

Chapter 4 Training Identity Assumption

Introduction

As described above, since 2011 the authors have been involved in the UK national 'Pilgrim' training programme for specialist online investigators. This course was designed by individuals in a Technical Intelligence Development Unit in a UK police force and run by a private police training company with strong links to that force. This course has now been replaced by other training with slightly different cohorts of trainees, but for the period of this study trainees on the programme were already experienced undercover police officers. They were, however, new to undercover operations in the online environment. The programme covered a number of issues including legal authorisations and ethics, open source intelligence gathering, technical learning – for example of Bitcoin, and TOR networks – and alongside this we provided linguistic training to assist with online identity assumption. Once trained, officers within this role engage in a variety of tasks and in a variety of policing contexts, and one of these contexts is the policing of online child abuse and exploitation in all its forms.

Within the context of online child sexual abuse there are a number of undercover tasks that can be undertaken. One such task is the infiltration of online fora and chat rooms dedicated to the production and exchange of child abuse images and videos. In such fora officers may be tasked to interact with offenders as if they too were offenders interested in acquiring indecent images. The purpose of such operations is to gather intelligence to help identify victims so that they can be rescued, to identify offenders so that they can be arrested and prosecuted, or, if neither of these aims is possible, to engage in general intelligence gathering on the networks and methods of such individuals and to disrupt their activities. Where an infiltration operation has been successful in identifying an offender, and where that offender is arrested, a situation may arise where news of the arrest is kept quiet for a period of time, and then a UCO may be tasked with assuming that offender's identity. In these

identity assumption or account takeover operations the UCO might be authorised to continue the interactions of the arrested offender, to gather further intelligence, seek new arrests and ultimately secure the rescue of further victims.

A different kind of operation can arise when a child has been discovered engaging in sexual activity online with an anonymous offender – the scenario set out at the beginning of this book. This sexual activity might be discovered by a parent or carer, who then reports it to the police. In these situations the child and the offender may have been communicating many times a day over many days, and such an intervention could mean a sudden and unexplained break in communication as far as the offender is concerned. From the child's point of view they will be taken offline and to a place of safety such that they can be offered proper care and support, and they will also be interviewed by specially trained officers. Where the child is unable to help identify the offender, then a police officer may be tasked with replacing that child in their online interactions. In this situation the UCO may have a limited time period to study the identity of the child in order to assume their identity. Only part of the analysis phase of the operation will be linguistic. The UCO must also master details of the child's life including their family and school life, and from any captured chat logs understand what the child has told the offender and *vice versa*. Chiang and Grant (2017) show that suspicious offenders can and do engage in conversational moves involving the assessment and management of risk, which sometimes occurs through quizzing the 'children' on aspects of their previous interactions. For the linguistic portion of their analysis, the UCOs must understand a specific child's performed identities within the interaction so that they can successfully deceive the offender. The primary objectives for the UCO is to collect intelligence to identify the offender and to identify any further children at risk. Where possible they may, as the child, try to arrange to meet with the offender so that the offender can be arrested. Under Section 15 of the Sexual Offences Act 2003, and amendments contained in Section 67 of the Serious Crime Act 2015¹ in order to address preparatory offences, a person commits an offence if they arrange to meet someone they believe to be under the age of 16 where they intend to engage in sexual activity with that child. In these cases the chat log between the UCO and the offender might

be used to provide evidence both of knowledge of the child's age and intention of sexual activity. If this can be achieved in the conversation there can be the additional benefit of protecting the child from the further harm of the court appearance, as the UCO alone can then appear as the prosecution witness against the offender.

Under UK law these two different tasks, of infiltration and account takeover, require separate and specific authorisations, and the legal learning on the Pilgrim course addresses the issues of both law and ethics in these operations. Chapter 6 discusses this legal context in more detail. A training distinction can also be drawn, on the one hand between the infiltration tasks where an officer invents a linguistic persona and accompanying background 'legend', and on the other hand identity assumption tasks where a UCO is asked to take over a particular individuals' account and interactions. Both these tasks can be conceived of as a reversal of the more traditional forensic linguistic task of authorship analysis (see e.g. Grant, 2013) and might be thus be characterised as authorship *synthesis*. Authorship synthesis is distinct from authorship obfuscation where a writer is simply attempting to hide points of their own linguistic style as a countermeasure to potential authorship analysis. Although authorship synthesis is different from authorship obfuscation it might of course be a strategy to achieve obfuscation. For linguistic legend-building the focus is on creating a credible linguistic identity through which the UCO can perform consistently over time and which is sufficiently different from other legends they may have built and different from the UCO's own everyday identity performances. The risk for the UCO is that their various legends and true identity might be linked as originating from the same individual. For the identity assumption task, authorship synthesis relies on an analysis of the target identity's previous interactions so that they can be successfully described, understood and later performed by the undercover officer. The risk here for the UCO is that of noticeable inconsistency with the target identity which would make them vulnerable to detection.

The linguistic input to the Pilgrim training comprises five hours of taught input through which some basic linguistic insights and analysis tools are provided. The trainee UCOs are taught an analysis protocol that enables them to systematically

analyse an interactant's contribution to an online chat. The protocol is built round a structured pro-forma for authorship analysis derived in part from Herring's (2004) levels of analysis for CMC as described in Chapter 2. This input session includes a substantial amount of practice analysis and culminates in an assessed three-hour simulation exercise, described below.

