, to

recognising that an individual may be able to access more or fewer power resources in any given interaction and thus be able to perform more or less powerfully. Their access to power resources may be constrained by their own history or experience, or by the nature of the interaction in terms of topic or content. Some individuals may have a broad repertoire of power resources to draw on, but they may not be relevant for use in a specific interaction or community. The idea of a powerful individual is thus reformulated to be an individual who has access to multiple power resources and finds themselves in an interaction in which there are opportunities to use them. An individual with a greater repertoire of power resources is thus more likely to be able to perform powerfully within any given interaction.

Our model thus far has identified nine main power resources across the three communities (with some additional subcategories), but we make no claim that this is a closed set or even an exhaustive list for these fora. What we have shown is that there are a recurring set of power resources that can be found across the three different online communities with very different interests and cultures. This highlights the transferability and potential generalisability of our resource model of power in interaction. Further work on the model requires coding of larger datasets and to achieve this we are exploring semi-automated coding of the power resources. This will open up analyses of a range of different fora and the opportunity to pursue different research agendas, such as investigating how interactants orient to high resource individuals, or how individuals' claims of power develop over time.



Our working definition of power, as noted in the introduction, highlighted that in our online contexts, an interactional definition of power would be more helpful than an institutional or social definition of power. Further to this, we also noted that power, both in definition and operation, varies considerably by context. In this study, the most notable contextual factors are that the data are from online interactions and that the participants are, to varying degrees, anonymous. As noted in the introduction, the study of power in CMC cannot depend upon correlations with traditional institutional or social categories of power but must mostly arise from analysis of the linguistic content of the interaction. This is true of most CMC interactions but becomes particularly important in forensic studies of highly-anonymised dark web fora. Our approach accounts for Culpeper's observations that there may be many different types of power at work in interaction, which can intersect in complex ways. By focussing on the resources that individuals can draw on in interaction to assert their claim to be listened to, we further contribute to this burgeoning body of pragmatic research in a way that can be particularly useful in CMC analyses. Furthermore, by conceptualising individuals' access to power in interaction as part of a resource model, our theory aligns with current writing on language and identity (Grant & MacLeod, 2020; Johnstone, 2009). This could be an intriguing direction for research in this field.

Whilst we continue to develop this approach to analysing power in online fora, we believe it can already provide valuable insights for the forensic linguistic analysis of online criminal interactions. Understanding a person in terms of their power resource provides a new way of describing different types of users in online fora, which may be useful for investigation and for the better understanding of specific roles and hierarchies of power in online fora where there is an absence of explicit institutional and social power structures.

References

- Ainsworth-Vaughn, N. (1998). *Claiming power in doctor-patient talk*. Oxford University Press.
- BAAL. (2021). *The British Association for Applied Linguistics: Recommendations on Good Practice in Applied Linguistics*. Retrieved 2024-02-16, from <https://www.baal.org.uk/wp-content/uploads/2021/03/BAAL-Good-Practice-Guidelines-2021.pdf>
- Blum-Kulka, S. (1990). You don't touch lettuce with your fingers: parental politeness in family discourse'. *Journal of Pragmatics*, 14(2), 259–288.
- Bourdieu, P. (1991). *Language and Symbolic Power*. Harvard University Press.
- Bousfield, D., & Locher, M. A. (Eds.). (2008). *Impoliteness in Language: Studies on its Interplay with Power in Theory and Practice*. Mouton de Gruyter.
- Brown, P., & Levinson, S. C. (1978). Questions and Politeness. In E. N. Goody (Ed.), *Universals in language usage: Politeness phenomena* (pp. 56–289). Cambridge University Press.
- Brown, P., & Levinson, S. C. (1987). *Politeness. Some Universals in Language Usage*. Cambridge University Press.
- Cialdini, R. B. (2001). The science of persuasion. *Scientific American*, 284(2), 76–81.
- Conley, J., O'Barr, W. M., & Lind, E. A. (1978). The Power of Language: Presentational Style in the courtroom. *Duke Law Journal*, 6, 1375–1399.
- Culpeper, J. (2008). Chapter 2: Reflections on impoliteness, relational work and power. In D. Bousfield & M. A. Locher (Eds.), *Impoliteness in Language: Studies on its Interplay with Power in Theory and Practice* (pp. 17–44). Mouton de Gruyter.
- Fairclough, N. (1984). *Language and Power*. Sage Publications.
- Grant, T., & MacLeod, N. (2020). *Language and Online Identities: The Undercover Policing of Internet Sexual Crime*. Cambridge University Press.
- Haworth, K. (2010). Police interviews in the judicial process: Police interviews as evidence. In M. Coulthard, A. Johnson, & R. Sousa-Silva (Eds.), *The Routledge Handbook of Forensic Linguistics*. Routledge.
- Johnstone, B. (2009). Stance, style and the linguistic individual. In A. Jaffe (Ed.), *Sociolinguistic Perspectives on Stance*. Oxford University Press.
- Kacwicz, E., Pennebaker, J. W., Davis, M., Jeon, M., & Graesser, A. C. (2014, March). Pronoun Use Reflects Standings in Social Hierarchies. *Journal of Language and Social Psychology*, 33(2), 125–143. Retrieved 2024-02-15, from <http://journals.sagepub.com/doi/10.1177/0261927X13502654> doi: 10.1177/0261927X13502654
- Lakoff, R. T. (1973). Language and Woman's Place. *Language in Society*, 2(1), 45–80.
- Leech, G. (1983). *Principles of Pragmatics*. Longman.
- Locher, M. A. (2004). *Power and Politeness in Action: Disagreements in Oral Communication*. Mouton de Gruyter.
- Locher, M. A., & Graham, S. L. (Eds.). (2010). *Interpersonal Pragmatics*. Berlin and New York: De Gruyter Mouton.
- Sillence, E. (2013). Giving and receiving peer advice in an online breast cancer support group. *Cyberpsychology, Behavior, and Social Networking*, 16(4), 480–485.
- Watts, R. J. (2003). *Politeness*. Cambridge University Press.

Jack Unterweger An authorship analysis of the notorious killer's autobiography “Fegefeuer oder die Reise ins Zuchthaus”

Karoline Marko ¹ and Alesia Locker ²

¹University of Graz, karoline.marko@uni-graz.at

²Royal Danish Defence College, ales@fak.dk

Abstract

Jack Unterweger, a notorious killer and celebrated author, was a highly skilled manipulator and was found more than once to have plagiarized poems and short stories (Herwig, 2022; Leake, 2010). Praised for his work as an author, and particularly for his autobiography “Fegefeuer oder die Reise ins Zuchthaus”, Unterweger was released from prison in 1990. Almost 40 years after the publication of his autobiography, rumors have begun circulating that Unterweger might not have written his autobiography by himself and that Sonja von Eisenstein, one of Unterweger’s benefactors and supporters, might have been involved in the writing of it (Herwig, 2022). Thus, based on a corpus of nine books by Unterweger and von Eisenstein, this study sets out to investigate whether these rumors are potentially true. The analysis is carried out with the help of a combination of quantitative (HCA, PCA) and qualitative (stylistic) methods. The analysis shows that Unterweger’s writing style is similar to that of his supposed autobiography, but that another individual’s writing style is also present.

Keywords: Principal component analysis, N-grams, Stylistics, Function words, Questioned authorship.

Resumo

Jack Unterweger, famoso assassino e célebre autor, era um manipulador muito habilidoso, tendo sido descoberto, mais do que uma vez, que plagiou vários poemas e contos. Elogiado pelo seu trabalho como autor e, em particular, pela sua autobiografia “Fegefeuer oder die Reise ins Zuchthaus”, Unterweger foi libertado da prisão em 1990. Quase 40 anos após a publicação da sua autobiografia, começaram a circular rumores de que Unterweger poderia não ter escrito a sua autobiografia sozinho e que Sonja von Eisenstein, uma das benfeitoras e apoiantes de Unterweger, poderia ter estado envolvida na sua redação.

*Assim, com base num corpus de nove livros de Unterweger e von Eisenstein, este estudo pretende investigar se estes rumores são potencialmente verdadeiros. A análise é efetuada com a ajuda de uma combinação de métodos quantitativos (HCA, PCA) e qualitativos (estilísticos). A análise mostra que o estilo de escrita de Unterweger é semelhante ao da sua suposta autobiografia, mas que o estilo de escrita de outro indivíduo também está presente. **Palavras-chave:** Análise de Componentes Principais, N-gramas, Estilística, Palavras funcionais, Autoria questionada.*

1. Introduction

Jack Unterweger was a notorious killer, malignant narcissist, psychopath (Haller in Leake, 2010; Widler, 2019), and celebrated author of a variety of books, plays, poems, and children's stories (Holzer & Reibenwein, 2022). Growing up in Wimitztal, he was a speaker of a Carinthian dialect, but he also gained rudimentary skills in English. In the late 1970s, he received a life sentence for the murder of an 18-year-old woman. During his 15 years in prison, Unterweger began to take writing classes and caught up on education he had missed¹. Unterweger also initiated a biannual journal called "Wortbrücke" (Wortbrücke, n.d.), which he edited and published while in prison (Leake, 2010). He reportedly had "the charm of a psychopath" (Haller in Leake, 2010, p. 401) and was a highly skillful manipulator. More than once, he was caught selling other people's work as his own, such as a poem by Herman Hesse and children's stories by Sonja von Eisenstein (Leake, 2010). Sonja von Eisenstein, who was an author herself and who initially fostered and encouraged Unterweger's education, was one of several benefactors he managed to attract the attention of during his time in prison (Herwig, 2022).

In the late 1980s, a petition was established to release Unterweger from prison as an example of successful rehabilitation (Holzer & Reibenwein, 2022). The petition fell on sympathetic ears and Unterweger was released in 1990 (Busch, 2019; Widler, 2019). Despite not having received any therapy during his time in prison, he was judged to have been completely rehabilitated (Leake, 2010). Upon his release, Unterweger began to work as a freelance journalist, gave readings of his books throughout Austria, and was a welcome guest on TV and various radio channels (Herwig, 2022; Leake, 2010). However, a series of murders of prostitutes in various parts of Austria that began shortly after his release started to attract the attention of some investigators. Because the *modus operandi*² in these murders was similar to the way Unterweger had killed his first victim, he soon became the prime suspect. In line with his psychopathic and narcissistic character traits, Unterweger, now working as a freelance journalist, began to report on the murders and even held interviews with the leading police officers to receive inside information on the status of the investigation (Leake, 2010). Upon finding out that he

¹It is reported that he graduated from 'Hauptschule' during his time in prison, a level of education usually obtained by 13-14 year-olds in Austria.

