Linguistic analysis of suspected child sexual offenders' interactions in a dark web image exchange chatroom

Chiang, Emily; Nguyen, Dong; Towler, Amanda; Hass, Mark; Grieve, Jack

[Link to publication on Research at Birmingham portal](#)

Linguistic analysis of suspected child sexual offenders' interactions in a dark web image exchange chatroom.

Emily Chiang, Aston Institute for Forensic Linguistics – corresponding author
Aston University
Birmingham
B4 7ET
e.chiang2@aston.ac.uk

Dong Nguyen, Alan Turing Institute
British Library
96 Euston Road
NW1 2DB London
dnguyen@turing.ac.uk

Amanda Towler, Hyperion Gray
atowler@hyperiongray.com

Mark Haase, Hyperion Gray
mhaase@hyperiongray.com

Jack Grieve, University of Birmingham
Department of English Language and Linguistics
University of Birmingham
Edgbaston
Birmingham
B15 2TT
j.grieve@bham.ac.uk

Abstract

It is increasingly recognised that child sexual offenders convene in dark web spaces to exchange indecent imagery as well as advice and support. In response, law-enforcement agencies deploy undercover agents to pose as child sexual offenders online to gather intelligence and evidence on these offending communities. Currently, however, little is known about how offenders interact online, which raises significant questions around how undercover officers should 'authentically' portray the persona of a child sexual offender. This paper presents the first linguistic description of authentic offender-offender interactions taking place on a dark web image exchange chatroom. Using a linguistic framework for discourse analysis known as Move Analysis*, we analyse the rhetorical strategies of chatroom users. We then model the move sequences of different users and user types using Markov chains, allowing us to make comparisons between their linguistic behaviours. We find the predominant moves characterising this chatroom are Offering Indecent Images, Greetings, Image Appreciation, General Rapport and Image Discussion, and that different rhetorical strategies are employed by users depending in part on their experience in exchanging indecent images on the dark web.*

Keywords

1

1   Introduction

Child sexual abuse (CSA) has received increasing attention from researchers since at least the 1980s (e.g. Finkelhor, 1984; Marshall & Barbaree, 1990), and in the last few decades, much of this work has focused specifically on online contexts (e.g. O'Connell, 2003; Craven et al., 2007; Briggs et al., 2011; Black *et al.*, 2015; Kloess *et al.*, 2017; Grant & MacLeod, 2018; Chiang & Grant, 2018). Although most research in this area comes from psychology, there is growing interest in linguistics, much of which focusing on computational methods of language analysis which aim to detect and classify 'grooming' conversations (e.g. Pranoto et al., 2015; al-Khateeb & Epiphaniou, 2016; Gunawan *et al.*, 2016; Cardei & Rebedea, 2017). So-called 'grooming' conversations seem to have taken centre-stage in online CSA research from all domains, in part because of the wealth of data available from sites like perverted-justice.com, which display transcripts of interactions between convicted child sex offenders and adult 'decoys' posing as child victims (see, e.g. Black *et al.*, 2015; Pranoto et al., 2015; Chiang & Grant, 2017). Aside from the various limitations of working with data featuring adult decoys rather than genuine victims, the overwhelming focus on offender–victim interactions has obscured other important types of online interaction that serve to facilitate CSA — namely, offender–offender interactions.

There is growing acceptance that adults with a sexual interest in children regularly convene in online spaces to exchange indecent imagery of children (IIOC), as well as advice and support regarding abusive behaviours and practices (Davidson & Gottschalk, 2011; Westlake & Bouchard, 2016). The 'dark web' — a collection of encrypted websites including chatrooms and fora that enables users to retain complete anonymity — facilitates interactions between offenders and grants a new level of protection against detection by law enforcement (MacLeod & Grant, 2017). Combating the problem is therefore extremely difficult, particularly because tech-savvy offenders are able to move quickly across different fora, switching between any number of usernames and identities. In response, law enforcement agencies around the world are increasingly using undercover officers (UCs) to pose as offenders online (Urbas, 2010; Mitchell et al., 2012; MacLeod & Grant, 2017). This permits the gathering of intelligence and evidence on the individuals operating in these spaces and the illegal activities they might be engaging in (e.g. producing and consuming IIOC), which in turn supports the disruption of offending networks and prosecution of suspected offenders.