Data

The training in language analysis for identity assumption was developed using the general corpus of chat logs of genuine child sex abuse interactions, as described in Chapter 2. The data discussed in this chapter, however, are the chat conversations generated through the simulation exercise carried out at the end of the Pilgrim language analysis training, described here and elsewhere as the 'training' data. These comprise IRC conversations between trainee UCOs and their trainers, collected via the Yahoo! chat client as used during the Pilgrim training programme. The simulation exercise begins when the UCOs' published training schedule for the day is apparently disrupted as if for an urgent operation. The trainees are told that a 14-year-old girl had been discovered engaging in a sexually explicit IM conversation with an unknown adult, that the offender is due online in two hours' time, and that the offender wants to arrange to meet the girl for sex for the first time. The trainees are provided with chat logs of the conversations apparently captured from the girl's computer (these were in fact anonymised chat logs from the 'Genuine IM' data set and referred to in training and in this chapter as the 'historic' chat logs). The trainees receive a briefing that they are to assume the child's identity and engage with the offender in order to gather intelligence which might help identify the offender, and ultimately to arrange a meeting at which the offender could be arrested. Working in pairs the trainees are thus given a two-hour time frame to prepare for the encounter, and must make their own prioritisation decisions about how much of this time they spend on linguistic analysis as opposed to the other forms of analysis that they have been taught as part of the programme. Each pair then has a one-hour online chat (half an hour each) during which they engage with the 'offender'. The offenders' parts in the interaction are played by a group of the trainers, situated in a different room in the building. During the simulation exercise,

a number of operational complications are created and must be dealt with appropriately. To pass the Pilgrim programme the trainees have to pass the simulation exercise, and to pass the simulation they must pass an assessment based on their competence at linguistic identity assumption.

A pre/post training model was adopted for the evaluation described below, and thus for each participating UCO there exist two sets of chat logs. One set was collected prior to any linguistic training being received. In this condition the trainees were drawing on their previous training and experience as offline UCOs. The second set of chats was collected three months later at a repeat of the simulation exercise, which occurred after the linguistic training had been delivered.

Because of the practicalities of running the study within the context of the set training programme the pre- training set is structurally different from the post-training set. In the pre-training dataset there are five chats with a single trainee UCO, and one chat between the 'offender' and a pair of trainee UCOs. In this latter chat the point at which one trainee takes over from the other is clearly marked in the text. In the post-training set, there are six conversations, all of which are produced by a pair of trainee UCOs each with the point at which one takes over from the other clearly marked in the text (the instruction to swap over is explicitly provided by the assessor approximately half-way through the conversation). Because of these differences and because the two groups do not comprise exactly the same set of individuals the pre-post comparison is carried out at the group level rather than being matched at the individual level. The length of the chat logs ranges from three to nine sides of A4. The historic and pre- and post-training logs were all analysed using bespoke software tools developed as part of the study.

As described above in both the pre- and post-training conditions the UCOs had a period of time allocated for preparation for the task. For these preparation sessions and during the subsequent online chats researchers on the project sat quietly in the room, making unstructured notes observing how the UCOs addressed their tasks and noting any commentary they made on the linguistic analyses. These notes provide a further dataset, which contributes to an understanding of the chat logs themselves.

Framework for Analysis

As discussed above in Chapter 2 Herring (2004:18) suggests that CMD can be useful analysed at the four levels of structure, meaning, interaction, and social behaviour. The training we provide is structured through these same categories with a view to providing a sufficiently complete model of a linguistic individual for identity assumption and the assessment of students in both their simulation exercise and in our evaluation of their learning uses the same structures. For ease the table from Herring reproduced in Chapter 2 is provided again here.

Table Error! No text of specified style in document..1: Four domains of CMD (adapted from Herring, 2004:18)

	Phenomena	Issues	Methods
Structure	typography, orthography, morphology, syntax, discourse schemata	Genre characteristics, orality, efficiency, expressivitiy, complexity	Structural/Descriptive Linguistics, Text Analysis
Meaning	Meaning of words, utterances (speech acts), macrosegments	What the speaker intends, what is accomplished through language	Semantics, Pragmatics
Interaction	Turns, sequences, exchanges, threads	Interactivity, timing, coherence, interaction as co-constructed, topic development	Conversation Analysis, Ethnomethodology
Social behaviour	Linguistic expressions of status, conflict, negotiation, face-management, play; discourse styles, etc.	Social dynamics, power, influence, identity	Interactional Sociolinguistics, Critical Discourse Analysis

The task presented to the trainees is essentially an authorship analysis task, and as we observed in Chapter 3 most current research and practice in authorship analysis

principally occurs at Herring's structural level of analysis. This may be because most reported authorship analysis questions concern largely monologic texts, which has perhaps driven the discipline towards focussing on low-level features at the structural level. Some work has been done at a functional level of analysis (Argamon and Koppel, 2012; Nini and Grant, 2013). However, because the use of structural features has been shown to be sufficiently effective in many attribution tasks, there appears to have been little motivation to include analysis of discursive patterns in addressing traditional authorship attribution problems.

When approaching questions of identity assumption in authorship synthesis it is clearly important to be able to replicate the target's language at the structural level. However the question then arises as to whether disguise can be successful in a case where structural level features are accurately reproduced but patterns of meaning, interaction and social behaviour differ from the target's past performance. As well as structural level analyses our training focuses on pragmatic analyses rooted in speech act theory (Searle, 1969), to account for some aspects of meaning, and analyses of topic introductions and responses via a simplified form of conversation analysis (Sacks, 1972) and floor management analysis (Gumperz, 1982). In the case of these three levels of analysis the aim was to provide training to enable the trainee UCOs to carry out the analyses in a limited timeframe and this, along with other operational constraints, requires that the nature of the training input at each level of analysis is somewhat simplified.