²The *modus operandi* he was known for is the strangling of his victims with pieces of their own clothing using a very specific and rare type of knot, and he used the same knot on himself when he committed suicide (Leake, 2010).

had been identified as the prime suspect in the killings, Unterweger fled to the United States but was soon extradited to Austria. In Graz, he was tried for the murder of eleven women and was convicted of the murder of nine of them³ in and around Austria as well as in the Los Angeles area. The conviction, however, was never spent, as Unterweger committed suicide in his cell after the announcement of the verdict (Busch, 2019; Leake, 2010).

Unterweger’s path to freedom was largely paved by the success of his 1983 autobiography called “Fegefeuer oder die Reise ins Zuchthaus”, which was also made into a movie in 1989 (Leake, 2010). Almost 40 years after the publication of the book, doubts about its authorship have arisen, and suspicions have emerged that Sonja von Eisenstein might have been involved in the writing of it (Herwig, 2022). As these doubts have so far largely been based on intuition and gut feeling, the present paper sets out to explore this issue from a more scientific standpoint and determine whether Unterweger’s autobiography is really his own.

1.1. Authorship Analysis

In order to investigate the proposed research question, this paper takes an authorship comparison approach. Forensic authorship comparison aims to analyze the writing styles of two or more authors with the objective of determining whether the authors of the investigated texts are likely to be the same or different (Ehrhardt, 2018). This can either be done largely qualitatively (e.g., Fobbe, 2020, 2022; Larner, 2014; Turell, 2010; Turell & Gavalda, 2013; Wright, 2013), largely quantitatively (e.g., Argamon, 2008, 2018; Argamon & Levitan, 2005; Belvisi, Muhammad, & Alonso-Fernandez, 2020; Johnson & Wright, 2014; Juola, 2021; Koppel, Schler, & Argamon, 2009, 2011; Stamatatos, 2009, 2013; Swain, Mishra, & Sindhu, 2017; Wright, 2017), or through a combination of both these approaches (Locker, 2019; Nini & Grant, 2013; Wright, 2013).

Qualitative approaches to authorship analysis focus on stylistic aspects of individuals’ writing styles, such as features relating to text format, the use of numbers and symbols, abbreviations, punctuation, capitalization (or lack thereof), spellings and misspellings, word formation, syntax, discourse, and mistakes and errors (McMenamin, 2002). Approaches based on a stylistic tradition, however, suffer from one disadvantage, namely that they are often considered to be too subjective (Solan, 2013). Quantitative approaches, on the other hand, are generally considered to be more objective compared to qualitative approaches, yet they frequently lack explicability (Locker, 2019). Such approaches consider the frequencies with which features appear rather than their actual use in context. It needs to be mentioned, however, that neither approach can definitively be said to be superior to the other (Ainsworth & Juola, 2018; Solan, 2013). In fact, the advantages of one method might be considered disadvantages of the other – thus, they are rather complementary, and the choice of one or the other as well as their application depends almost entirely on the situational and circumstantial elements of the respective case. That is to say, even though some automated approaches reportedly have low error rates (Koppel et al., 2009), their success rate largely depends upon the length of the available textual material (Eder, 2013, 2017). In contrast, qualitative approaches with

³Two of the bodies were too decayed, so not enough evidence was found to convict Unterweger (Leake, 2010)

detailed manual analyses are rather impractical to apply to long texts. Further, quantitative methods can miss qualitative differences and/or similarities, while qualitative methods can miss larger patterns or distributions of linguistic features (Solan, 2013). Therefore, we have chosen to apply a combination of both quantitative and qualitative methods to our dataset (see also Locker, 2019).

2. Data Methodology

The dataset for this analysis consists of five books by Jack Unterweger in addition to his supposed autobiography, and of three books by Sonja von Eisenstein. Prior to the analysis, these books were digitalized and subsequently proofread to correct OCR (Optical Character Recognition) errors, as these could add random noise to the dataset.

In total, the corpus comprises 103,606 words in the Unterweger corpus, 96,862 words in the von Eisenstein corpus, and 66,724 words of the Q-text, which all together amount to 267,192 words. Table 1 provides an overview of the dataset.

Table 1. Overview of the whole dataset.

Author	Date of Publication	Title	Short title for in-text references	Wordcount	Genre
Q	1983	Fegefeuer oder die Reise ins Zuchthaus	Fegefeuer	66,724	Autobiography
Unterweger	1982	Tobendes Ich	Tobendes Ich	6,023	Poems
Unterweger	1987	Reflexionen	Reflexionen	855	Poems
Unterweger	1988	Wenn Kinder Liebe leben	Kinder	28,303	Short stories for children
Unterweger	1990	Kerker	Kerker	49,752	Novel
Unterweger	1990	Mare adriatico	Mare	18,673	Novel
				103,606	
von Eisenstein	1992	Verschlafete eure Träume nicht	Verschlafete	51,605	Advice book
von Eisenstein	1993	Wenn die Seele Märchen erzählt	Seele	27,712	Fairy tales for children
von Eisenstein	2014	Ein Dickkopf mit zwei Flügeln	Dickkopf	17,545	Children's story
				96,862	

To establish the authorship of the supposed autobiography, a two-stage analysis was undertaken. Stage 1 involved the statistical identification of strong co-occurring lexico-grammatical patterns in each of the texts. Due to the length of the texts, we split the individual files into chunks of 5,000 words in order to test whether our model could identify linguistic features that differentiate the writing styles of the texts known to be written by the two authors before using the model on the text whose authorship is questionable. In our analysis, we focused on function words, which are described in more detail below. After presenting the statistical identification of patterns, we provide

a graphical visualization of the relationship between the texts based on the similarities identified. In Stage 2 we then conducted a qualitative inspection of the patterns identified in Stage 1 in order to interpret the functional and idiosyncratic bases underlying the identified patterns. To investigate the use of the identified features in context, we used the corpus program AntConc (Anthony, 2022) to generate concordance lines⁴, MaxQDA for the qualitative annotation of the texts, and #LancsBox (Brezina, Weill-Tessier, & McEnery, 2002) for the comparison of n-grams (lockgrams). This combination of quantitative and qualitative methods allows us to describe the author’s writing styles in great depth and to support our findings with linguistic theories.

2.1. N-Grams Function Words

N-grams are strings of either words or characters of varying length (Wright, 2017). For example, word 4-grams are strings of four consecutive words, and character 4-grams are strings of four consecutive characters. Even though it is still unclear why character n-grams appear to be successful in some authorship analyses (Kestemont, 2014), the successful use of word n-grams in authorship analyses can potentially be explained. It has been argued, for instance, that the linguistic input that individuals are exposed to differs from other individuals because of their social backgrounds, including their families, network of friends, the acquisition of first and additional languages, education, profession, hobbies, etc. (see, e.g., Ehrhardt, 2018; Nini, 2018b). Thus, due to their sociolinguistic histories (Grant & MacLeod, 2018), individuals will subconsciously be primed for the combination of particular words (i.e., collocations) if they occur with high frequencies in their linguistic surroundings (Larner, 2014; Wright, 2017).

It has been suggested that the “formulaic sequences” (Wray, 2002, p. 9) that individuals are exposed to are stored as “holistic units” (Larner, 2014, p. 2) in the minds of the speakers and can thus be useful indicators of an individual’s style. This results in what Coulthard (2004, 2007) has termed “preferred co-selections”: the potentially individuating preference for the co-selection of particular words, the combination of which can differentiate one person’s style from that of others. Evidence for this theory comes from different linguistic disciplines, such as psycholinguistics (Wray, 2002), sociolinguistics (Coulmas, 1979), corpus linguistics (Moon, 1998), and language acquisition (Peters, 1984). Such formulaic sequences become a part of an individual’s vocabulary, and, with that, a part of their idiolect (Coulthard, 2004). Schmitt, Grandage, and Adolphs (2004), for instance, say that

every person has their own unique idiolect made up of their personal repertoire of language, and as part of that idiolect, it seems reasonable to assume that they will also have their own unique store of formulaic sequences based on their own experience and language exposure. (p. 138)

In the present paper, we have chosen to focus on the use of word n-grams, because their use is explicable and can complement the qualitative analysis of texts. As Wright (2017) has pointed out, however, it is possible that different lengths of n-grams will have different success rates for different authors, and that for some authors, n-grams generally work better than for others.

⁴A concordance is (...) a list of all of the occurrences of a particular search term in a corpus, presented within the context that they occur in” (Baker, 2006, p. 71).

Function words such as determiners, conjunctions, and prepositions have been found to be less topic-dependent than content words (i.e., nouns, verbs, adjectives, and adverbs) and can technically be considered unigrams. An individual's use of function words is more likely to be indicative of their writing style rather than the content of the texts they produce (e.g., Argamon & Levitan, 2005; Miranda García & Calle Martín, 2007; Pennebaker, 2011). Importantly, the use of function words is less likely to be consciously controllable by an individual, while content words seems to be reflected upon more by the author and chosen more deliberately (Argamon & Levitan, 2005). Further, function words particularly lend themselves to authorship analysis because they appear with high frequencies in language, which also makes them easily quantifiable (Kestemont, 2014). The inclusion of pronouns in the analysis, in contrast, is less widely agreed upon, as evidence has been found that some of them are highly influenced by topic and genre (Biber, Conrad, & Reppen, 2006; Burrows, 1987a; Herring & Paolillo, 2006; Hoover, 2004), which might make them less reliable as markers of authorship, even though evidence of linkage between the author's personality and pronoun choices also exists (Newman, Groom, Handelman, & Pennebaker, 2008; Pennebaker, 2011; Rude, Gortner, & Pennebaker, 2004).

The use of the relative occurrence of function words as input for statistical analysis is among the oldest and best-known computational authorship analysis methods. Studies that have used this technique to investigate authorship date back to 1984 (Mosteller & Wallace, 1984). An advantage of using function words is that they have very little semantic content, and therefore they are less dependent on the topic of the text. On the other hand, function words express grammatical relationships among the words within a sentence and can consequently give us some insight into the speaker's or writer's grammatical preferences. The use of function words is also grounded in functional linguistics (Biber & Conrad, 2009), which gives our analysis interpretability. The relative stability of the frequency of function words has also been demonstrated in other studies. An interesting study in this respect and one which resembles our case to some extent was undertaken by Binongo (2003). Binongo compared the styles of Frank Baum, the creator of the Land of Oz, to the style of his successor, Ruth Thompson, who continued the book series after Baum's death. Interestingly, although Thompson's early Oz novels are often considered faithful to Baum's canon, the Principal Component Analysis revealed that the styles of the two authors were very different in their use of function words.

