

***Language and Online Identities: The Undercover  
Policing of Sexual Crime***

**Tim Grant and Nicci MacLeod (2020)**

Cambridge University Press. 195 pp

*Reviewed by Emily Chiang*

It is hard to imagine a more crucial context for exploring the relationship between language and identity than the undercover policing of online sexual crime. In this book, Tim Grant and Nicci MacLeod draw on their extensive experience of working with undercover police units to demonstrate how writers perform various aspects of identity (age, gender, experience, etc.) through linguistic choices, and how understandings of linguistic identity performance can assist in undercover policing tasks, forensic authorship analysis and other forensic contexts. Drawing on this work, they bring together divergent theories of identity to develop their novel and unified ‘resource-constraint’ model and demonstrate its practical and theoretical implications. All the research presented in this book is done against a backdrop of clearly explained legal context and thoughtful discussions on the operational and theoretical ethics of undercover work and tasks engaged in by the forensic linguist.

The book is organised thematically, beginning with an introductory chapter describing the problem of online child sexual abuse and the practical, legal and ethical difficulties associated with policing these sorts of crimes. It offers an overview of research in computer-mediated discourse and identity, raising important

---

**Affiliation**

Aston University, UK  
email: e.chiang2@aston.ac.uk

questions around identity in online spaces, perhaps the most central of these concerning our capacity to 'be whoever we want to be' online. The authors introduce their own *resource-constraint* model which reconciles the cognitivist notion of idiolect ('every native speaker has their own distinct and individual version of the language they speak and write ...' (Coulthard 2004: 432)) with the resource model of the 'linguistic individual' (individuals draw on a range of resources developed over the course of their sociolinguistic histories (Johnstone 1996; Kredens 2002)). The resource-constraint model, then, proposes that individuals are both enabled and constrained by the various resources available in a given moment of interaction, including (but not limited to) sociolinguistic history, the immediate communicative context, the communities we partake in, and physicality (including cognitive function). The authors demonstrate how this understanding of identity relates to the various tasks that forensic linguists might assist the police with, namely authorship profiling, comparative authorship analysis, and identity assumption.

Chapter 2 ('Data and Methods') outlines the various data sets used in Grant and MacLeod's research, before describing the authors' methodological approaches. Data sets include instant messaging (IM) conversations between i) a convicted child sexual offender and multiple victims, ii) large groups of offenders, iii) undercover police officers (UCOs) and their targets, iv) UCOs and their trainers in a role-playing exercise, and v) participants (university students and UCOs) under experimental conditions. The impressive range and diversity of data types explored make for extremely rich analyses, enabling the authors to tackle theoretical and practical questions around online identity construction and obfuscation, including, for example, the specific linguistic cues that alert individuals to identity deception. The rarity of the data sets obtained for this research speaks to the authors' unique position of working closely with policing units and is part of what makes this research truly novel. Grant and MacLeod's methodological approach is based on Herring's (2004) computer-mediated discourse analysis toolkit and centres on language analysis at four levels: structure, meaning, interaction and social behaviour. Through illustrative examples from their data, the authors demonstrate how interactants perform various aspects of identity at each level, from low-level features like spelling variants to discoursal features such as conversation openings and closings. The final section of this chapter is dedicated to a discussion on ethics. In addition to considerations around research ethics (relating to the protection of victims, suspects, offenders, participants and researchers), the authors importantly discuss at some length the issues inherent to the activities involved in undercover policing. They reasonably conclude that undercover work is morally justifiable 'in cases of serious crime as a proportionate response to the harm done to victims ...' (p. 56).