One of the main challenges to this approach is that CSA-focused dark web communities tend to be highly suspicious of potential law enforcement infiltration (MacLeod & Grant, 2017), and so it is paramount that UCs tasked with posing as offenders are able to do so as 'authentically' as possible. Furthermore, UCs' freedom to perform as child sexual offenders is hindered by various operational constraints. For UCs in the UK (and other jurisdictions), one of the most significant of these — especially in communities in which IIOC exchange is an expected practice – is that they are not legally sanctioned to possess or share IIOC (Martellozzo, 2015). It is easy to see how this could hamper the attempt to perform as interested and enthusiastic CSA offenders.

In online environments where few paralinguistic cues are available, the success of UCs is largely dependent on their careful use of language. As such, there is a clear need for linguists to help law enforcement combat crimes in this area. In particular, we know very little about how offenders interact with each other online, so our understanding of how UCs might attempt to

communicate as offenders is limited. Of course, abuse-focused interactions between CSA offenders typically do not occur in open, clear web spaces, which makes access difficult. To our knowledge, only two such studies have analysed dark web interaction between CSA offenders. McManus *et al.* (2016) explored the themes of conversations between child sexual offenders using content analysis methods, finding dominant themes of child sexual interest and rapport, as well as adult relationships, the media, and the sexual self. The authors compared the use of these themes between offenders who had engaged in contact abuse and those who perpetrated online abuse exclusively, and found no significant differences between the two groups. A second study by Chiang (2018) had a slightly narrower focus, and explored through a form of discourse analysis known as move analysis (Swales, 1981; 1990) a corpus of forum posts written by self-identifying 'newbies' attempting to join existing communities of CSA offenders on various dark web sites. Chiang found performances of competence and expertise in relation to offending behaviours to be a central element of newbies' attempts to persuade other forum users into granting community membership.

Understanding how language is used in these dark web spaces provides a basis for supporting UCs in the task of performing as CSA offenders online. Studies on adult-child interactions dominate research in this domain, while linguistic behaviours of whole CSA communities remain poorly understood, and it is this that drives our overall research aims. Greater insight into CSA interactions at the community level can support the disruption of large groups of offenders; a task which is not well addressed by analysing texts involving a single offender. Primarily, we will provide the first description of the common rhetorical strategies used in online interactions between suspected child sexual offenders, based on a move analysis (Swales, 1981; 1990) of a sample of chat-logs from a single chatroom. These chat-logs provide a rare opportunity to gain insight into how groups of suspected CSA offenders (who we refer to as either 'suspected offenders' or (chatroom) 'users') interact with each other online, making them a clear area of interest for law-enforcement, particularly in the undercover context. We will also compare the rhetorical strategies used by different types of users to better understand the communicative landscape in the chatroom and the different persona types encountered in this community. Two particularly important groups comprise those offenders who offer images and those who exert authority, as these users are likely among the most harmful in terms of perpetuating illegal image-sharing (and, in turn, contact abuse), and the most influential on the community as a whole. Our specific research questions are as follows:

1. What are the rhetorical moves used by suspected CSA offenders in group interactions and how frequently are they used?
2. What are the common move sequences of suspected CSA offenders in group interactions?
3. What are the differences in move sequences used by image offerers and non-image offerers?
4. What are the differences in move sequences between authority figures and non-authority figures?

In the remainder of this paper, we first describe our dark web corpus and the methods we used to describe the rhetorical structures in this corpus and to compare the rhetorical strategies used by different user types. Specifically, in addition to introducing move analysis, we will describe techniques for quantifying the rhetorical structures characterising different user types by measuring transition probabilities between moves and then for visualising, examining and comparing these transition matrices, using Markov chains. We then present the results of our analysis, including our full move system for this domain, and the comparison of image offerers and non-offerers and authority figures and non-authority figures using these visualisation techniques. Finally, we conclude by considering the practical consequences of our results for

law enforcement agencies conducting undercover investigations in this domain, and by discussing directions for future research.

2 Methods and Materials

2.1 Data

The dataset analysed for this study consists of transcripts of chat-logs scraped from a single dark web chatroom focused on the exchange of IIOC, posted between 2015 and 2016. The data were collected by Hyperion Gray (https://www.hyperiongray.com). Notably, all illegal content (i.e. images, videos, gifs) were stripped from the database, ensuring that there was no risk of exposure to researchers during analysis. We have chosen to refrain from providing any real detail about the chatroom itself, other than to note that its primary purpose is to facilitate the exchange of IIOC. We have chosen not to name the chatroom, or make our dataset publicly available. All textual examples we provide to illustrate the moves observed have been carefully selected so as to avoid giving any identifying or explicit/disturbing details. This means that examples may have been clipped from longer utterances or be generic approximations of actual utterances. All usernames, victim names, dark website names, place names and image links (already disabled by the data provider) have been removed and replaced with descriptive terms, e.g. *subject name*, *site name*, *link*, etc.