With regard to the level of analysis Herring labels 'social behaviour', this did not form part of the linguistic element of the Pilgrim training. This decision was made for two reasons. First, because the trainees on the Pilgrim programme are already experienced UCOs in offline encounters, meaning this aspect of interaction is to a degree considered already part of their expertise. Second, other aspects of the Pilgrim programme include psychological inputs which address issues of rapport building and persuasive techniques, which were considered more directly useful to the trainee UCOs and which might be considered to be at this level of 'social behaviour'. Related linguistic research is, however, being undertaken as part of the wider AIO programme investigating offenders' strategies and communicative

purposes using Swales' (1981) move analysis as a framework (see Chiang and Grant 2017, 2018).

In this chapter we thus take each of *Structure*, *Meaning* and *Interaction* in turn as three separate areas for analysis. For each of these three levels, we first present our analysis of the historic chat log describing some of the features identified and then examine how well the trainee UCOs performed in aligning their language to reproduce these features before and after linguistic training.

The Structural Level

Lexical and Morphological Analysis of the Target Persona

The structural analysis of chats was based the feature taxonomy from MacLeod and Grant (2012) described in Chapter 2. Although the original taxonomy has been substantially developed and modified since this earlier research, it covers the key structural feature types and provided a sound point of departure for the current work.

An overview of some of the victim's style choices in relation to a selection of some of the structural features identified appear in Table 4.2 below.

Table Error! No text of specified style in document..2: Examples of victim's vocabulary and spelling choices

Feature	Examples used by victim
Initialism	brb; nm (not much); tbh
Vowel Omission	yh; dnt ; wht; dwnstairs; bk; pls
Shortening	sis; pics; probs; sec; convos; morro; cause; cuz (for cousin); cam
Emoticon	:L
Substitution	u; r; y
Prosodic stylisation	noooo; noo; plsss
Phonetic stylisation	yup; yeah; thanx; aint (but also haven't); nahh; nope; outta; dunno; gunna; wanna; sorta; kinda
Misspelling	there (for they're); too (for to); of (for have)

g- clipping	fingerin; tossin BUT nothing; minging; anything; doing (x2); showing; leaving; blackmailing; watching; talking; fucking; sorting; something; working; joking.
--------------------	--

One of the clear findings arising from a structural analysis, is that a flat list description of any individual's vocabulary is insufficient to capture this aspect of their identity. As shown in Table 4.3 below, this individual uses three variants for the word 'please', the standard spelling which she uses twice, 'pls' with vowel deletions has four uses, and 'plsss', used just once, is similar to this but with 's' repetition for prosodic emphasis. This is typical. We commonly find that individuals use different variants for the same term. As we shall see, this ratio of variation for individual terms may be a significant style feature over and above the occurrence of any specific spelling.

Table Error! No text of specified style in document..3: Variants of use in victim's vocabulary

Word	Variant
Are	Are (3), r (1)
Because	Because (1), cuz (1)
Dnt	dont: 5, dnt: 1
Know	know: 4, kno: 1,
No	no: 23, nahh: 5
Ohh	ohh: 1, ohhh: 1
Ok	ok: 6, k: 2
Please	pls: 4, please: 2, plsss: 1
So	so: 5, soo: 1
Ur	ur: 8, ya: 3
why	why: 8, y: 5
would	would: 6, wud: 2
yes	yh: 28, yes: 3
you	u: 45, you: 5

This proportionate observation of an individual's variation in use can be extended to develop understanding of individualised lexico-grammatical patterns or rules for a particular studied identity. For the above identity one such rule might focus on the degree to which g-clipping is used as a stylistic choice— 'g-clipping' is the omission of the -g from the -ing ending. We observe broad differences between individual use where some individuals apply g-clipping in both verbs and in words that are not verbs but that happen to end -ing (e.g. 'anything'). In contrast other

individuals might clip just verbs, or even just a few specific verbs (such as 'goin'). Further individual variation can be observed such that for some individuals g-clipping can be noted as occurring more narrowly, that is to say occurring in one particular lexico-grammatical function (e.g. when the word is used as a participle but not as a gerund) or for others it can be used more generally disregarding the grammatical context of the -ing ending.

As can be seen in Table 4.2, the victim's language in this text largely avoids g-clipping with the two exceptions being 'fingerin' and 'tossin'. It may be notable that the two verbal occurrences of g-clipping are both main verbs for sexual acts. The context of occurrence of 'fucking' is adjectival and for the other verbs in the list the full -ing ending is used only in progressive forms. For this chat log it would be difficult to describe this as a strong rule for the individual, the data is simply too slim to warrant such a conclusion, however observation of the pattern does allow the UCO to follow and extend the habits of the target identity. It should also be noted that the paucity of data in this case is absolutely typical for UCOs engaging in language analysis for identity assumption - improvising from marginal observations such as these is the required task.

Further study of the use of vocabulary makes apparent that semantic or pragmatic contextual use is also vitally important to understand. For example, a vocabulary analysis (for the offender in the historic chat log described above) showed a 3:2 ratio of use between the terms 'orgasm' and 'cum'. Further analysis revealed that for this individual the former was only used in context of the female child interactant but the latter referred to himself. We might speculate that this offender had a gendered understanding of the two terms. Other offenders in our data can be seen to use the two words interchangeably for either gender. As a pattern of vocabulary use this too can be used as marker of individual style in the chat logs that needs to be used and extended by the UCO.

Thus we can see that in the same way that a detailed vocabulary analysis of an individual's style can be important in authorship analysis cases, it is also of great assistance in the authorship synthesis task. By carrying out an analysis such as this the UCO can become more familiar with the individual persona and use this

information to be consistent with the historic chats when they attempt to emulate their identity performances in the interactions with the offender.

Structural Level Performance Pre- and Post-training.