2.2. Principal Component Analysis (PCA) Hierarchical Cluster Analysis (HCA)

Principal Component Analysis (PCA) is a statistical method that helps an analyst to "clarify a complex mass of data" (Grant & Baker, 2001, p. 77). The method was first adapted and introduced to authorship attribution tasks by Burrows (1987b, 1992). Through the extraction of correlations between linguistic features (Gries, 2022), PCA reduces the list of features for the comparison of potential authorship markers to only those markers that differ in the writing styles of the authors investigated (Grant & Baker, 2001). In this paper, PCA thus facilitates the identification of features in the writing styles of the two authors that separate them from each other before the document whose authorship is questionable is introduced into the analysis. Once the features that differentiate Unterweger's writing style from von Eisenstein's style have been identified, the document under investigation is added in order to determine whether the

writing style in Unterweger’s supposed autobiography is more similar to Unterweger’s style or to von Eisenstein’s style. PCA has previously been used by other scholars such as Holmes (1992), and Baayen, van Halteren, and Tweedie (1996) for authorship attribution as well.

A Hierarchical Cluster Analysis (HCA) is frequently used in identifying similarities and dissimilarities in large datasets. As Gries and Stefanowitsch (2010) point out:

this is done by (i) comparing pairwise (dis)similarities between the items in a (dis)similarity matrix, (ii) successively amalgamating all items into clusters such that the items within a cluster are highly dissimilar to each other and at the same time highly dissimilar from all other items and clusters, and (iii) representing the resulting structure in the form of a tree-like dendrogram (p. 77).

Both methods are utilized in the present analysis in order to obtain an overview of potential linguistic patterns in the text, which can be used as a starting point for further, more in-depth investigations.

2.3. Limitations

The main limitation of our dataset is that even though we have gathered and collected all the available books by the two authors under investigation, some books are out of print. Thus, the books that we include in the analysis are of different genres, which can pose a problem for the analysis (see, e.g., Overdorf & Greenstadt, 2016). Interestingly, however, even though the document examined in our study is officially an autobiography, it contains several elements of fiction, as it has been found that not all the reported events happened in the way they are portrayed in the narrative (e.g., Herwig, 2022). This makes the text under investigation similar to Unterweger’s known novels, which in turn also contain some biographical elements and references⁵. Also, the oldest book in the corpus dates from 1982, while the newest one dates from 2014, which represents a 32-year difference. This is particularly an issue for Sonja von Eisenstein, whose books date from 1992, 1993, and 2014 – the time span between the oldest and the newest book can potentially explain some intra-author differences. For Unterweger, the dating of the books is less of an issue, as they were all published within an eight-year period and during and shortly after his time in prison, which means that the situational aspects remain relatively constant. Lastly, even if the results of the analysis were to show some similarities to Sonja von Eisenstein’s writing style, the presence of a third possible author cannot entirely be ruled out (see also Grant & Baker, 2001), and it is also possible that other people, such as Unterweger’s teacher(s), had a strong influence in this writing. Similarly, it is possible, as with any book, that the texts have been proofread and edited by the publishers, which can also interfere with the author’s writing style (Fobbe, 2011). However, this is perceived to be less of an issue in the present case, as we believe that editing and proofreading was kept to a minimum with Unterweger’s texts⁶. This argument is supported by the presence of multiple spelling mistakes and grammatical errors in his books, as will be discussed in more detail in the qualitative analysis.

⁵In one interview, Unterweger even refers to them as sequels to the autobiography (Mose, 2015).

⁶In a podcast by Herwig (2022), one of Unterweger’s editors states that Unterweger’s style was retained in the books and that only minor changes were introduced by the editor.

3. Analysis

The analysis section is separated into three parts. First, we will provide some general information about the dataset before we proceed to an overview of the quantitative examination of the investigated material. Based on these findings, we have identified words and phrases of particular interest, which are examined in a follow-up qualitative analysis. The qualitative analysis also focuses on potentially idiolectal markers, as well as on mistakes and errors.

3.1. Stage 1: Pre-evaluation of the texts

First, we consider sentence and word lengths as well as sentence complexities measured using the Flesch Readability Ease score (Bachmann, 2014), which provides some indication of the complexity of the texts. Table 2 shows the results of this analysis:

Author	Short title	Readability Ease (RE)	Average word length	Average sentence length
Questioned	Fegefeuer	70	4.9	13.1
Unterweger	Tobendes Ich	n/a*	6.3	n/a*
Unterweger	Reflexionen	n/a*	5.5	n/a*
Unterweger	Kinder	69	4.8	15.1
Unterweger	Kerker	61	5.1	12.1
Unterweger	Mare	60	5.2	15.9
von Eisenstein	Verschlafet	46	5.4	18.4
von Eisenstein	Seele	71	4.8	13.9
von Eisenstein	Dickkopf	63	5.1	13.4

Table 2. Overview of Readability Ease (RE), average word lengths, and average sentence lengths. Low RE scores indicate higher complexities; higher scores indicate less complexity.

*For the collection of poems, no RE scores and average sentence lengths are provided, as the poems contain many sentences without punctuation. Imposing punctuation marks would skew the results to a similar degree to calculating the values and scores with the texts as they are.

As shown in Table 2, the RE score of almost all the books in the corpus is between 60 and 71, which indicates that the sentence structures in all of the texts are rather simple. The only exception is von Eisenstein's "Verschlafet eure Träume nicht", which, with a score of 46, is more complex than the others. No scores could be obtained for Unterweger's books of poems ("Tobendes Ich" and "Reflexionen"), as the lack of punctuation and incomplete sentences would have skewed the analysis. In terms of word length, both Unterweger's and von Eisenstein's fairy tales contain shorter words; von Eisenstein's children's stories also show decreased sentence lengths compared to her advice books, whereas Unterweger's fairy tales contain the third-longest sentences of his corpus, with sentences similar in length to "Mare", and it is the novel "Kerker" that contains the shortest sentences. After this initial evaluation of the texts, a quantitative examination was conducted.

3.2. Stage 2 – Quantitative examination

Our statistical analysis investigates frequency patterns of function words and their co-occurrence in the books of Jack Unterweger and Sonja von Eisenstein. Subsequently, the results of the PCA and HCA will be presented.

3.2.1. Pre-tests with Known Texts

The statistical AA was carried out in R (R Core Team, 2022), a programming language for statistical analysis, utilizing the Stylo package for computational stylistics (Eder, Rybicki, Kestemont, & Pielstrom, 2022). To identify the frequency patterns which distinguish the styles of Jack Unterweger and Sonja von Eisenstein, a list of the 200 most frequently occurring words in all of their known texts was created. In computational AA, the length of the feature list is not agreed upon, with research suggesting numbers from the 50 most frequent words to all the words in the corpus (see Stamatatos, 2009, pp. 540–541). The number in our study was chosen to include the majority of frequently occurring function words. It is noteworthy that AA based on frequency patterns is susceptible to the modelling of the topic, since many of the lexical features are highly topic dependent. The wordlist was reviewed to exclude content words and pronouns, as it has been demonstrated that the topics favored by authors can significantly influence the authorship attribution accuracy and result in artificially high scores (Layton, Watters, & Dazeley, 2015). The words which occur in fewer than 50% of the texts were excluded from the analysis, since sparse data would lead to poor statistics. The removal of the content words and rare words shortened our feature list to 114 words.

The statistical AA was carried out in R (R Core Team, 2022), a programming language for statistical analysis, utilizing the Stylo package for computational stylistics (Eder et al., 2022). To identify the frequency patterns which distinguish the styles of Jack Unterweger and Sonja von Eisenstein, a list of the 200 most frequently occurring words in all of their known texts was created. In computational AA, the length of the feature list is not agreed upon, with research suggesting numbers from the 50 most frequent words to all the words in the corpus (see Stamatatos, 2009, pp. 540–541). The number in our study was chosen to include the majority of frequently occurring function words. It is noteworthy that AA based on frequency patterns is susceptible to the modelling of the topic, since many of the lexical features are highly topic dependent. The wordlist was reviewed to exclude content words and pronouns, as it has been demonstrated that the topics favored by authors can significantly influence the authorship attribution accuracy and result in artificially high scores (Layton et al., 2015). The words which occur in fewer than 50% of the texts were excluded from the analysis, since sparse data would lead to poor statistics. The removal of the content words and rare words shortened our feature list to 114 words.

The length of the books allowed them to be split into blocks of 5000 words. This provided us with a substantial number of samples of undisputed authorship to serve as a baseline against which we could assess the success of the AA while still leaving enough data for the computational analysis to identify patterns in the stylistic preferences of Unterweger and von Eisenstein.

In order to further investigate the extent of the influence of genre, an analysis of the whole corpus was undertaken. The dendrogram in Figure 1 shows that Unterweger’s children’s stories are similar in the use of function words to von Eisenstein’s texts. As

mentioned before, there are rumors that some of these stories might have been stolen from von Eisenstein; however, to avoid exaggerating the implications of our findings, it has to be emphasized that the genre of children’s stories could play a part in this categorization as it could, to a certain extent, affect the distribution of function words. Due to the potentially disputable authorship of “Wenn Kinder Liebe leben”, we exclude this book from further analysis.

The rest of the books by Unterweger can be successfully clustered together despite differences in genre and topic. Even though “Verschlafet eure Träume nicht” by von Eisenstein is a self-help book aimed at the adult reader, it is revealed by the analysis to be more similar to the children’s books by the same author than the books for adults by Unterweger.