We extracted 200 samples from the site by collecting the content of every 50th chat-log page, working backwards from the most recent page. Chat-log pages each consist of 100 consecutive time-stamped conversational turns, except for the final page which has 79 turns, and often involve ten or more users. In total, our corpus contains a total of 5,590 users, 19,974 turns and 21,770 transcript lines.

This systematic approach to data sampling enabled us to build a broad general picture of the conversational topics and linguistic strategies employed in these group interactions, which is important to our primary goal of providing an overall linguistic description of the chatroom for undercover investigative purposes. However, we recognise certain limitations with selecting our data in this way: for example, 200 consecutive chat-log pages would have enabled us to account better for patterns of ongoing interaction including issues of turn-taking and topic management. Having said this, systematic sampling still allows us to trace individual users' linguistic strategies and consider the more typical moves and move transitions of users of various 'types'.

2.2 Move Analysis

Move analysis is a linguistic framework for discourse analysis that seeks to capture the discrete communicative functions performed by particular segments of language in a given communicative context (Swales, 1990). Moves represent the overall goals of a text and may be broken down into lower-level goals (termed 'strategies' (Bhatia, 1993)) which work to achieve those moves. The method was originally developed to describe academic genres for pedagogical purposes (Moreno & Swales, 2018), the most famous example being set out in Swales' (1990) Create A Research Space (CARS) model, which illustrates the moves and strategies (which he termed 'steps') of the introductory sections of academic papers.

Moves	Steps

1. Establishing a territory	1. claiming centrality, and/or 2. making topic generalisations, and/or 3. reviewing items of previous research
2. Establishing a niche	1. counter-claiming, or 2. indicating a gap, or 3. question-raising, or 4. continuing a tradition
4. Occupying the niche	1. outlining purposes, or 2. announcing present research 3. announcing principle findings 4. indicating RA structure

Table 1. CARS model for research article introductions. (Adapted from Swales, 1990, p. 141).

The model has since gained popularity in numerous areas of research and been applied to a diverse range of genres, including online genres like product reviews (Skalicky, 2013), LinkedIn profiles (Bremner & Phung, 2015) and crowdfunding discourse (Liu & Deng, 2016). It is still most commonly applied to monologic texts, but its focus on the functions or actions performed by language means it also serves as a useful method for exploring how conversational goals are pursued dialogically. For example, Chiang & Grant (2017) observed that a commonly identified move in online child grooming conversations is rapport-building, which may be achieved through strategies such as giving compliments and eliciting statements of trust. The goal-oriented nature of move analysis makes it especially useful in forensic contexts where the communicative goals in question are often highly specific and unlawful.

Following methods used in Chiang & Grant (2017; 2018), our approach to developing a system for move analysis in this domain involved a manual turn-by-turn interpretation of the most likely function(s) of each user utterance. Identified functions were then grouped into higher-level moves and lower-level strategies. The data-driven coding system was developed iteratively as chat-logs were coded and re-coded, functions grouped and re-grouped, and new data continually introduced. The main advantage to this intuitive approach is that our analysis is based on careful linguistic judgement and expertise, which has allowed us to develop a system which we believe captures communicative functions in detail while remaining reasonably simple to use, and to train others in using. A downside, however, is that this sort of manual analysis is time-consuming (which is why our dataset was limited to 200 transcript pages). In the interest of simplicity we also decided to focus on what we perceived to be the most pertinent function of each utterance (while recognising that there may be additional functions). Additionally, we split individual turns where a sequence of two or more functions were clearly being realised in a single turn. We also included a category for marking 'unclear' moves, which was used in cases where an utterance conceivably had more than one possible interpretation, but it was unclear which was more prominent, or where an utterance had no clear interpretable function at all. This approach resulted in a system of 20 moves, including moves such as *Greetings*, *Offering IIOC* and *Image Discussion* (for the full system see section 3.1 below).