In the pre-training condition untrained UCOs were observed to devote some effort to the observation and notation of vocabulary used, and notes taken by researchers suggest that even without any linguistic training the UCOs were aware that accuracy at this structural level would be important in the identity assumption task. UCOs were thus observed to write lists of useful vocabulary items and phrases they noted from the historic chat logs. None of the UCOs were observed in this pre-training condition to make detailed notes of how specific terms were used in context or to note the proportion of use of different variants of a specific term.

Structural Analysis of the Pre-training Chats

Although some linguistic preparation by the UCOs was observed, their actual performance as the victim appeared to take little note of these preparations. Table 3 below provides just a few examples from the feature categories identified above. For each category we give examples from the historic chat of the victim’s identity as performed in the historic chat and then examples of how the trainee UCOs are consistent or inconsistent with that identity when attempting the identity assumption task. In each case we have aggregated the UCOs’ language choices.

Table Error! No text of specified style in document.4: Examples of UCOs' structural features prior to linguistic training

Feature	Historic chat	UCOs' use - consistent with historic chat.	UCOs' use - inconsistent with historic chat.
Initialism	brb; nm (not much); tbh	brb	lol [5 of 7 UCOs used 'lol']; wuu2; wbu; atm
Substitution	u; r; y	U	b4; 4ward; 4; 2; C; l8r

			[4 of 7 UCOs used digit substitutions]
Phonetic stylisation	aint (but also haven't); dunno; gunna; kinda; nahh (x2); outta; sorta; plsss; thanx; wanna; yup; yeah;	aint; dunno; wanna;	betta; BOVRD; gonna; innit; lata; loool; ma; numba; propa; sed; SOMFIN; summat; sumtimes; wana; waz; wen;
Vowel Omission	yh; dnt ; wht; dwnstairs; bk; pls	Wht	CNT; fuk; gt; hrd; luk; nxt snd; tke wil; WNT;

Table 4.4 suggests that, prior to any linguistic training, the trainees stereotype and overestimate the victim's use of what they characterise as 'netspeak' or 'textspeak' likely to be used by a teenaged girl. In their use of initialisms the trainee UCOs make use of a markedly wider range of choices than is evident in the victim's historic chats. Five of the seven UCOs have the victim using 'lol' which never occurs in the historic log. While the historic log shows that the victim never uses a number to substitute a syllable, three of the trainees make use of '4' in this manner, as well as a use of '8' in 'l8r'. In general the historic log shows the victim's phonetic stylisations to be representative of fairly standard speech, while a number of the trainees select spellings with strong accent stylisation, as well as various representations of the sound of laughter, which do not appear anywhere in the victim's own writing. Furthermore, while the victim's use of vowel omissions is only apparent in six separate words in the historic chats, there is a substantially higher level of vowel omission in the trainees' attempts to recreate her style. Finally, some trainees shift into uppercase part way through their conversations, while the victim makes no use whatsoever of upper case in her writings. The failure to accurately assume the victim's identity is also evident in their use of g-clipping, where, despite the victim employing this feature in only 3 out of 21 possible instances, all the trainees used the dropped form more often than not.

The issue before training was not only this stereotyping of online linguistic forms. Trainees also failed to accurately replicate spellings of individual terms using for example 'mom' instead of 'mum', 'wat' or 'wot' instead of 'wht', and so on. There was little observation of punctuation features; whereas the historic log showed minimal use of commas and apostrophes, some of the UCOs could not help but use a near standard form of punctuation.

Structural Analysis of the Post-training Chats

The linguistic training on structural features demonstrated how to analyse chat logs to elicit the kinds of features described above. This resulted in considerable improvement but not entirely accurate replication of the persona from the historic chat logs. Taking the same feature categories as used above we can see some of the patterns of improvement.

Table Error! No text of specified style in document..5: Examples of UCOs' structural features after linguistic training

Feature	Historic chat	UCOs' use - consistent with historic chat.	UCOs' use - inconsistent with historic chat.
Initialism	brb; nm (not much); tbh	brb ; nm	atm; lol [2 out of 7 UCOs used 'lol']
Substitution	u; r; y	u; r; y	b; 4; 2morro; 2nite [2 out of 7 UCOs now used digit substitutions]
Phonetic stylisation (excluding g-clipping)	aint (but also haven't); dunno; gunna; kinda; nahh (x2); outta; plsss; sorta; thanx;	aint; gunna; kinda; nahh; plsss; thanx; wanna;	2mora; beta; gonna; innit; lata; loool; ma; numba; propa; sed; sumtimes; wana; waz; wen;

	wanna; yup; yeah;		
Vowel	yh; dnt ;	bk; dnt; pls;	cnt; forwrđ;
Omission	dwnstairs; bk; pls wht;	wht	fuk; gd; gt; hm; hrd; jst; knw; kl; luk; nd (for need); nything; nxt; skl; snd; thn; thr; tht; tke tme; txi; wil; whr; wnt; yng

The main finding of the vocabulary analysis described in Table 4.5 is that training does improve the ability of UCOs to identify and adopt the victim’s vocabulary choices. In this post-training condition all of the UCOs analysed show considerable improvement, and not only with the selection of appropriate terms. UCOs also improved in observing the proportionate use of some words; as well as picking up the use of the more unusual variants such as ‘dnt’ and ‘pls’ the trainee UCOs now alternate these forms with the more standard spellings ‘don’t’ and ‘please’. For any individual UCO there was insufficient data to tell whether they were using the terms in comparable ratios to the historic chat, but that there was some degree of alternation was a clear improvement.