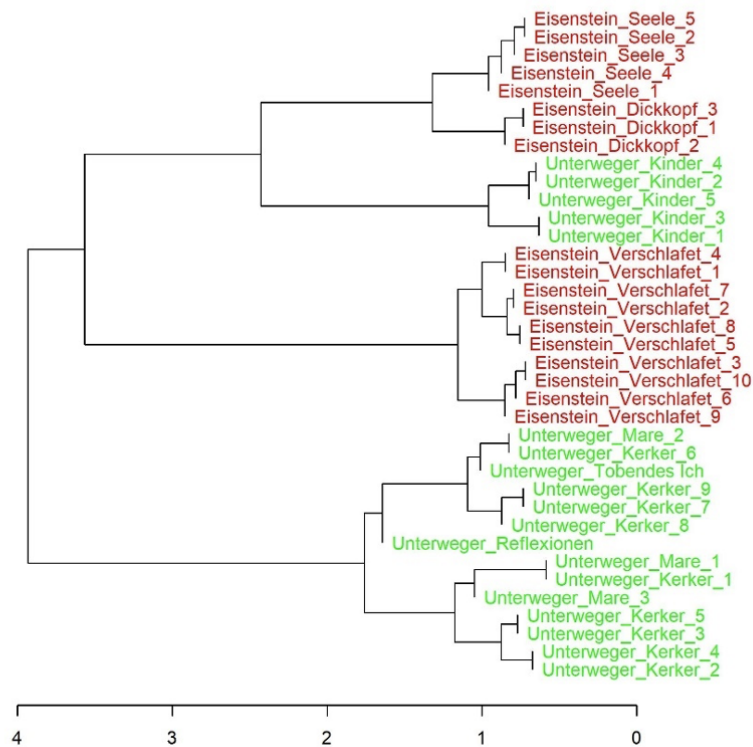


Figure 1. Cluster analysis of the known texts by Unterweger and von Eisenstein.

3.2.2. Adding the Q-Text

After having established that the list of 114 functional features distinguished Unterweger’s writing style from von Eisenstein’s style, the autobiography was added to the analysis. While the style of the books known to be written by Unterweger is very consistent, as indicated by the blending of the text blocks from several books (see Figure 2), it is remarkable that the autobiography forms a separate tight cluster of its own. This means that the frequency profiles of this text’s blocks are consistently distinct from the books known to be by Unterweger, although they are still more similar to Unterweger’s style than to that of von Eisenstein.

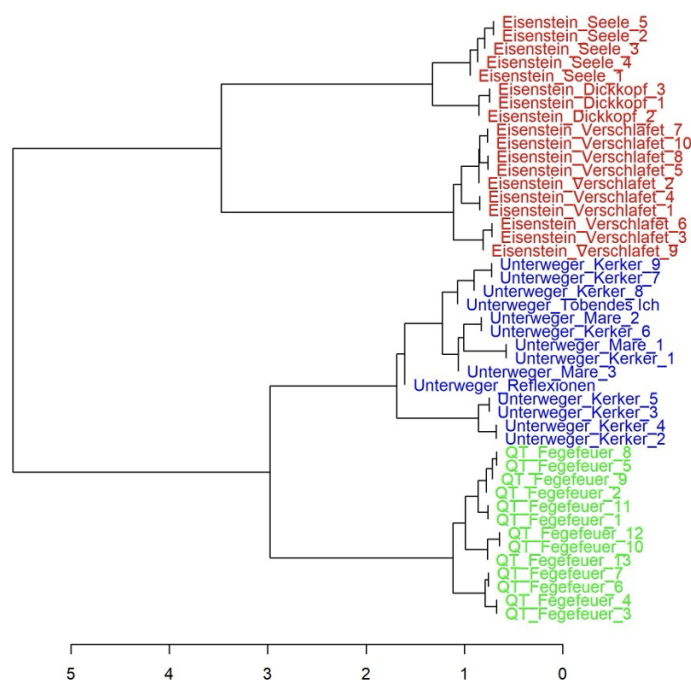


Figure 2. Cluster analysis: the questioned autobiography and the known texts by Unterweger and von Eisenstein.

Since we cannot exclude the possibility that Unterweger’s later books were influenced by authors other than himself, seven of his longer personal letters used in a different study (Marko & Leibetseder, 2023) were added to the analysis for the following reason: these private letters were addressed to the same person, and it is known that they were authored by Unterweger himself. Thus, this analysis aimed at finding out whether the books known to be by Unterweger or his supposed autobiography is more similar to the letters in style, or whether these letters would form a separate cluster like the text under investigation. The letters were not included in the initial analysis, because they are of a very different genre. However, since the cluster analysis did not identify either Unterweger or von Eisenstein as the author of the text under examination, and because we wanted to investigate whether this text or Unterweger’s known books were closer to his “true” writing style, we added the personal letters. Figure 3 illustrates that Unterweger’s letters cluster with his known texts. Despite being prose, the distribution of function words in his letters is more similar to Unterweger’s poem collection “Reflexionen” than to his supposed autobiography. This provides support for the hypothesis that the style of the autobiography is different from the style of his other books and personal letters.

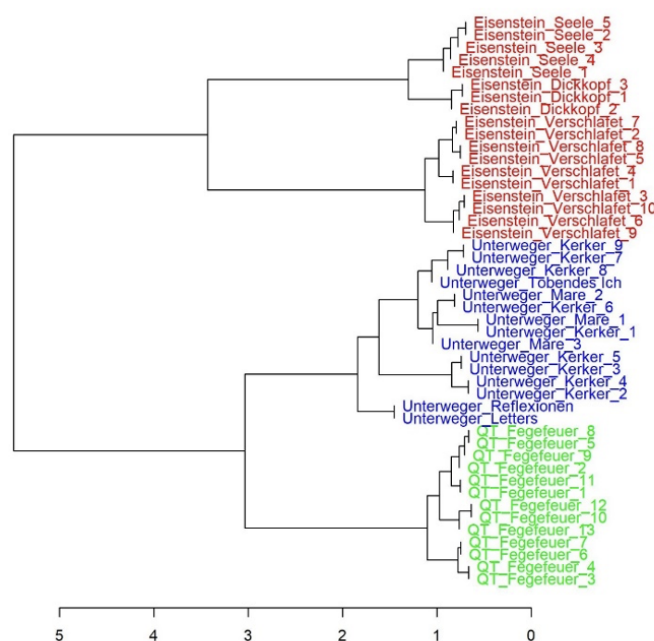


Figure 3. Dendrogram with personal letters.

To further understand which of the function words are most informative in differentiating between the styles of the authors, we refer to the exploratory statistical technique of PCA. This analysis boils down the frequencies of the 114 function words from our feature list to a single point on the plot that represents a specific text (see Figure 4a). The texts with similar frequency profiles are grouped together, while the loadings (see Figure 4b) help us more deeply examine how strongly specific function words contribute to this process and which function words frequently occur together in these groups. It is clearly evident that the analysis separates the authors quite well. The words with higher values in Figure 4b, positive or negative, have a larger effect on the results of the separation. Generally, the words on the right of this plot are the ones that are more indicative of Unterweger's writing style and the words on the left are characteristic of von Eisenstein's writing style.

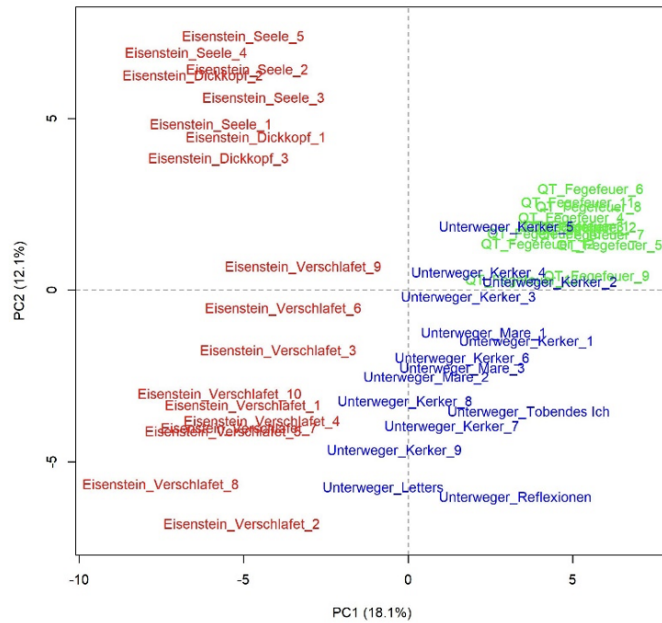


Figure 4a. PCA of the corpus. Plot with principal component scores of the text blocks.

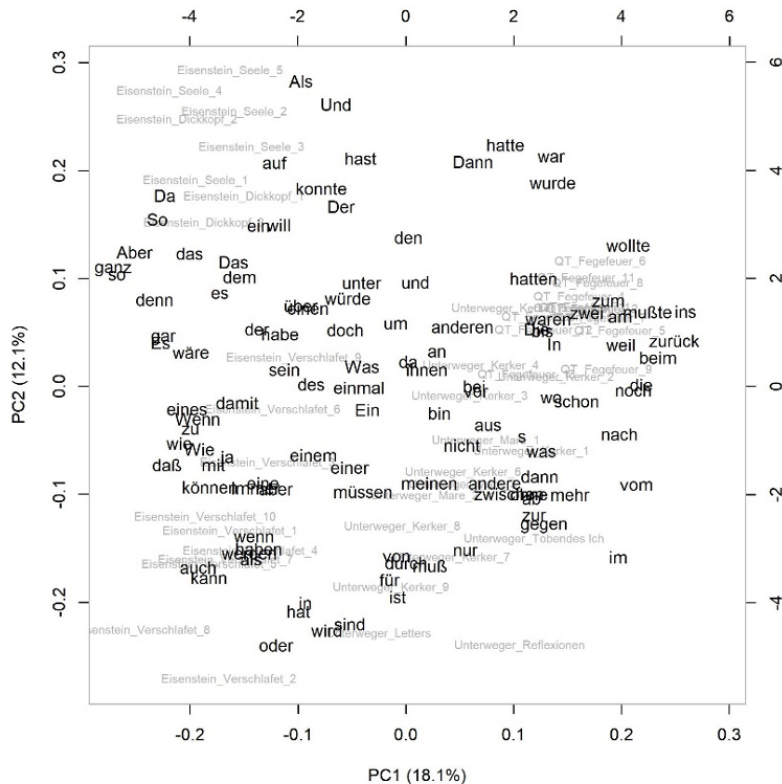


Figure 4b. PCA biplot. The loadings (top and right axes) show how strongly each feature influences the principal components and how they correlate with one another.

According to the PCA, similarities between the text under investigation and Unterweger’s other books can be found in the frequent occurrence of prepositions and conjunctions: “ins” (“into”), “beim” (“with”), “am” (“on”), “weil” (“because”), and “zum”

“to”) all occur more frequently compared to the books by von Eisenstein. If we inspect the data, we notice that the relative frequencies of these features are higher in the supposed autobiography than in the texts known to be by Unterweger (see Table 3). Interestingly, a similarity between the autobiography and von Eisenstein’s style of children’s stories can be found in the frequent appearance of the conjunction “Und” (“And”) and adverb “Dann” (“Then”) at the beginning of the sentence, which does not seem to be the case for the texts known to be by Unterweger.