Chapter 3 ('Experimental Results') first provides an overview of research on variation in online discourse, highlighting the importance of perceived authenticity in online interactions in light of the numerous resources provided by digital media for 'identity play' – a useful term the authors provide to mark the deliberate and conscious performance of a particular identity (Tagg 2015). Against this backdrop, Grant and MacLeod describe their experimental work designed to determine participants' accuracy in detecting the point at which a known online interlocutor had been replaced by another impersonating the first. This aspect of their work demonstrates the complementarity of experimental, elicited linguistic data with that often described as 'naturally occurring'. Through these experiments, the authors were able to establish the distinct linguistic features that caused participants to suspect a change of interlocutor (and therefore those most salient in identity construction) as well as participants' confidence levels in their assertions, and the effects of different preparation conditions for the task (see Figure 3.1 on p. 68). In focusing solely on participants' abilities to detect *when* a switch has taken place, however, an opportunity is missed; arguably it would be equally useful to assess participants' competence in detecting *whether* a switch had taken place at all, and this might be addressed in future research. One of the most important ideas to come from the experimental research is the notion of 'identity leakage' – the observation that while individuals may successfully emulate some aspects of a target's language, residual linguistic habits indicative of their 'home identity' can be difficult to suppress, resulting in the performance of 'hybrid identities' (p. 82). This finding is particularly crucial for UCOs, for whom the suppression of the institutional police officer identity is paramount when attempting to impersonate, for example, a 14-year-old abuse victim, or indeed an abuser.

Chapter 4 ('Training Identity Assumption') describes the linguistic identity training that Grant and MacLeod developed and delivered as part of a national training course for UCOs. The training programme, heavily informed by the resource-constraint model, aimed to assist officers in the various undercover tasks involved in investigating online child sex abuse, including the infiltration of abusive online communities and the assumption of individual target identities, i.e., offenders and victims. The programme involved educating police officers in language analysis at the levels of structure, meaning and interaction, and evaluating their performance before and after various online identity assumption tasks. Crucially, the authors demonstrate that linguistic training led to an overall improvement in UCOs' competence in accurately impersonating a target identity by, for example, adopting the target's vocabulary items, speech acts and conversational topics. In short, they had successfully acquired some of the linguistic resources necessary to convincingly assume a target identity. Grant and

MacLeod are careful, however, not to overstate their findings, and highlight those areas where further training is required. Common problems included the use of 'extended runs of interrogatives' (p. 106) and difficulty initiating sexual topics, even where this was a demonstrably usual behaviour for the target victim identity. It is clear that observable findings like these really evidence the practical application of linguistic analysis to the policing of online crimes. Regarding the initiation of sexual topics, the authors raise an important discussion on the possibility of UCOs acting as agents provocateurs (the English and Welsh law counterpart to the US entrapment defence) by instigating or participating in sexual conversation. They highlight the 'unexceptional opportunity' test that examines whether police presented a defendant with more than an unexceptional opportunity to commit a crime (*R v. Loosely* [2001]), as being particularly important. They argue convincingly that being able to demonstrate that initiating sexual topics is within the usual repertoire of a victim's linguistic behaviour (as their data sometimes shows) works towards a reasonable defence against accusations of acting as an agent provocateur. Another interesting point arises with the novel term 'authorship synthesis' (p. 92), which Grant and MacLeod have coined to refer to the creation of credible and consistent linguistic identities and describe as an inversion of forensic authorship analysis (wherein the main goal is to identify an author). Conceiving of identity assumption and authorship analysis as two sides of the same coin makes clear that any theory of linguistic identity will necessarily address both tasks, and perhaps encourage broader thinking in their execution.

Reflecting on the outcomes of the experimental research and police training, Chapter 5 ('Resources and Constraints in Abuse Identity Performance') revisits the discussion on language and identity, expounding the authors' resource-constraint model in detail. In particular, it demonstrates how those aspects of identity especially pertinent in child abuse interactions (age, gender, relationships and abuse communities) are discursively constructed, and describes the types of resources and constraints which impact participants' capacity to perform these roles. All of this work is illustrated with snippets of transcripts from the authors' various data sets. One of the main take-home messages from this discussion is that identity performance is bounded. Even when online spaces afford so much opportunity for identity play, we are, according to Grant and MacLeod, limited by (among other things) the extent of our sociolinguistic resources. But importantly, we are able to expand our repertoires of available identities by acquiring new resources; the authors show UCOs doing this consciously through training and exposure to online conversations involving their target identities.

In light of previous discussions, Chapter 6 ('Contexts for Linguistic Investigative Advice') considers how forensic linguistics can continue to assist in online investigations, with particular reference to the tasks of comparative authorship

analysis, sociolinguistic profiling and online identity assumption. Each task is discussed in terms of its broad purpose(s) and illustrated with examples from resolved cases, for example, the identification of a notorious child sex offender in part through his use of the unusual greeting term ‘hiyas’. These examples give the reader a more concrete picture of the sort of casework that a forensic linguist might undertake. The discussion importantly shows that forensic linguistic work goes beyond authorship analysis and the provision of courtroom evidence, and the authors strongly get across the message that linguistic analysis, whether evidentiary or investigative, should be rigorous and rich in explanation.