Following initial coding of the full dataset by the first author, we conducted an inter-rater reliability test on a small subset of the data (543 lines or 2.5%) which showed two coders (Chiang and Grieve) reached 84% agreement regarding the primary functions of each utterance. This is well within the satisfactory range of agreement (Stemler, 2004), but we believe it is especially good here, given the relative complexity of our system and the complexity

of this domain in terms of numbers of active users at any point, and considering the lack of information denoting interactional structure as is characteristic of online fora (e.g. Reddit), as opposed to chatrooms. Where discrepancies arose, these were typically around distinguishing between expressions of rapport-building, image appreciation and user appreciation, as well as how minimal responses should be coded. These issues are addressed further in the results section, though it is worth noting that some amount of overlap between moves, especially in dialogue, is to be expected (see Chiang & Grant, 2017; 2018).

2.3   Comparison of Move Sequences for Different User Types

Based on this linguistically coded data, we compared the move sequences that characterise the communicative repertoires of users, both individually and in the aggregate. To this end, we extracted all the turns from each user in each of the samples, and then represented each of these user instances (users may occur in multiple samples) as a string of moves. For example, a user might be represented as a sequence of three moves, ignoring intervening moves: *Greeting*, *Offering IIOC*, *Sign Off*.

To model user move sequences, for all 3,881 user instances consisting of at least two moves, we have focused on computing transition probabilities – the likelihood of one move being followed by another in one user's move sequence (sequences of a single move would not provide any transition probabilities). For example, if one move is always followed by a second move across all users' move sequence, then the transition probability between those two moves is 1, whereas if one move is followed by a second move half the time and a third move half the time, then the transition probabilities between the first move and the second and the third move is .5. Computing the transition probabilities between all pairs of moves across one or more user instances yields a probability transition matrix for that sample of move sequences. Note that we do not consider intervening turns from other users in this analysis, primarily because it is not simple to reconstruct who is talking to whom in these chatrooms, which often involve ten or more users engaged in multiple conversations taking place simultaneously.

Our basic assumption is that such a probability transition matrix provides valuable information about the rhetorical strategies of the user or users they represent. Furthermore, we assume that by comparing the transition matrices for different users or different sets of users (e.g. image offerers vs non-image offerers), we can better understand different rhetorical strategies of user 'types' in this online community.

In particular, we have focused on visualising these matrices by generating Markov chains (Spedicato et al., 2019, Csárdi, 2019), which are common stochastic models that describe a sequence of events where the likelihood of any given event occurring depends solely on the previous event in the sequence (Gagniuc, 2017). In a Markov chain, each state — in this case a move — is represented by a vertex, and non-zero transitions between moves are represented by an arrow extending from the first vertex to the second vertex. By representing transition matrices as Markov chains, we are able to better understand the complex discourse patterns in the language of the different user types, facilitating comparisons between user types and how they tend to navigate rhetorically through this chatroom. For simplicity, across all the Markov chain visualisations we have trimmed any transitions less than .1, to facilitate interpretation by allowing us to focus on more common transitions. Notably, our visualisations do not depict the strength of transition probability, e.g. a .2 transition and a .5 transition are both indicated with a line (see figures 1-5).

3   Results and Discussion

3.1 The Move System

We identified a total of 20 moves (plus the 'unclear' category) across the 200 chat-log pages. Table 2, which is organised thematically in roughly the order in which the moves are discussed in the following section, summarises these moves and presents basic information on their frequencies across the dataset (i.e. percentage of conversational turns in which they are observed). Themes or 'Move types' arose by observing those moves which seemed to naturally group together to describe a more general function. We recognise that other groupings may be interpreted; ours are presented merely to aid the summary of identified moves rather than being central to our analysis. We also briefly discuss a few very rare moves of potential interest to law-enforcement. A table illustrating the full move system in order of frequency, including strategies and examples can be found in the appendix. Textual examples (often generic approximations) are provided where possible.

Move type	Move	Abbreviated name	Frequency
Rapport	Greeting	Greet	15.84%
	General Rapport	Rapport	8.81%
	Sign off	Leave	0.86%
IIOC exchange	Offering IIOC	Offer	30.19%
	Requesting IIOC	Request	3.40%
Image Discussion & Appreciation	Image Discussion	Img Dis	7.44%
	Image Appreciation	Img App	12.89%
	User Appreciation	User App	1.58%
Assistance	Seeking Assistance	Seek Help	4.04%
	Providing Assistance	Give Help	2.92%
	Denying Assistance	Deny Help	0.06%
Other Less Common	Expressing Opinion/Preference	Opinion	2.21%
	Describing Experience	Experience	2.18%
	Exerting Authority	Authority	2.43%
	Showing Deference	Defer	0.26%
	Judging Character	Judge	0.38%
	Seeking User/Interaction Type	Seek User	1.94%
	LEA Accusation	Accuse	0.03%
	Meeting Planning	Meet	0.04%
	Law Discussion	Legal	0.16%
Unclear			2.33%

Table 2. Rhetorical moves of chatroom users.