	Und	und	Dann	dann	weil	am	ins	zurück	beim	muß	mußte	wurde	wollte
Eisenstein_Dickkopf	0.60	2.99	0.16	0.18	0.04	0.11	0.05	0.05	0.03	0.00	0.00	0.16	0.11
Eisenstein_Seele	0.79	3.57	0.14	0.21	0.14	0.18	0.06	0.09	0.03	0.04	0.04	0.16	0.10
Eisenstein_Verschlafet	0.10	2.67	0.07	0.34	0.10	0.14	0.06	0.07	0.03	0.08	0.07	0.11	0.11
QT_Fegefeuer	0.37	2.75	0.18	0.37	0.25	0.31	0.30	0.20	0.13	0.07	0.30	0.26	0.30
Unterweger_Kerker	0.15	3.58	0.04	0.26	0.17	0.15	0.14	0.13	0.10	0.13	0.07	0.11	0.10
Unterweger_Mare	0.24	3.48	0.05	0.22	0.19	0.18	0.17	0.09	0.07	0.12	0.13	0.12	0.14
Unterweger_Tobendes Ich	0.15	3.20	0.08	0.42	0.13	0.30	0.13	0.17	0.10	0.08	0.03	0.12	0.05
Unterweger_Reflexionen	0.00	3.17	0.00	0.23	0.12	0.12	0.00	0.35	0.00	0.23	0.00	0.00	0.12

Table 3. Characteristic features of the autobiography and their relative frequencies.

On its own, each of these features may have little significance for AA, but their combined weight pulls the autobiography into a separate cluster in the plots. While it is difficult to make any claims about authorship at this stage, we can argue that the autobiography, for whatever reason, was not written in the same style as Unterweger’s other books. Subsequently, we will move on to the qualitative part of the analysis.

3.3. Stage 3 – Qualitative description

The qualitative analysis presented below proceeds as outlined in Fobbe (2021), and thus begins with an error analysis followed by a stylistic analysis.

3.3.1. Error Analysis

As mentioned previously, several mistakes were identified in the texts. These errors are largely orthographical, but the texts also contain some syntactic mistakes and some relating to idiomatic language use. This indicates only minimal interference of an editor. Sonja von Eisenstein’s books, in contrast, do not contain any of these errors or issues with idiomatic language use. Some examples of Unterweger’s texts are shown in Table 4.

	German – Error	German - Corrected	English translation	Source Text
1	Wörthsee	Wörthersee	Lake Wörth	Fegefeuer (QT)
2	Fibrieren	Vibriieren	Vibrate	Fegefeuer (QT)
3	Fridhof	Friedhof	Cemetery	Fegefeuer (QT)
4	Radstatt	Radstadt	<i>Name of a city</i>	Fegefeuer (QT)
5	Steh’ns auf	Stehn’S auf	Stand up	Fegefeuer (QT)
6	Artbeitskittel	Arbeitskittel	Work coat	Kerker
7	Agression	Aggression	Aggression	Kerker

Table 4. Examples of orthographical errors in Unterweger’s books.

Table 4 above shows orthographical errors, but, as mentioned, other mistakes were also identified, for example in relation to idiomatic language use. These mistakes might point towards Unterweger’s basic level of education and his relative inexperience with written language up to the point at which he decided to become an author. Examples of his unidiomatic use of words and expressions can be found in Table 5 below:

	Example	Intended	Translation (literal)	Source Text
8	Minute auf Minute wird Vergangenheit		Minute after minute becomes the past	Kerker
9	steirischer Österreicher		Styrian Austrian	Kerker
10	schönwörtlicher Neubeginn	Sprichwörtlicher Neubeginn	Proverbial fresh start	Kerker
11	Vorsichtsabstand	Sicherheitsabstand	Safe distance	Kerker & Mare
12	ich rede Worte		I talk words	Mare
13	hassende Wut		Hating anger	Kerker & Mare

Table 5. Examples of unidiomatic language use in Unterweger’s known texts.

Similar salient words and expressions can be found in Unterweger’s supposed autobiography (see Table 6), but they do not appear in von Eisenstein’s book.

	Example	Intended	Translation (literal)	Source Text
14	In dieser Nacht <u>erbrach</u> ich alle Geschäftsstände (...)	In dieser Nacht brach ich in alle Geschäftsstände (...)	In this night, I <u>vomited</u> all of the booths (...) [intended: broke into]	Fegefeuer (QT)
15	Versprechungen waren <u>zerronnen</u>		Promises melted away	Fegefeuer (QT)
16	hämmernde Vergewaltigung		Hammering rape	Fegefeuer (QT)
17	Rauflustigkeit	Rauflust	Rowdiness	Fegefeuer (QT)

Table 6. Examples of unidiomatic language use in the autobiography.

After this overview of errors and mistakes, we will subsequently examine the stylistic features detected using PCA and HCA in more detail, in order to find out if there is an explanation for their use in the text under investigation.

3.3.2. Stylistic analysis

The next step in the qualitative part of the analysis is based on the PCA and HCA described above. As mentioned, the features that were significantly more common in Unterweger’s supposed biography than in his known texts are the words “zum” (“to”), “weil” (“because”), “Und” (“And”), “ins” (“into”), “beim” (“with”), and “am” (“on”), which are prepositions and conjunctions. These features were investigated with AntConc, i.e., concordance lines were created to see how the words are used in context in order to find an explanation for why they appear so much more frequently in the autobiography

compared to Unterweger's known texts. Table 7 shows an example of the concordance lines for the preposition "zum" ("to"):

#	File		KWIC	
182	Q.txt	Ich schwieg und fuhr mit ihm	zum	Bezirksgericht der Stadt. "Wir hatten Glück, der
		I kept silent and drove him	to	the circuit court of the city. "We were lucky, the
183	Q.txt	Kaffee. "Ist er das?" sagte der Mann	zum	Amtsrat. "Stimmt, Herr Rat." Der zweite setzte
		coffee. "Is that him?" said the man	to	the councilor. "Yes, Mr. Councilor." The second sat
184	Q.txt	und Jeans, kichernd und flüsternd, trippelnd vorbei	zum	Greifen nahe und doch unerreichbar.
		and jeans, giggling and whisperingly, tripping past	within	reach and yet unattainable.
185	Q.txt	der Küche haben sie sicher noch was	zum	Essen", sagte der Cheferzieher und schob mich
		the kitchen, you sure have something	to	eat", said the teacher and pushed me
186	Q.txt	die Küche, die geben ihm noch was	zum	Essen", sagte der älteste, ruhigste, kleine,
		the kitchen, they give him something	to	eat", said the oldest, quietest, short
187	Q.txt	älteste, ruhigste, kleine schnauzbärtige Erzieher vom Dienst	zum	Blondgelockten. Ich müßte vor ihm gehen,
		oldest, quietest, short mustachioed teacher	to	the one with the blond curls. I would have to walk in front of him,
188	Q.txt	sagte mir der Erzieher, "da hast Zeit	zum	Nachdenken!" Ich dachte nach, in meinem Kopf
		said the teacher to me, "there you will have time	to	ponder!" I thought, in my head
189	Q.txt	hat... In der zweiten Woche kam ich	zum	Arzt. Nachdem ich eine gute Stunde
		had... in the second week I came	to	the doctor. After an hour
190	Q.txt	ich als gesund registriert. Anschließend kam ich	zum	Psychologen. Zusammen mit sieben anderen
		I was registered as healthy. Afterwards I came	to	the psychologist. Together with seven others
191	Q.txt	hat, brauchen wir drei neue Häuser." Und	zum	Vorführbeamten: "Die haben ihm drei Tage auf
		has, we need three new houses". And	to	the clerk: "they gave him three days

Table 7. Examples of the use of the preposition "zum" ("to"), original and translation.

After the extraction of concordance lines, the concordance lines were imported into MaxQDA, which allows for the easy coding of datasets. Using MaxQDA, we attempted

to identify patterns in the use of these words which would help us find an explanation for their frequent occurrence in the text under investigation. The patterns that were found are related to physical localizations and movements, the description of moods and states, and to temporal information. Many of the prepositions are used with nominalized constructions, which suggests that the writing style employed in the autobiography is more nominal compared to Unterweger’s known texts, at least in respect to the description of movements. The difference in the usage of these words across the autobiography and known texts is of interest insofar that all three of the texts (the autobiography and the novel-like books) focus on travelling, movement and people. Thus, the fact they occur significantly more frequently in the autobiography is a remarkable stylistic difference. Whether or not this difference is due to differences in authorship is yet unclear.

Another interesting finding was discovered in relation to the conjunction “Und” (“And”) with a capital letter, as has been pointed out above. The spelling of this word with a capital letter indicates that it is used sentence-initially. A close examination of the texts further reveals that it is used for the incorporation of direct speech (i.e., turn-taking sequences in conversations) in the autobiography, which apparently is not the case in the texts known to be by Unterweger.

A close reading of the texts has further shown that Unterweger has a very distinctive writing style characterized by words he creates on the fly, such as compounds (mostly nouns but also adjectives), and he uses nominalizations that are not part of standard German language. Examples can be seen in Table 8 below and are also provided in Table 4 above. In order to establish how common these words are in a general reference corpus, the DWDS 1900-1999 (DWDS Kernkorpus, 2023), which comprises 121,494,429 words from a variety of genres, was consulted. A corpus search revealed that Unterweger’s word creations are indeed very rare⁷ in the corpus, if they appear at all. While these words serve as illustrations for Unterweger’s creative language use, this does not mean that these exact words appear in more than one of the books. If they appeared in more than one of his writings but rarely occurred in the reference corpus, this would lend support to the idea that he is indeed the author of all his books. Thus, in order to investigate whether any of his salient uncommon words appear in the autobiography and at least one other book, a 1-gram analysis with #LancsBox was undertaken (see below).

	German word	English translation (literal)	Occurrences in the DWDS reference corpus
18	Antwortgebung	Answerprovision	0
19	Blondhaar	Blondhair	49 (i.e., 0.4 times per million words)
20	Augenkrieg	Eyewar	0
21	Freundenmorgen	Delightmorning	0
22	Einsammachung	lonely-making	0

Table 8. Examples of salient uncommon words in Unterweger’s writings.

⁷(Nini, 2018a, p. 628) refers to frequencies of 10-18 times per million words as “quite rare”.