In contrast to this positive finding we also noted that the rate of use of terms which were inconsistent with the historic chat barely reduced. In the first condition the UCOs tended to use terms, which were either derived from their stereotype of the victim’s chat (or possibly from their own style). As Table 4.5 demonstrates, examples of these inconsistent items persisted post-training and particularly in feature categories which produce the non-standard spellings stereotypical of ‘netspeak’. In terms of vocabulary selection the trainee UCOs largely achieved a hybrid style including terms typical of their original attempts but also with more of the actual historic vocabulary.

With regard to derivation of morphological rules the UCOs vocabulary showed that most of them noticed that the victim used both some g-clipping and some vowel

deletion. There was an issue here however with the overextension of these. As discussed above the trainee UCO's tend to overuse both vowel omissions and g-clipping in the pre-training condition. Part of training encouraged trainees to spot these kinds of pattern in the historic chat usage and these featured in some preparation notes. The result of this however was that in the post-training condition there were even more examples of vowel deletion than in the pre-training condition. A similar pattern could be observed for the extent of g-clipping, noticing that the victim had used g-clipping in the historic chat seemed to be the basis of a rule but the trainees failed to build into the rule the fact that the victim had considerable use of the standard non-drop form too.

Overall vocabulary training was shown to have some positive effect on the trainee UCO's ability at this structural level of identity assumption but clearly there is more work to be done. Subsequent training events are already feeding these findings back to the UCOs to encourage them further to respond accurately to the historic chat logs and suppress their stereotypes of linguistic performance. With a view to the resource theory of identity proposed throughout this book the difficulty experienced by the UCOs can be characterised as being that they are still drawing on either their own resources for their habitual identity performances, or on their own resources which feed into stereotypes of the teenaged girl online. The aim of the training is to enable the UCOs to develop a new resource that specifically derives from analysis of the historic chat log and is specifically drawn on in their performance of the assumed identity. Their success or failure in the identity assumption task can be helpfully characterised in terms of how well they develop and use such a resource.

The Meaning Level

Speech Act Analysis of the Target Persona

Herring's (2004) framework identifies a level of analysis of the 'Meaning' in an interaction. The current project has a minimal interest in how individual terms are used to acquire meanings in the interaction but considerable interest in how 'intent' is performed between interactants in online chat, and in Herring's words 'what is accomplished through language' (2004: 18). As for the other levels of analysis, the

task was to use a form of analysis that would sufficiently capture this aspect of interactional meaning whilst being straightforward enough to offer in an effective way to the trainee UCOs.

As described in Chapter 2 our solution to this has been to focus on a high-level categorisation of speech acts. The categories used are as follows:

Statements or Assertives: Conversational turns that have the potential to be 'true' or 'false' because they aim to describe a state of affairs in the world.

Instructions or Directives: Conversational turns that attempt to make the other person's act in a certain way. These can vary from weak or tentative requests through to direct and coerced orders.

Interrogatives: Refers to requests for information, usually in the form of a question.

Commissives: Conversational turns that commit a speaker to a future course of action. Again the commitment can be specific and strong, as in a promise, or weak or vague. Threats, seen as a promise to harm, can be seen a particular example of commissive.

Expressives: Conversational turns that express an attitude or feeling about an person, object or situation. Expressive intent is often manifest in single word turns *e.g.* 'lol', 'cool' and *emoticons* are often used expressively in online chat.

Acknowledgements: Statements that acknowledge another participants' contribution. This is included to account for the minimal receipt turns observable in the data, *e.g.* 'OK', 'Yes'.

This categorisation system recognises that that different expressions can be used to express different several types of pragmatic force and the types of force within an utterance can carry different weights. It is a platitude of speech act theory that structural assertives can be pragmatically used as directives but some UCOs were observed to struggle to grasp this aspect of language use. Thus in our data there is a turn where the offender writes to the victim 'u have too much clothes on though'. Given the context we would encourage the UCOs to code this utterance as

containing assertive force, expressive force and also directive force — in the latter case recognising that it might act as a request, to remove some items of clothing. When analysing the language use of a particular individual the teaching encouraged the trainee UCOs to code all the speech acts that might be present and where possible to rank them as to which speech act was primary in any one utterance. In the broader context of the conversation the utterance ‘u have too much clothes on though’ was primarily an observation, in this case was coded primarily as an assertive .

A second feature of this kind of analysis is that *patterns* of speech act use are specific to different conversational pairings and situations. That is to say, observable patterns between any two interactants are a function of that particular relationship and may be specific to it. Thus in the historic chat logs the offender in the first instance developed a moderately sexualised relationship with the victim (to the extent that she shows him her breasts on webcam). The child has been groomed at this stage to believe that she is in a relationship with this ‘Offender 1’ and she believes that this is normal behaviour for such a relationship. A pattern of interaction develops which can be characterised using the speech act coding system and which maintains a degree of consistency throughout the chats between the girls and ‘Offender 1’. Later in the interaction the offender adopts a second online identity, and from within this identity starts interacting with the child. From the girl’s understanding this is a completely different person – ‘Offender 2’. Using this second identity the offender claims to the victim that he has hacked her earlier chats and has obtained compromising pictures of the victim. ‘Offender 2’ says that he will distribute these pictures through the girl’s social media networks unless she does as he directs. In the interactions with this second identity the observable patterns of speech acts from the victim are, as might be expected, very different. Within the original dyadic relationship the girl can be observed to be relatively consistent in her use of speech acts with ‘Offender 1’, but with ‘Offender 2’ this pattern is disrupted and a new distinct pattern of interaction develops. Interestingly there is no equivalent shift at the structural level, e.g. in vocabulary use between these two conversational phases, either on the part of the victim or the offender.