A general point about the move system worth noting is that (as seems common in dialogue-based move analyses, see Chiang & Grant, 2017; 2018), there is a certain amount of strategy overlap across the moves. This is because a single utterance may perform different functions in different contexts. A simple "yes", for example, could function to express agreement with another user, in which case it would be seen as a strategy of *General Rapport*, whereas if used in response to an inquiry about, say, efficient download methods, it would function as a strategy of *Providing Assistance*. Furthermore, one of the biggest difficulties of applying move analysis to dialogue — and particularly multi-user dialogue — is that it is not always clear how utterances are functioning from the conversational context. This is especially problematic in relation to minimal positive responses like 'cool', 'nice', 'sweet', and 'thanks', which occur at a relatively high frequency and which are largely used without specification of what they might refer to (e.g. an image, comment, story, piece of advice, etc.). In some cases, the distinction lies in the force of the utterance, for example warnings or instructions issued about community rules could function either as *Providing Assistance* or *Exerting Authority*. The coder must interpret the most likely function of the utterance based on whatever contextual information is available. It is these sorts of difficulties that make the 'unclear' category necessary.

3.1.1 Rapport moves

Greeting moves are used to indicate a user's presence in the chatroom, initiate interaction with other users, and respond to other greetings. These mostly involve typical greeting terms, such as 'hey', 'hi' and 'hello', and users may or may not address the group as a whole, e.g. 'hello all', 'hi room'. This move type accounts for 15.84% of all utterances and is the second most common move observed across the dataset.

General Rapport moves function to facilitate the building and maintenance of social relationships and the smooth/cooperative exchange of IIOC, and involves strategies like well-wishing, e.g. 'hope you're all well', politeness terms, e.g. 'sorry, i haven't got those pics' (emphasis added), and positive minimal responses, e.g. 'cool', 'lol', 'ok haha'. This move type accounts for 8.81% of utterances, making it the fourth most common move across the dataset. We acknowledge that various other identified moves (e.g. *Providing Assistance*, *Offering IIOC*) also likely work towards the development of rapport, so this move accounts for the more general conversational strategies involved in rapport-building.

Taken together, these two moves account for nearly a quarter of all user utterances and seem in themselves unremarkable in the online chat context. But in working to initiate, facilitate and sustain continued interaction between users, they provide an important background context against which the more specific goals associated with IIOC exchange are pursued, making them highly important aspects of this communicative context.

Finally, **Sign Off** moves function in a similar way to these other two rapport moves and, like *Greetings*, tend to be realised by typical sign off terms (e.g. 'bye', 'later all'). However, they are far less common, occurring in only 0.86% of utterances, indicating that it is more important to announce your presence in this chatroom than to signal your departure. This likely reflects a formal rule of user behaviour in this particular chatroom; some conversations in our data indicate that regular verbal contributions are required in order to sustain chatroom access.

3.1.2 IIOC Exchange

Offering IIOC is the move by which users either provide a link to indecent material (most commonly a still image or gif), or (far less frequently) make a verbal offer to provide such a link, e.g. 'wanna see something?'. Offers may be to the group in general or a specified user. Accounting for nearly a third of all utterances (30.19%), this is the most common move observed, situating *Offering IIOC* as the central activity and primary function of the chatroom. While this overriding purpose may have been clear from the outset, the sheer volume of image offers we observed — 6,550 in just the relatively small proportion of dialogue we analysed — demonstrates how serious a problem chatrooms of this nature pose.

Requesting IIOC moves function to directly request images of a specific type, e.g. 'any *subject name*?' or indirectly express general interest in receiving images, e.g. 'who here likes to share?'. IIOC requests are far less frequent than offers, accounting for only 3.4% of utterances, suggesting that *Offering IIOC* may be a more effective approach to obtaining further material than merely requesting it. The relatively low frequency of this move suggests it is not an effective strategy for the acquisition of IIOC or for general social cohesion, and this is reflected in users' responses to this move which frequently involve some form of reprimand.