As mentioned, nominalizations such as those in Table 8 are frequently found in Unterweger's texts. In addition to nominalizations of compounds, temporal nominalizations of adverbs such as "das Jetzt" ("the now"), "das Gestern" ("the yesterday"), and "das Irgendwann" ("the anytime"), which provide the text with the air of a literary register, are also present. In addition, nominalized constructions using verbs and adjectives, such as "ein Verjagtwerden" ("the chasing away"), "die Hinterfragung" ("the scrutinization"), and "Herbstlichkeit" ("autumn-ity") can be detected.

Before moving to the 1-gram analysis of uncommon words that appear in more than one of Unterweger's books, it has to be pointed out that the qualitative analysis shows a large lexical overlap between "Kerker" and "Mare", which, upon closer reading, can easily be explained: large parts of these two books are identical with only slight modifications, and can thus be considered as instances of self-plagiarism. Based on this finding, the lexical similarities between these two texts must not be overrated; this, however, does not affect the comparison of the autobiography and Unterweger's known texts. The following examples in Table 9 serve as an illustration:

	Mare	Kerker	Translation
23	Ich war dreiundzwanzig und gierte nach Leben, Lust. Hot-Love! In den Gedanken: Aggression , Wut, Ohnmacht und irreale Allmacht . Das Erwachen, vermischt mit wilden Gefühlen, in der Zeit der Morgendämmerung. An diesem Morgen, dieser späten Nachtstunde, kreischten meine Nerven wie ein rostiges Sägeblatt im nassen Buchenholz und verhinderten ein weiteres Bettspiel. Später ertönte die Sirene dreimal, grell, Gänsehaut erzeugend und befehlend: Tagwache!	Ich war 23 und gierte nach Leben, Lust, Hot Love! In den Gedanken: Agression [sic], Wut, Ohnmacht und Allmacht . Das Erwachen vermischt mit wilden Gefühlen in der Zeit zur Morgendämmerung. An diesem Morgen, dieser späten Nachtstunde, kreischten meine Nerven und verhinderten ein weiteres Bettspiel. Später ertönte die Sirene dreimal, grell, Gänsehaut erzeugend und befehlend: Tagwache!	I was 23 and craved for life, passion, hot love! In my mind: aggression, anger, impuissance, and omnipotence. The awakening mixed with wild feelings during dawn. On this morning, at this late time of night, my nerves screeched and prevented another bed game. Later, the siren sounded three times, gratingly, giving me the shivers, and commanding: a wake-up call!
24	Für mich war es wieder einmal ein letztes Mal . Ich wartete auf meine siebente Entlassung aus einem Gefängnis.	Für mich wieder einmal zum letzten Mal . Für diesmal . Ich erwartete meine sechste Entlassung , (...)	For me, it was another last time. For this time. I was awaiting my sixth release, (...)

Table 9. Examples of self-plagiarism in "Kerker" and "Mare adriatico"

As mentioned, we will now consider which salient uncommon words appear in the autobiography and in at least one of the books known to be written by Unterweger.

First of all, the use of the adjective “berockt” (“skirted”, as in “dressed in a skirt”) in the autobiography and both of Unterweger’s other books is of interest. A search in the DWDS corpus reveals only two hits (0.016 times per million words) of this word, which indicates that it is very rare indeed. Similarly, the adjective “tränenblind” (“tearblind”), which only has three hits (0.02 times per millions words) in the DWDS corpus, appears twice in the autobiography and once in “Kerker”; “Vergangenheitsbilder” (“pastimages”) only has one hit (0.008 times per million words), “Erinnerungsbilder” (“memoryimages”) has no hits in the DWDS, and “schmutziggrau” (“dirtygray”) has 20 hits (i.e., 0.16 per million words), but all of them appear in both the autobiography and in at least one of Unterweger’s known texts. The same is true for the expressions “jovialer Ton” (“jovial tone”), which has only one hit (0.008 times per million words) in the DWDS, and “Au-Schrei” (“ouch-cry”) and “Au-Rufe” (“ouch-call”), which do not appear in the DWDS at all, appear in the autobiography and at least one of Unterweger’s known texts. Importantly, none of these appear in von Eisenstein’s book. The appearance of all these lexical items in the autobiography and at least one of Unterweger’s known texts provides evidence for Unterweger being the author of all of them.

As with the use of nouns and compounds, Unterweger is similarly creative in his use and creation of adjectives, many of which end in -los (-less), and most of which are also not part of the standard vocabulary of German. Examples thereof can be seen in Table 10 below:

	German word	English translation (literal)	Occurrences in the DWDS reference corpus
25	blicklos	Gazeless	68 (0.6 per million words)
26	fluchtlos	flightless	0
27	gegenwartslos	presentless	2 (0.016 per million words)
28	kollegenlos	colleagueless	0

Table 10. Examples of adjectives ending in -los (-less) in Unterweger’s books.

Additionally, Unterweger’s use of color terms is a salient feature of his writing style. He appears to have a preference for the combination of two colors, or a color plus another adjective, but, as he does with compound nouns, he combines these words largely without hyphens. To conclude this part of the stylistic analysis, it is remarkable that Unterweger employs a very nominal writing style generally characterized by nominalizations, compounds, attributions, subordinate attributions, and extended participial attributes (Hennig, 2020), as well as a high level of creativity.

3.3.2.1. Lockgrams

The stylistic analysis in the previous section can be complemented by an examination of lockgrams in order to investigate potential overlap between n-grams of different lengths across all the books in the corpus. Using the corpus program #LancsBox, lists of n-grams were generated and compared among the books. As mentioned before, Lockgrams are n-grams that appear with similar frequencies in the investigated texts (Brezina, 2018, p. 79-87), which means that they might indicate common authorship. The resulting list of lockgrams for each of the books was then checked in the DWDS to

determine how frequent they are in the reference corpus. An overview of the number of lockgrams in the books is provided in Table 11.

First, the autobiography and Sonja von Eisenstein's book were compared. Although 800 words that appear in the autobiography also appear in von Eisenstein's book, none of these are rare, salient, or non-standard (unlike Unterweger's non-standard words, as discussed previously). However, 2,221 2-grams from the Q text, 869 3-grams, 74 4-grams, nine 5-grams, and even one 6-gram appear in the autobiography and von Eisenstein's book. An investigation of these word sequences in the DWDS reveals only two hits for two of the 5-grams (i.e., 0.016 per million words).

	1-grams	2-grams	3-grams	4-grams	5-grams	6-grams	7-grams
Fegefeuer + Verschlafet (von Eisenstein)	800	2,221	869	74	9	1	-
Fegefeuer + Kerker (Unterweger)	1,314	4,399	1,447	237	46	7	1
Fegefeuer + Mare (Unterweger)	332	517	74	1	-	-	-

Table 11. Lockgrams in the dataset. The following settings in #LancsBox were used: relative frequencies, dispersion = Coefficient of Variation (CV), and Simple Math statistic.

The overlap between Unterweger's books and the autobiography in terms of n-grams is as follows: similarities between the autobiography and "Kerker" are more extensive than between the autobiography and "Mare"; however, as discussed above, some 1-grams are particularly uncommon in the reference corpus but appear in the autobiography and in Unterweger's known texts. The n-gram analysis is of importance insofar that it points towards the possibility of von Eisenstein having had some influence (be it active or passive) on the production of Unterweger's autobiography. The implications of these findings will be discussed below.

4. Discussion

This paper has attempted to investigate the questioned authorship of Jack Unterweger's autobiography. We were able to show that although Unterweger and von Eisenstein have different writing styles, the writing styles in the text under investigation is different from Unterweger's and von Eisenstein's other books, as it forms a separate cluster in the PCA and HCA. Since Unterweger's novels and poems cluster together, and his supposed autobiography is similar in genre to the novels, an influence of genre on the analysis can largely be excluded. Based on this, we attempted to identify features that make the autobiography different from texts known to be written by a particular author. Even though individual words were identified, no explanations for their appearance in the autobiography were found, other than that they contribute to the nominal writing style, which might be related to Unterweger's attempts to sound more formal and professional, as will be discussed below.

An influence of Sonja von Eisenstein in the use of n-grams might be indicated by the lockgram analysis. This analysis reveals several longer strings of words that appear

in von Eisenstein’s text and in the autobiography which do not appear in the reference corpus. The nature of Sonja von Eisenstein’s influence on the production of the autobiography remains unclear; it is also possible that the elements that point towards her involvement might have come from passive influence or from collaboration. It is known that she was an early supporter of Unterweger, and thus the close interaction between them might have left some traces in Unterweger’s own writing style. On the other hand, it is also known that Unterweger took some of her children’s stories and sold them as his own; consequently, some of her wording might have been integrated into his autobiography as well. Lastly, it is also a possibility that Sonja von Eisenstein played a role in the editing or revision of the book, which would also explain the presence of some of her stylistic features in the book. The results, however, remain inconclusive with regard to her involvement.

Our qualitative analysis has revealed that Unterweger employs a very distinctive writing style characterized by the creative use of words, particularly with respect to nouns and adjectives. This style is consistent across his novel-like books, and it can also be found in the autobiography. Examples illustrate his linguistic creativity in terms of non-standard words and expressions, colloquial, and informal expressions. In many instances, these words and expressions collide with an otherwise formal, literary, or intellectual style typified by a strong tendency towards nominalizations, which is generally characteristic of German “*officialese*” (Hennig, 2020). The incorrect use of some words, the use of unidiomatic expressions, and the preference for a nominal style might therefore indicate an attempt at appearing more educated than Unterweger actually was. This strategy might be considered a form of disguise, which is difficult to maintain in a longer text and often results in errors (Dern, 2008), such as those found in Unterweger’s books.

Based on our analyses, we conclude that the probability that Unterweger’s books and his supposed autobiography were produced independently is rather small. The lexical similarities provide strong evidence that Unterweger is indeed the (main) author of his autobiography. However, the autobiography still forms a neat separate cluster in the PCA and HCA, which indicates some stylistic differences. The nature of these differences cannot be conclusively explained, but a possible explanation is that Unterweger’s writing style had changed in the seven years between the publication of his autobiography in 1983 and his other novels in 1990. This difference could possibly be accounted for by the education he obtained while in prison, which included several writing classes. Moreover, it is also a possibility that, given that the autobiography was one of his first books (and definitely his most important), he put more effort into it than into the other books which were published when he had already become well-known. On the other hand, his poems, published in 1982, a year before his autobiography, are more stylistically similar to his later novels despite being of a different genre – a fact that again highlights that the autobiography is different.