The final chapter (‘Implications and Future Directions’) explores the implications of Grant and MacLeod’s work for identity assumption tasks, considers theoretical questions around linguistic identity and highlights areas in which future work is most needed. Here, the unified nature of the ‘neither deterministic nor wholly interactionist’ (p. 167) resource-constraint model is really emphasised as the authors reiterate that identity is born of both an individual’s sociolinguistic history and cognitive capability. Interestingly, it is only at this point that Grant and MacLeod begin to tackle the question of persistence, i.e., if identity is fluid and changeable, how do we construct a consistent identity across moments of interaction or over the course of many interactions? The answer offered is that some of our available resources (e.g., the immediate communicative context, audience, etc.) are more dynamic and subject to change, while others (e.g., our sociolinguistic histories, physicality and cognitive functioning) are more stable. This is a useful concept which could have been introduced earlier on in the book, giving the reader a chance to consider these different dynamic and stable resource types while working through the chat log examples. Regardless, it is another crucial insight for UCOs tasked with constructing and maintaining credible online personas and has further implications for forensic authorship and profiling tasks.

This small book presents an impressive amount of research, tackling big questions around the nature of identity and its construction online, and demonstrating how an understanding of linguistic identity can support police work in the highest-stakes scenarios. Grant and MacLeod present some novel concepts with clear implications for undercover policing, in particular, the problem of identity leakage (see p. 168) and the description of the whole linguistic individual as performed at all levels of language production (not just conspicuous surface features). One of the most important things this book does is demonstrate the successful acquisition of linguistic resources by UCOs and their consequent improvement in assuming target identities, because it is one way in which the authors show that the dark web is not, as some believe, unpolicable.

One area that might have received more attention is the perspective of the non-UCO interlocutor in undercover scenarios and how UCOs’ performances

are perceived. The UCOs' accuracy in adopting specific target forms is well demonstrated and provides good evidence for their success in assuming other online identities, but do the identity positions that UCOs attempt to perform (e.g., age, abusing experience, etc.) always land as intended from the perspective of the other interlocutor? Given the importance of UCOs maintaining a credible online persona, more discussion might be dedicated to the perlocutionary aspect of these interactions.

There is much that is novel about this book, not least the incredibly rare types of data it draws on. But its real strength lies in the careful interweaving of research, application and theory, presented with conscientious detail throughout. The authors provide relevant legal context for the investigative scenarios discussed (i.e., police and courtroom processes, statutes), examples from real forensic linguistic casework, and textual examples from the data to support their observations. All of this results in a rich and rigorous account of online identity construction in the context of online child abuse, presented in a clear, methodical and self-critical way. In particular, while the authors show forensic linguistic techniques to be powerful in a range of investigative scenarios, they are always careful not to overplay their significance, making clear that linguistic analysis most commonly plays a small, auxiliary part in a larger picture. Its theoretical and practical contributions will no doubt make this book a valuable resource for undercover police and a longstanding reference work for any researchers with an interest in linguistics (or any subdiscipline), identity and undercover police practice.

## References

- Coulthard, M. (2004) Authorship identification, idiolect, and linguistic uniqueness. *Applied Linguistics* 25(4): 431–447. <https://doi.org/10.1093/applin/25.4.431>
- Herring, S. C. (2004) Computer-mediated discourse analysis: an approach to researching online behavior. In S. A. Barab, R. Kling and J. H. Gray (eds) *Designing for Virtual Communities in the Service of Learning* 338–376. New York: Cambridge University Press. <https://doi.org/10.1017/CBO9780511805080.016>
- Johnstone, B. (1996) *The Linguistic Individual: Self Expression in Language and Linguistics*. Oxford: Oxford University Press.
- Kredens, K. (2002) Idiolect in authorship attribution. In P. Stalmaszczyk (ed.) *Folia Linguistica Anglica* vol. 4, 191–212. Łódź: Łódź University Press.
- Tagg, C. (2015) *Exploring Digital Communication: Language in Action*. Abingdon: Routledge. <https://doi.org/10.4324/9781315727165>