3.1.3 Image Discussion and Appreciation

Image Discussion moves function to describe and discuss the images being exchanged in the immediate interaction. The focus of this move is on more objective details than positive evaluations. Prominent strategies include describing the content or subject of images, e.g. 'red head', and providing meta-information regarding image posting, e.g. 'last one for today *link*'. This move accounts for 7.44% of utterances, making it the fifth most common move.

Image Appreciation moves account for those utterances which work to express a sense of appreciation or gratitude specifically in relation to posted imagery. Prominent strategies include complimenting images or victims, e.g. '*subject name* is perfect', positive evaluations of images or victims, e.g. 'hot', 'nice', 'cute' and expressing sexual pleasure or physical arousal derived from image consumption, e.g. 'mmmm'. This move is the third most common, accounting for 12.89% of all user utterances. The high frequency of this move and of *Image Discussion* reflects the strongly social nature of this community, i.e. images are not exchanged in a straightforward, business-like fashion, rather, they are typically consumed and enjoyed together as a social practice.

User Appreciation moves are similar but refer only to users themselves rather than imagery. They are far rarer than *Image Appreciation* moves, accounting for 1.58% of utterances, suggesting that these interactions tend to focus more on the images in question than individual users.

These three move types often proved difficult to tease out from each other, with many utterances overlapping. While they could have been grouped into a single more general 'Discussion' move, overall, it was deemed important to be able to demonstrate the more specific functions of each move where they were clearly observable.

3.1.4   Assistance

Seeking Assistance moves are used to obtain help, support or advice from other users in a range of areas. Common strategies include inquiring about the content or subject of an image, e.g. 'anyone know this girl's name?', how to locate and access imagery, e.g. 'can anyone help me find vids?', and about technical details of imagery, e.g. 'anyone happen to know what set # this is?'. Other common strategies involve inquiring about the safety and security of the

chatroom and similar online environments, e.g. 'is *site name* safe?', and inquiring about common practices and rules to which users are expected to adhere, e.g. 'can I make a request here?'. This move accounts for 4.04% of utterances, making it the sixth most common.

Providing Assistance moves are not as quite frequent, accounting for 2.92% of utterances. Common strategies of this move include addressing inquiries about image content, chatroom rules and security, warning or advising about potentially risky behaviours, e.g. 'no names', and suggesting image access locations and methods, e.g. 'click the links button'. The less frequent use of this move compared with *Seeking Assistance* is likely because requests for assistance are often met with other more prominent moves, e.g. *Offering IIOC* and may involve *Providing Assistance* but as more of a secondary move.

Denying Assistance moves are far less common than either of the two previous moves, accounting for only 0.06% of utterances in our dataset. This suggests that in general, users are motivated to assist each other in participating in the chatroom and accessing imagery. It would be interesting, however, to isolate from a larger dataset instances of *Denying Assistance*, to explore the methods of requesting help that might provoke this move.

3.1.5   Other Less Common Moves

A range of other observed moves (in descending order of frequency) include *Exerting Authority*, *Expressing Opinion/Preference*, *Describing Experience*, *Judging Character*, *Showing Deference*, *Law Discussion*, *Meeting Planning* and *LEA Accusation*.

Expressing Opinion/Preference moves are used to express opinions, preferences and stances, and debate issues which are not directly related to the immediate images being shared. It accounts for 2.21% of utterances. **Describing experience**, which accounts for 2.18% of utterances, is similar in that it gives some sort of information about the user (whether this information is true or fabricated) but wider reaching, functioning in a number of ways. First, it is used to share previous, current and planned sexual/abusive experiences and events (again, whether real or imagined). Second, it serves to demonstrate levels of (in)experience, (in)competence or (in)expertise regarding CSA offending, chatroom use and IIOC image exchange practices. Finally, it works to share personal information regarding users' lives, activities, technological practices and aspects of identity. Both moves work to tell us something about the users, and so combined, they may be the most useful in terms of building user profiles. Having said this, together they account for only 4.39% of the utterances observed, showing that users are not particularly forthcoming with regards to personal information, which is expected given the high level of risk, and low level of trust in this environment.