Another promising direction for further research is the use of pronouns. As mentioned, all pronouns were removed from the analysis prior to testing, since they can be highly register and topic dependent and thus lead to biased results. However, since the use of pronouns can partially reveal people’s conceptualization of themselves and others (e.g., Orvell, Gelman, & Kross, 2022), they might be particularly interesting in the present context, as Unterweger was diagnosed with malignant narcissism and psy-

chopathy (Leake, 2010; Widler, 2019; see also Marko & Leibetseder, 2023). Additionally, the use of modal verbs and the perspectives taken in the books (past tense vs. present tense) appear to provide interesting grounds for future research.

5. Conclusions

This paper aimed at investigating whether or not there is linguistic evidence for the hypothesis that Jack Unterweger is not the (sole) author of his well-known autobiography “Fegefeuer oder die Reise ins Zuchthaus”. By using a combination of quantitative and qualitative methods, we were able to show that while differences between Unterweger’s other books and his autobiography do exist, many similarities can also be detected. The quantitative investigation has suggested stylistic differences between the autobiography and Unterweger’s novel-like books; the qualitative analysis has revealed lexical similarities, the latter of which provides support for the argument that Unterweger did in fact write the autobiography. The differences in style might be explained by the time that had passed between the writing of the autobiography and the writing of the other books. This finding, however, raises important (forensic) linguistic questions such as: how much does (or can) a person’s writing style change over a period of seven years? How much can writing classes (or other types of education) influence a native speaker’s language competence within seven years? The answers to these questions have practical consequences in terms of the comparability of texts produced several years apart. The analyses in this paper have also shown that the inclusion of pronouns in the analysis, even though they are often dependent on the genre, might actually be useful for certain authors and certain writing styles. In this specific context, it might even be argued that the inclusion of pronouns would be highly valuable, since Unterweger was diagnosed with psychopathy and malignant narcissism (Leake, 2010; Widler, 2019) – two psychological disorders that have been shown to be reflected in general language use (e.g., Fine, 2006; Viding, 2019), and in pronoun usage (e.g., Pennebaker, 2011) in particular. Therefore, another important avenue for future research in authorship analysis is how psychological disorders are reflected in written language, whether or not psychopathic individuals with high manipulative skills are also adept at manipulating their written language (e.g., in terms of adapting to others’ styles), and how pronoun use is influenced by genre, topic, audience, and psychological disorders.

To conclude, Unterweger’s writing style is characterized by distinctive and salient configurations of register, words and expressions. Whether or not his style can be considered “good” or “literary” is beyond the scope of this paper, but as Bandelow (2013, p. 24, translation) puts it, it is unlikely that “his [Jack Unterweger’s] books would [...] ever have had the slightest chance hadn’t there been the dark side to them that obviously bewitched the public.”

References

- Ainsworth, J., & Juola, P. (2018). Who Wrote This?: Modern Forensic Authorship Analysis as a Model for Valid Forensic Science. *Washington University Law Review*, 96(5). Retrieved from https://openscholarship.wustl.edu/law_lawreview/vol96/iss5/10
- Anthony, L. (2022). *AntConc*. Waseda University. Retrieved from <https://www.laurenceanthony.net/software/antconc/>
- Argamon, S. (2008). Interpreting Burrows’s Delta: Geometric and Probabilistic Foundations. *Literary and Linguistic Computing*, 23(2), 131–147. Retrieved 2024-03-02, from <https://academic.oup.com/dsh/article-lookup/doi/10.1093/llc/fqn003> doi: 10.1093/llc/fqn003
- Argamon, S. (2018). Computational Forensic Authorship Analysis: Promises and Pitfalls. *Language and Law / Linguagem e Direito*, 5(2), 7–37.
- Argamon, S., & Levitan, S. (2005). Measuring the Usefulness of Function Words for Authorship Attribution. In *Proceedings of the Joint Conference of the Association for Computers and the Humanities and the Association for Literary and Linguistic Computing* (pp. 1–3).
- Baayen, H., van Halteren, H., & Tweedie, F. (1996). Outside the Cave of Shadows: Using Syntactic Annotation to Enhance Authorship Attribution. *Literary and Linguistic Computing*, 11(3), 121–131.
- Bachmann, C. (2014). *Leichtlesbar.ch*. Retrieved from <http://www.leichtlesbar.ch/html/>
- Baker, P. (2006). *Using corpora in discourse analysis*. London & New York: Bloomsbury.
- Bandelow, B. (2013). *Wer hat Angst vorm bösen Mann?* Hamburg.
- Belvisi, N. M. S., Muhammad, N., & Alonso-Fernandez, F. (2020). Forensic Authorship Analysis of Microblogging Texts Using N-Grams and Stylometric Features. , 1–6.
- Biber, D., & Conrad, S. (2009). *Register, genre and style*. Cambridge: CUP.
- Biber, D., Conrad, S., & Reppen, R. (2006). *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.
- Binongo, J. (2003). Who wrote the 15th Book of Oz? An application of multivariate analysis to authorship attribution. *Chance*, 16(2), 9–17.
- Brezina, V. (2018). *Statistics in corpus linguistics. A practical guide*. Cambridge: Cambridge University Press.
- Brezina, V., Weill-Tessier, P., & McEnery, A. (2002). *#LancsBox*. Retrieved from <http://corpora.lancs.ac.uk/lancsbox>.
- Burrows, J. F. (1987a). *Computation into criticism: a study of Jane Austen’s novels and an experiment in method*. Oxford.
- Burrows, J. F. (1987b). Word-Patterns and Story-Shapes: The Statistical Analysis of Narrative Style. *Literary and Linguistic Computing*, 2(2), 61–70. Retrieved 2024-03-02, from <https://academic.oup.com/dsh/article-lookup/doi/10.1093/llc/2.2.61> doi: 10.1093/llc/2.2.61
- Burrows, J. F. (1992). Not Unless You Ask Nicely: The Interpretative Nexus Between Analysis and Information. *Literary and Linguistic Computing*, 7(2), 91–109. Retrieved 2024-03-02, from <https://academic.oup.com/dsh/article-lookup/doi/10.1093/llc/7.2.91> doi: 10.1093/llc/7.2.91
- Busch, V. (2019). Jack Unterweger. In *Psychologie des Guten und Bösen* (pp. 359–369). Berlin/Heidelberg: Springer.
- Coulmas, F. (1979). On the sociolinguistic relevance of routine formulae. *Journal of Pragmatics*, 3(3-4), 239–266. Retrieved 2024-03-02, from <https://linkinghub.elsevier>

- .com/retrieve/pii/037821667990033X doi: 10.1016/0378-2166(79)90033-X
- Coulthard, M. (2004). Author identification, idiolect and linguistic uniqueness. *Applied Linguistics*, 25(4), 431–447.
- Coulthard, M. (2007). By their words shall ye know them: on linguistic identity. In C. R. Caldas-Coulthard & R. Iedema (Eds.), *Identity trouble* (pp. 143–155). Palgrave Macmillan.
- Dern, C. (2008). Wenn zahle nix dann geht dir schlecht“. Ein Experiment zu sprachlichen Verstellungsstrategien in Erpresserbriefen. *ZFGL*, 36(2), 240–265. Retrieved 2024-03-02, from <https://www.degruyter.com/document/doi/10.1515/ZGL.2008.018/html> doi: 10.1515/ZGL.2008.018
- DWDS Kernkorpus. (2023). *DWDS Kernkorpus*. Retrieved from <https://dwds.de/d/korpora/kern>
- Eder, M. (2013). Does size matter? Authorship attribution, small samples, big problem. *Digital Scholarship in the Humanities*, 30(2), 167–182. Retrieved 2024-03-02, from <http://academic.oup.com/dsh/article/30/2/167/390738/Does-size-matter-Authorship-attribution-small> doi: 10.1093/llc/fqt066
- Eder, M. (2017). Short samples in authorship attribution: a new approach. *DH*, 1–5.
- Eder, M., Rybicki, J., Kestemont, M., & Pielstrom, S. (2022). “*Stylo*”: a package for stylometric analyses. Retrieved from https://github.com/computationalstylistics/stylo_howto
- Ehrhardt, S. (2018). Authorship attribution analysis. In J. Visconti & M. Rathert (Eds.), *Handbook of Communication in the Legal Sphere* (pp. 169–200). Berlin: De Gruyter. Retrieved 2024-03-02, from <https://www.degruyter.com/document/doi/10.1515/9781614514664-010/html> doi: 10.1515/9781614514664-010
- Fine, J. (2006). *Language in psychiatry. A handbook of clinical practice*. London: Equinox.
- Fobbe, E. (2011). *Forensische Linguistik*. Tübingen: Narr.
- Fobbe, E. (2020). Text-Linguistic Analysis in Forensic Authorship Attribution. *International Journal of Language & Law (JLL)*, Vol 9: Forensic Linguistics: New Procedures and Standards. Retrieved 2024-03-02, from <https://www.languageandlaw.eu/jll/article/view/78> (Publisher: [object Object]) doi: 10.14762/JLL.2020.093
- Fobbe, E. (2021). Forensische Linguistik. Eine kriminaltechnische Disziplin in Deutschland. *SIK-Journal – Zeitschrift für Polizeiwissenschaft und polizeiliche Praxis*(4), 18–27. Retrieved 2024-03-02, from https://doi.org/10.7396/2021_4_B doi: 10.7396/2021_4_B
- Fobbe, E. (2022). Authorship identification. In V. Guillén-Nieto & D. Stein (Eds.), *Language as evidence. Doing forensic linguistics* (pp. 185–218). Palgrave Macmillan.
- Grant, T., & Baker, K. (2001). Identifying reliable, valid markers of authorship: a response to Chaski. *International Journal of Speech, Language & the Law*, 8(1), 1350–1771.
- Grant, T., & MacLeod, N. (2018). Resources and constraints in linguistic identity performance: a theory of authorship. *Language and Law / Linguagem e Direito*, 5(1), 80–96.
- Gries, S. (2022). How to use statistics in quantitative corpus linguistics. In A. O’Keeffe & M. J. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 168–181). London/New York: Routledge.
- Gries, S., & Stefanowitsch, A. (2010). Cluster analysis and the identification of collexeme classes. In S. Rise & J. Newman (Eds.), *Empirical and experimental methods in*