Victim's Speech Acts in the Historic Chat Log

The historic chat log is thus broken down into different sections. In the initial sessions with the offender's first identity, the victim can be observed to use a mixture of assertive statements, minimal acknowledgements of the offender's contributions, and occasional interrogative questions: 'where u live?' and directives: 'go on cam but dnt be dirty sis is next to me'. When confronted by the offender's second identity and threats to publish intimate pictures, she responds with a series of interrogatives: 'why are u on my fb?', 'wht do u want then'. There are also a number of directives here in the form of requests: 'yes just pls dont post them' and assertives 'i do have 2 sisters 4 and 7'. There then follows a sexually explicit exchange with Offender 2 comprising on the victim's side mainly assertives and minimal acknowledgements. Later when engaging in sexual activity with Offender 2, use of expressives become more frequent: 'i wanna fuck u', 'this feels good' and commissives 'yh want me too use something too finger', 'wanna hear'.

There is further complexity created when the offender briefly also engages with the victim using two further identities. The final conversations, between the victim and the offender's third identity, comprises mainly assertives from the victim with some interrogatives such as: 'do i know u' before she realises that it is the same person as 'Offender 2'.

Given such complexity it is hard to describe the victim's characteristic use of speech acts but there are some patterns that do emerge. The victim uses a significant proportion of directive as well as commissive speech acts and even within the coercive interaction she attempts to bargain (using utterances which contain a combination of directive and commissive force). At the start of the coercive interaction with Offender 2 there is a long run of interrogatives where she is trying to get information about his identity and intentions. She is also highly expressive (mostly about the sexual acts). Perhaps one main characteristic of this individual is that she draws on a wide range of linguistic resources expressing herself through a mix of different speech acts. This is not true of all the individuals we have analysed; others demonstrate a preference of use of just one or two types of speech act in these types of interaction.

With regard to the resource model the victim's natural variation in identity performance between interactions with apparently different offenders raises a point of interest. This natural fluctuation in identity through interaction is entirely expected but nevertheless begs the question of the range of performances available to the victim across a wider range of interlocutors. Just as when considering the structural level of analysis it was possible to note that different spelling variants were used in differing proportions we may speculate that an individual has a range of preferences of how to act in different interactions. To explore this performative range further will require examination of one or more individuals interacting with a large and varied number of conversational partners. Whilst this might make an interesting research question, this kind of variation is not provided for by the project's data and so remains unstudied here.

Speech Act Performance Pre- and Post-training

Pragmatic Analysis of the Pre-training Chats

Prior to training the UCOs showed very little awareness of potential variation in the speech acts and their analysis notes show no evidence that they considered anything like this in their preparation for engagement. Perhaps because of this lack of awareness the trainee group show considerable individual variation in their engagement as the girl.

Several trainees, for example, uses a high proportion of interrogatives in an attempt to pin the offender down to a time and location for their meeting, and to elicit his phone number. This contrasts with the victim's historical chat, where there tended to be a more even mix of speech acts, (with the exception as noted with the start of the conversation with Offender 2). For other trainees there is a clear difficulty in playing the role of the victim and using directives. In the historic chat the girl does tell the offender what to do, including directing him in online sexual activity. Another aspect where the trainees perform less well is in their use of expressives as a way of deflecting the apparent suspicion of the offender.

Where some trainee UCOs clearly struggle to consistently assuming the girls identity against this pragmatic criteria others do naturally perform better. One

trainee was observed to use a fairly high but appropriate number of directives: 'i wana meet u propa'; 'giv me ur numba'; 'wana lose my virginity', and also a number of interrogatives: 'who r u agan', 'how will i kno its u'. The way these were used was in keeping with the victim's online identity as recorded in the historic chat logs.

Pragmatic Analysis of the Post-Training Chats

The described variation in ability to assume the victim's identity at the pragmatic level clearly marks a training need and post training there was more consistency with the historic chat logs but also some individual UCOs who clearly struggled with their analysis and performance at this pragmatic level.

The best trainees could be observed in the preparation phase of the simulated operation using the linguistic input and attempting to understand better the way the victim used language in the interactions. In our feedback to the trainees we note that for one UCO *'A good mix of assertives, directives and commissives, in keeping with the style of the victim. The interrogatives are well spaced, and refer to the proposed sexual activity that evening as well as to travel arrangements and requesting contact details.'* Other trainees, however, persisted in using extended runs of interrogatives that are not generally characteristic of the historic chat. This tendency may well relate to the operational tasking of intelligence gathering (should the attempt at meeting the offender fail) but it does mark a point of difference between the actual persona of the girl and the officer assuming that persona. As such it marks a point of potential discovery.

These findings are clearly mixed and we maintain the need for training in this feature of identity assumption, and are looking to develop the training we offer to better develop the UCOs skills in this area.

The Interaction Level

Topic Development and Control by the Target Persona

Moving on to topic management (at Herring's (2004) 'interaction' level), the aim here was to provide a form of analysis for topic development and control that would allow for a good understanding of the target identity whilst remaining simple and

flexible enough to be applied in the operational context. As with the speech act analysis discussed above, topic analysis is a function of the relationship between the interactants. It is our view that this makes it no less a feature of the linguistic individual's persona as these personae are developed through interaction.

As well as ensuring consistency in identity performance there are strong operational reasons for accuracy with regard to identity assumption with regard to topic control. One concern is that undercover officers might leave themselves open to accusations of entrapment or acting as *agents provocateur* if they are seen to instigate or participate in sexual conversations whilst acting the part of the child. The operational issue is however that some children, including those represented in our wider data sets, have been so effectively groomed and sexualised that they instigate sexual topics of conversation and themselves initiate online sexual activity. Failure to act consistently with this past behaviour when performing as the child might risk alerting the perpetrator to the victim's replacement by the UCO. Indeed we have seen in our genuine data sets suspicions raised by offenders where UCOs are reluctant to engage in sexual activity that has been previously been normalised between offender and victim.