Some moves give potential clues about users' roles and statuses within the chatroom community. **Exerting Authority** moves (2.43%), for example, are principally about challenging other users or being generally uncooperative. This may be done at a group level, by issuing orders about chatroom practices and behaviour, or by challenging the behaviour of an individual user. Less common is **Showing Deference** (0.26%), a move approximately opposed to **Exerting Authority**. Its main functions include deferring to or complying with other users' commands, judgements and instructions, and showing extreme admiration or respect towards others. It is easy to see how the former might be associated with users of greater experience in both CSA offending in general and chatroom use/IIOC exchange, and the latter with less experienced, newer chatroom users. **Judging Character** (0.38%) is also related to these moves, as its main function is to gauge information about other users (which may or may not indicate relative experience or status), particularly regarding interests, experiences, histories, behaviours and tendencies. Characterised by inquiries, this move often functions reciprocally

with **Describing Experience**. Another related move is **Seeking User/Interaction Type** (1.94%), by which users seek out and initiate interaction with a user of specified characteristics, experiences or interests.

Certain moves identified, even very rare ones, may be of particular interest in the investigative context. **LEA Accusation** is perhaps the most important in terms of undercover tasks; clearly having some understanding of what triggers users to accuse others of having involvement with law-enforcement authorities would be extremely useful for UCs assuming offender identities online. Unfortunately, its scarcity in our dataset (0.03%) means we cannot investigate this in any detail, but its presence is nonetheless interesting. In future work we aim to target these accusations as a point of interest across both chatroom and forum data in order to explore accusation triggers explicitly. Other rare moves of interest are **Meeting Planning** (0.04%), which might provide an indicator of those individuals making genuine plans to co-offend in offline abusive events, and **Law Discussion** (0.16%), which can provide insights into how law-enforcement goals and methods are perceived by suspected CSA offenders online.

3.1.6   Move Analysis Outcomes

The majority of moves identified (15/20) each make up only a small percentage of the total content of the interactions (between 0.03% and 4.04%), although some of them certainly seem worthy of closer examination in the investigative context. The five most common moves (*Offering IIOC*, *Greetings*, *Image Appreciation*, *General Rapport* and *Image Discussion*) account for over 75% of all conversational contributions, and most clearly characterise these interactions in general. It seems fair to assume, then, that the successful participation in this particular community would likely involve at least some of these moves. This has positive and negative implications for covert operations. Arguably it is beneficial that the majority of interactional work undertaking by these chatroom users involves just a small set of somewhat straightforward linguistic behaviours; users will enter the chatroom, greet other users, possibly share some indecent material, comment on images shared by others, and use general rapport strategies to ensure they maintain a positive relationship with the community. So for undercover agents, perhaps with limited time to prepare for such an assignment, it is probably possible to 'pass' as a genuine chatroom user by focusing just on some subset of these five high-frequency moves, without having to expend lots of time learning how and when it might be appropriate to use the rarer, more nuanced moves. Of course, the biggest difficulty is that the single most common move (*Offering IIOC*) is legally unavailable to UCs. It is possible then that UCs would have to make up for this in other ways; perhaps by increasing their use of moves which demonstrate rapport-building, gratitude, and interest in others' posted content, so that they can still provide some kind of valuable contribution to the community.