- cognitive/functional research (pp. 73–90). CSLI Publications.
- Hennig, M. (2020). *Nominalstil*. Tübingen: Narr.
- Herring, S. C., & Paolillo, J. C. (2006). Gender and genre variation in weblogs. *Journal of Sociolinguistics*, 10(4), 439–459. Retrieved 2024-03-02, from <https://onlinelibrary.wiley.com/doi/10.1111/j.1467-9841.2006.00287.x> doi: 10.1111/j.1467-9841.2006.00287.x
- Herwig, M. (2022). *JACK. Gier frisst Schönheiten*. Retrieved from <https://www.ardaudiothek.de/sendung/jack-gier-frisst-schoenheiten/10385179/>
- Holmes, D. I. (1992). A Stylometric Analysis of Mormon Scripture and Related Texts. *Journal of the Royal Statistical Society*, 155(1), 91–120.
- Holzer, E., & Reibenwein, M. (2022). Jack Unterweger: Der Popstar unter den Serienmördern. Retrieved from <https://kurier.at/chronik/oesterreich/jack-unterweger-der-popstar-unter-den-serienmoerdern/401888249>
- Hoover, D. L. (2004). Testing Burrow’s Delta. *Literary and Linguistic Computing*, 19(4), 453–475.
- Johnson, A., & Wright, D. (2014). Identifying idiolect in forensic authorship attribution: an n-gram textbite approach. *Language and Law / Linguagem e Direito*, 1(1), 37–69.
- Juola, P. (2021, August). Verifying authorship for forensic purposes: A computational protocol and its validation. *Forensic Science International*, 325, 110824. Retrieved 2024-03-02, from <https://linkinghub.elsevier.com/retrieve/pii/S0379073821001444> doi: 10.1016/j.forsciint.2021.110824
- Kestemont, M. (2014). Function Words in Authorship Attribution From Black Magic to Theory? In *Proceedings of the 3rd Workshop on Computational Linguistics for Literature* (pp. 59–66).
- Koppel, M., Schler, J., & Argamon, S. (2009). Computational methods in authorship attribution. *Journal of the American Society for Information Science and Technology*, 60(1), 9–26. Retrieved 2024-03-02, from <https://onlinelibrary.wiley.com/doi/10.1002/asi.20961> doi: 10.1002/asi.20961
- Koppel, M., Schler, J., & Argamon, S. (2011). Authorship attribution in the wild. *Language Resources and Evaluation*, 45(1), 83–94. Retrieved 2024-03-02, from <http://link.springer.com/10.1007/s10579-009-9111-2> doi: 10.1007/s10579-009-9111-2
- Larner, S. (2014). A preliminary investigation into the use of fixed formulaic sequences as a marker of authorship. *International Journal of Speech Language and the Law*, 21(1), 1–22. Retrieved 2024-03-02, from <https://journal.equinoxpub.com/IJSLL/article/view/5898> doi: 10.1558/ijll.v21i1.1
- Layton, R., Watters, P. A., & Dazeley, R. (2015). Authorship analysis of aliases: Does topic influence accuracy? *Natural Language Engineering*, 21(4), 497–518. Retrieved 2024-03-02, from https://www.cambridge.org/core/product/identifier/S1351324913000272/type/journal_article doi: 10.1017/S1351324913000272
- Leake, J. (2010). *Der Mann aus dem Fegefeuer: das Doppelleben des Jack Unterweger*. Heyne.
- Locker, A. (2019). “Because the computer said so!” Can computational authorship analysis be trusted? *Journal of Language Works*, 4(1), 23–37.
- Marko, K., & Leibetseder, I. (2023). Linguistic Indicators of Psychopathy and Malignant Narcissism in the Personal Letters of the Austrian Killer Jack Unterweger. *Forensic Sciences*, 3(1), 45–68. Retrieved 2024-03-02, from <https://www.mdpi.com/>

- 2673-6756/3/1/6 doi: 10.3390/forensicsci3010006
- McMenamin, G. R. (2002). *Forensic stylistics*. CRC Press.
- Miranda García, A., & Calle Martín, J. (2007). Function Words in Authorship Attribution Studies. *Literary and Linguistic Computing*, 22(1), 49–66. Retrieved 2024-03-02, from <https://academic.oup.com/dsh/article-lookup/doi/10.1093/lc/fql048> doi: 10.1093/lc/fql048
- Moon, R. (1998). *Fixed expressions and idioms in English: a corpus-based approach*. Clarendon Press.
- Mose, J. (2015). *Jack Unterweger Interview im Gefängnis Stein Part 1*. Retrieved from <https://www.youtube.com/watch?v=kjYNVQjggto>
- Mosteller, F., & Wallace, D. L. (1984). The Federalist Papers As a Case Study. In *Applied Bayesian and Classical Inference* (pp. 1–15). New York, NY: Springer New York. Retrieved 2024-03-02, from http://link.springer.com/10.1007/978-1-4612-5256-6_1 (Series Title: Springer Series in Statistics) doi: 10.1007/978-1-4612-5256-6_1
- Newman, M. L., Groom, C. J., Handelman, L. D., & Pennebaker, J. W. (2008). Gender Differences in Language Use: An Analysis of 14,000 Text Samples. *Discourse Processes*, 45(3), 211–236. Retrieved 2024-03-02, from <http://www.tandfonline.com/doi/abs/10.1080/01638530802073712> doi: 10.1080/01638530802073712
- Nini, A. (2018a). An authorship analysis of the Jack the Ripper letters. *Digital Scholarship in the Humanities*, 33(3), 621–636. Retrieved 2024-03-02, from <https://academic.oup.com/dsh/article/33/3/621/4824843> doi: 10.1093/lc/fqx065
- Nini, A. (2018b). Developing forensic authorship profiling. *Language and Law / Linguagem e Direito*, 5(2), 38–58.
- Nini, A., & Grant, T. (2013). Bridging the gap between stylistic and cognitive approaches to authorship analysis using Systemic Functional Linguistics and multidimensional analysis. *International Journal of Speech Language and the Law*, 20(2), 173–202. Retrieved 2024-03-02, from <https://journal.equinoxpub.com/IJSLL/article/view/5925> doi: 10.1558/ijssl.v20i2.173
- Orvell, A., Gelman, S. A., & Kross, E. (2022). What “you” and “we” say about me: How small shifts in language reveal and empower fundamental shifts in perspective. *Social and Personality Psychology Compass*, 16(5), e12665. Retrieved 2024-03-02, from <https://compass.onlinelibrary.wiley.com/doi/10.1111/spc3.12665> doi: 10.1111/spc3.12665
- Overdorf, R., & Greenstadt, R. (2016). Blogs, Twitter Feeds, and Reddit Comments: Cross-domain Authorship Attribution. *Proceedings on Privacy Enhancing Technologies*, 2016(3), 155–171. Retrieved 2024-03-02, from <https://petsymposium.org/popets/2016/popets-2016-0021.php> doi: 10.1515/popets-2016-0021
- Pennebaker, J. (2011). *The Secret Life of Pronouns: What Our Words Say about Us*. Bloomsbury.
- Peters, A. (1984). *The Units of Language Acquisition*. Cambridge: Cambridge University Press.
- R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. Viena, Austria: R Foundation for Statistical Computing. Retrieved 2022-08-14, from <https://www.R-project.org/>
- Rude, S., Gortner, E.-M., & Pennebaker, J. (2004). Language use of depressed and depression-vulnerable college students. *Cognition & Emotion*, 18(8), 1121–

1133. Retrieved 2024-03-02, from <http://www.tandfonline.com/doi/abs/10.1080/02699930441000030> doi: 10.1080/02699930441000030
- Schmitt, N., Grandage, S., & Adolphs, S. (2004). Are corpus-derived recurrent clusters psycholinguistically valid? In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing, use* (pp. 127–151). John Benjamins.
- Solan, L. M. (2013). Intuition versus Algorithm: The Case of Forensic Authorship Attribution. *Journal of Law and Policy*, 21(3), 551–576.
- Stamatatos, E. (2009). A survey of modern authorship attribution methods. *Journal of the American Society for Information Science and Technology*, 60(3), 538–556. Retrieved 2024-03-02, from <https://onlinelibrary.wiley.com/doi/10.1002/asi.21001> doi: 10.1002/asi.21001
- Stamatatos, E. (2013). On the Robustness of Authorship Attribution Based on Character N-gram Features. *Journal of Law and Policy*, 21, 421–439.
- Swain, S., Mishra, G., & Sindhu, C. (2017). Recent approaches on authorship attribution techniques — An overview. In *2017 International conference of Electronics, Communication and Aerospace Technology (ICECA)* (pp. 557–566). Coimbatore: IEEE. Retrieved 2024-03-02, from <http://ieeexplore.ieee.org/document/8203599/> doi: 10.1109/ICECA.2017.8203599
- Turell, M. T. (2010). The use of textual, grammatical and sociolinguistic evidence in forensic text comparison:. *International Journal of Speech, Language and the Law*, 17(2), 211–250. Retrieved 2024-03-02, from <https://journal.equinoxpub.com/IJSLL/article/view/6001> doi: 10.1558/ijssl.v17i2.211
- Turell, M. T., & Gavalda, N. (2013). Towards an index of idiolectal similitude (or distance) in forensic authorship analysis. *Journal of Law and Policy*, 21(2), 495–514.
- Viding, E. (2019). *Psychopathy: A Very Short Introduction* (1st ed.). Oxford University Press. Retrieved 2024-03-02, from <https://academic.oup.com/book/28395> doi: 10.1093/actrade/9780198802266.001.0001
- Widler, Y. (2019). Psychiater Haller: “Gefährlich ist nicht der finstere Wald, sondern das eigene Heim”. *kurier.at*. Retrieved from <https://kurier.at/chronik/oesterreich/profiler-haller-gefaehrlich-ist-nicht-der-finstere-wald-sondern-das-eigene-heim/400637114>
- Wortbrücke. (n.d.). Retrieved 2022-08-08, from <https://www.onb.ac.at/oe-literaturzeitschriften/Wortbruecke/Wortbruecke.htm>
- Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge University Press.
- Wright, D. (2013). Stylistic variation within genre conventions in the Enron email corpus: developing a textsensitive methodology for authorship research. *International Journal of Speech Language and the Law*, 20(1), 45–75. Retrieved 2024-03-02, from <https://journal.equinoxpub.com/IJSLL/article/view/5944> doi: 10.1558/ijssl.v20i1.45
- Wright, D. (2017). Using word n-grams to identify authors and idiolects: A corpus approach to a forensic linguistic problem. *International Journal of Corpus Linguistics*, 22(2), 212–241. Retrieved 2024-03-02, from <http://www.jbe-platform.com/content/journals/10.1075/ijcl.22.2.03wri> doi: 10.1075/ijcl.22.2.03wri