Whilst other jurisdictions may vary, entrapment is not *per se* a defence in English and Welsh law, rather statute and case law frames this issue as a potential abuse of process affecting the fairness of trial. The legal effect of finding that a UCO acted as an *agent provocateur* would thus be that the evidence of the chat log might be declared inadmissible or a prosecution might be 'stayed'. *R v Loosely (2001)* is the most important recent judgement in this area. In *Loosely* Lord Nicholls writes that one issue is:

whether the police did no more than present the defendant with an unexceptional opportunity to commit a crime. I emphasise the word unexceptional. The yardstick for the purpose of this test is, in general, whether the police conduct preceding the commission of the offence was no more than might have been expected from others in the circumstances. Police conduct of this nature is not to be regarded as inciting or instigating crime,

or luring a person into committing a crime. The police did no more than others could be expected to do.

(R v Loosely [2001] UKHL 53, para 23)

This 'unexceptional opportunity' test is important in the current context. By carrying out a full analysis of the child's topic initiations and following the patterns derived from the historic chat, the UCO can put themselves in a strong position against any allegations of acting as *agent provocateur*. By evidencing how the child has previously acted they can go further than the test of doing 'no more than others are expected to do'. Through their analysis of topic development the UCO can govern their own interactions as the child by doing no more than *this child* has previously done.

Given these thoughts a simple tripartite framework for topic analysis was developed to record for each individual topics initiated in the historic chat, topics responded to and developed and topics which when introduced by the other participant were avoided or which provoked only a minimal response.

Topic Development and Control in the Historic Chat Log

As with the previous areas of analysis topic control was coded for the historic chat log. Table 4.6 shows the topic control patterns of the victim in her interactions with the offender (who himself is performing as Offender 2). Amongst the list of topics initiated are sexual topics including online sexual activity, and she also engages in sexual activity at the instigation of the offender.

Table Error! No text of specified style in document..6: Topic control in the historic chat log

Topics initiated	Topics responded to / developed	Topics avoided or attracting minimal response
Sexual contact with offender, Own arousal, use of objects. Offender's location health and appearance and	Video conversation - clothing and sexually explicit content. Offender's 'hacking'. Sister's age etc. Offender's instructions for sexual behaviour on camera.	Teasing (sister in room). Sexual activity involving sister.

**motives (for blackmail).
Own appearance (self-deprecating).
Arranging webcam chats**

Topic Analysis of the Pre-training Chats

As shown in Table 4.7 the chat logs of the trainee UCO's prior to the linguistic input show significant differences at the level of interaction to the target persona described above. In the historic chat log the child introduces sexual topics and sexual activity on several occasions. In all but one case the trainee UCO failed to do this. For some trainees, not only did they not initiate sexualised conversation they declined it when attempted by the offender. This natural reluctance to engage in online sexual activity whilst performing as a 14-year-old needs to be overcome by UCOs in these tasks, and some find this difficult to achieve. An explicit learning objective of the simulation exercise is to facilitate officers in doing this more easily whilst staying within their authorisation and avoiding going further than activity and discussions analysed in the historic chat.

Table Error! No text of specified style in document..7: Topic control in the pre-training chats

Topics initiated	Topics responded to / developed	Topics avoided or attracting minimal response
Sexual chat (only one trainee initiated sexual chat) Meeting including location of meeting, time of meeting, travel arrangements etc Offender's phone number and location.	Cameras Victim's clothing Some sexual talk, planned sexual acts, masturbation, arousal, sexual activity.	Sexual activity Victim's arousal, Victim's masturbation, oral sex, anal sex. Putting webcam on.

**Experience of 14 yr olds.
Offender's clothing and description,
offender's name
Offender's arousal
Being nervous**

A further feature of the pre-training chat is the nature and quantity of topics of operational interest to the UCOs. These types of operation can be very pressured and UCOs need to focus on a number of tasks simultaneously. As well as performing the child's identity in a sufficiently convincing manner they must try to obtain information about the offender which might identify them, be explicit about the child's age so that the offender falls within the terms of the Sexual Offences Act, try and arrange a location to meet away from other children who might be endangered by a sexual predator, and get a description of the offender so that they can be easily recognised at the meet. All these tasks may create points of inconsistency with the child's previous chat, and it is part of the skills that the officers develop to work these new topics naturally into the conversations.

Topic Analysis of the Post-training Chats

After training the officers showed consistent improvement in this area and particularly appreciated how language analysis can protect against accusations of acting as *agents provocateur*. As can be seen in Table 4.8 the points of inconsistency typically involved introduction of the operational issues as discussed above. In both pre- and post- training chats an additional operational issue was that the students did not respond to attempts by the offender to start a web cam conversation. This again reflects operational constraints.

Table Error! No text of specified style in document..8: Topic control in the post-training chats

Topics initiated	Topics responded to / developed	Topics avoided or attracting minimal response
Sexual activity initiated by the	Travel and plans for the evening.	Sharing webcam.

victim was much more apparent across all trainee UCOs with UCO proposing sexual activity and expressing sexual arousal. Offender's arousal. Present from offender - the offender had previously promised a present when they met. Victim's mum - an argument with the mother was proposed as a reason why the victim could not access her mobile phone. Travel arrangements Offenders clothing School uniform Victim's virginity School in morning Being nervous Offender's number.	Proposed activity Location of meeting / Hotel location Sexual activity talk Victim's location	sexual	Possible call	phone
--	--	--------	---------------	-------

What Arouses Suspicion?

The principal purpose for analysing the simulation data from the training exercise has been to examine how well identity assumption can be achieved in terms of emulating the language patterns of the target identity, and to examine which areas of training require further attention. In addition to this textual comparison with the historic chat log before and after training, the data collected allows us to interrogate the chat logs to identify what aspects of failure in identity assumption might be noticed. Whilst this is a more central question with the experimental data discussed in chapter 3 these simulation chats are also useful in this regard. The trainers who played the offender in the simulation exercise were explicitly told to challenge the trainee UCOs who were playing the victims if they noticed aspects of the identity

performance that were inconsistent with the previous performances of the child victim. In addition to the evidence of these challenges within the text of the chat logs the trainers also provide an evaluation of each trainee. Some of these evaluations mention linguistic successes or failures in identity performance.

Below is a selection of examples taken from trainers' turns in the role-players' chat that indicate they may be flagging behaviour from the trainees that have the potential alert a genuine offender to the deception. Bold indicates that the issue may have to do with **structural** issues such as spelling and/or vocabulary choices; underlining indicates the problem seems to be with topic, and italics indicate the issue is with *pragmatic force*.

This first set of examples are taken from the pre-training condition:

i can tell ur on ur best behavior **cuz ur typin is betta** than norm

what's wrong wif you today? (In response to limited sexual talk and *persistent questioning* about meeting location and offender's phone number)

whts up/; what u playing at; pissing me off (in response to trainee's *requests* for number and claims that battery has run out)

you *askin* a lot today; whats up wiv u you aint not been horny wiv me b4 is summat wrong; wtf y u no talk sexy wiv me; who are u? you taked and played sexy b4

no worri cutey u no scared when u on cams so what change

why you keep *asking* who i am?

you better fucking come; you said you loved me last week, now you aint got your phone and you wnt c2c; i dont feel good about this

In the chat log text there were few explicit challenges regarding the structural level of analysis – spellings, punctuation abbreviations etc. – although in the trainer's comments these aspects were clearly noted. Thus one trainer notes 'over use of text speak – not so much in original' and another trainer notes 'heavy use of punctuation'. These comments match our own comparison with the historic chat

logs in the pre-training condition, and we can speculate that the trainers find it harder to call out the trainees on this aspect of their writing whilst within the role play of the simulation exercise. It is clear from the examples given above, however, that trainers also flagged up *pragmatic* and topic issues as being likely to cause suspicion. One trainer wrote 'lots of questioning as a result felt very cold and clinical'. Comments such as these, along with the within-chat log challenges, suggest an assumption by trainers that a genuine offender would be alerted by language features such as a higher than usual degree of interrogatives, or by a reluctance to engage with sexual topics. Chapter 3 provides some support to this assumption.

As noted above the post-training set of interactions showed a reduction in structural level identity performance errors, and none were commented on by the trainers in their evaluations in the post-training condition. Again, *italics* indicate where the problem appears to be one relating to *pragmatic* force and underlining indicates the issue is with topic:

ur annoying me wont cam *dictating* where I have to go

u ok not like u not to talk dirty is this the sis or what?

There was a notable reduction in linguistically focussed challenges in the post-training condition and remaining challenges focused either on the content of previous chat or more behavioural differences observed by the trainers:

what school did you say

you hurt ur hand, u got slower at typing?

i fink ur avoiding me wont cam wont phone

if u cant jump in taxi then u not gonna do stuff we talked about..watsa goin on wt
u????

Behavioural differences with the historic performance (or between two UCOs alternating at playing the victim) are not issues that can be addressed with linguistic training. This combined with suggests a marked improvement in the trainees'

ability to emulate the style of the victim demonstrated by our own analysis suggest some success in the training.

Evaluation

The principal finding of this chapter is that linguistic identity assumption is challenging but can be trained. By taking a principled linguistic approach to the training we can hope to ensure that different aspects of identity performance can be analysed and then emulated and thus complement the skills of UCOs to provide more convincing and less detectable identity assumption. The aims of the trial were to evaluate the usefulness of the linguistic model for adopting a persona, and to determine when and to a lesser extent discover the basis on which trainers/instructors might detect or become suspicious of identity assumption by the trainee operative. It is our view that the provision of the structured linguistic analysis was shown to be trainable to these non-linguists and that subsequent to training we saw improvement at different levels of linguistic analysis. Prior to the linguistic training, the trainees tended to concentrate their efforts at identity assumption almost exclusively by mimicking vocabulary features, but in doing so they tend to overestimate the victim's use of 'netspeak' or 'textspeak' spellings. Subsequent to training this stereotyping was generally reined in and most students performed better and could recognise that the target identity would use a range of variants for some terms and that the ratio of use could be copied. Further to this prior to training the individual pragmatic patterns of use, the observable patterns in turn taking and topic control were almost entirely neglected by the UCOs in their attempts at identity assumption. After training we've shown the UCOs can emulate patterns of use in these areas and that doing so is important to avoid detection. The feedback from this trial was thus positive, indicating improved performance in assuming an identity using the model, and that detection of this identity disguise was more difficult.

In addition to this general positive conclusion we are also more aware of where training needs to be improved or changed (principally in the area of pragmatic analysis where students most notably struggled) and also potential areas where this training may be extended. We are, for example, considering whether a better

understanding of overarching strategies or linguistic moves analysis (see Chiang and Grant, 2017) would improve the identity assumption. When considering any changes to the currently delivered programme we must also bear in mind the practicalities. In terms of the wider Pilgrim training programme the linguistic training is a small element, and needs to be balanced in terms of time and effort with other requirements for online UCO training. Perhaps more importantly the linguistic analysis which is trained needs to be useable by UCOs in the field. Over-elaboration might render the system less usable and less useful, so any additional suggested analysis would need to be shown to improve the identity assumption.