3.2. Comparison of Rhetorical Strategies

3.2.1   All Users

We first explored typical move transitions of the user group as a whole, before looking to describe different user types. This provided a basis for more focused comparisons to understand how various defined user types differ from the general communicative patterns we have discovered in this domain. It also allowed us to compare how the communications of different user types vary from this general model. Figure 1 illustrates common move transitions of all 3,881 users across the 200 chat-log samples (labels used in the figures are abbreviated descriptive approximations of the full move names).
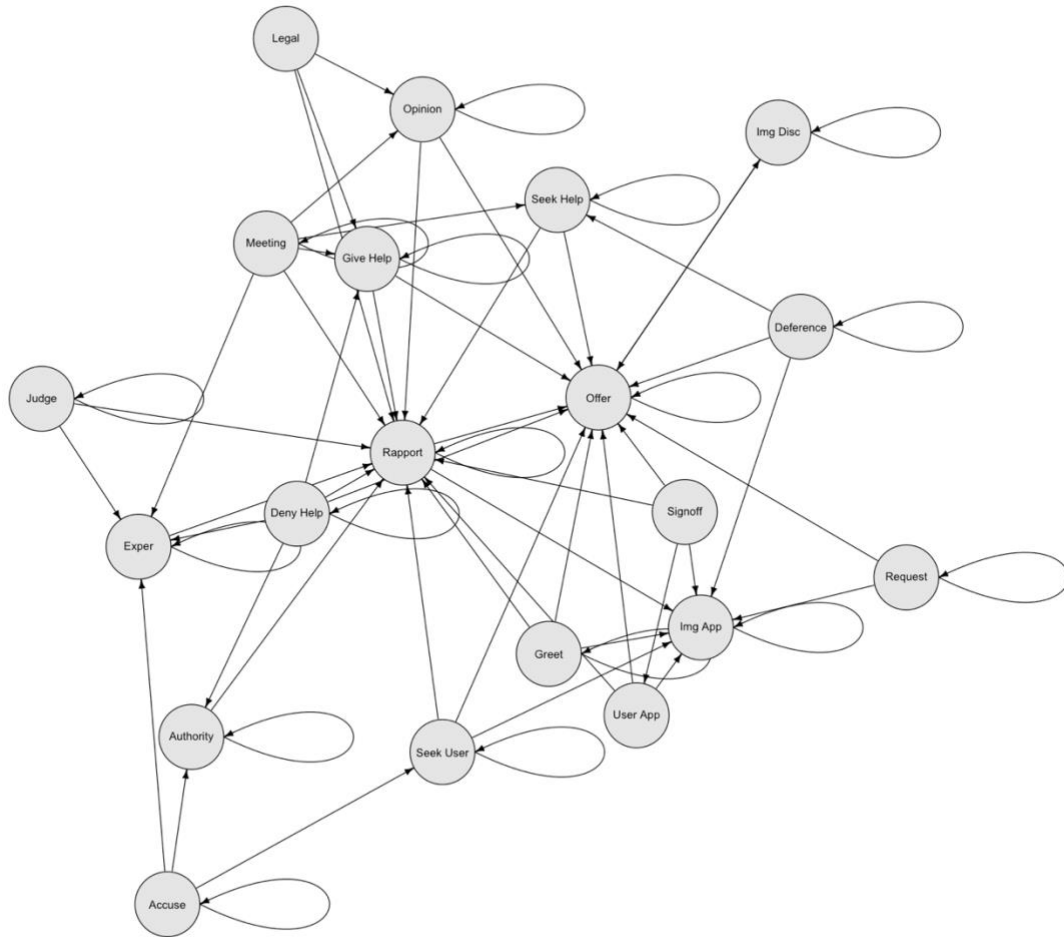
Figure 1. Common move transitions of all users (transition under .10 trimmed).

This Markov chain illustrates common move transitions for all user instances (3,881 in total). For example, users commonly follow *Image requests* with the moves *Image Appreciation*, *Offering IIOC* or a further image request, but not with, say, *Image Discussion*. What is immediately clear from this visualisation that while some moves might typically follow a small number of moves, others tend to follow a very wide range of moves, forming centralised moves, which we refer to as *attractors*. The strongest examples here are *Image Appreciation*, *Offering IIOC* and *General Rapport,* all of which tend to follow a wide range of different moves. To some extent these moves act as attractors for other moves because of their frequency: these are three of the top four most common moves across in the dataset. However, *Greeting* is the second most common move and does not act as an attractor, and the same is true of *Image Discussion*, which is the fifth most common move. These results demonstrate that these three attractor moves are employed with more freedom than others as they can follow many other moves, whereas, *Image Discussion*, for example, tends to only follow an *Offering IIOC* or itself.

A second notable point is that some moves commonly repeat themselves, i.e. users may employ the same move over two or more consecutive utterances. In the visualisations these are represented as loops (which we refer to as 'move loops'). From figure 1 we can see that loops occur with all moves except *Sign Off*, *User Appreciation*, and *Law Discussion*. It makes sense

that *Sign Offs* would not typically be followed by further *Sign Offs* but there is no inherent reason for the lack of looping with the other two moves.

Overall, the Markov chain visualisation shows that *Offering IIOC* and *General Rapport* are the most fundamental rhetorical goals for users of this chatroom, with *Image Appreciation* also functioning as an important auxiliary move. This supports the idea that in this particular chatroom, it is the norm to approach the goal of sharing and consuming IIOC in a friendly, discursive manner.

3.2.2 Image offerers and non-image offerers

Because it is of investigative value to understand which users are sharing IIOC, and because undercover agents working in the UK and many other jurisdictions are not legally permitted to share IIOC themselves, we are particularly interested in looking at users who offer imagery compared with those who do not. For undercover agents it seems especially useful to observe how non-image offerers tend to participate in this community. Searching just for user instances consisting of at least two moves where at least one of these moves involves *Offering IIOC* identified 1,313 user instances in total across the dataset. Associated move transitions for these users are illustrated in figure 2.
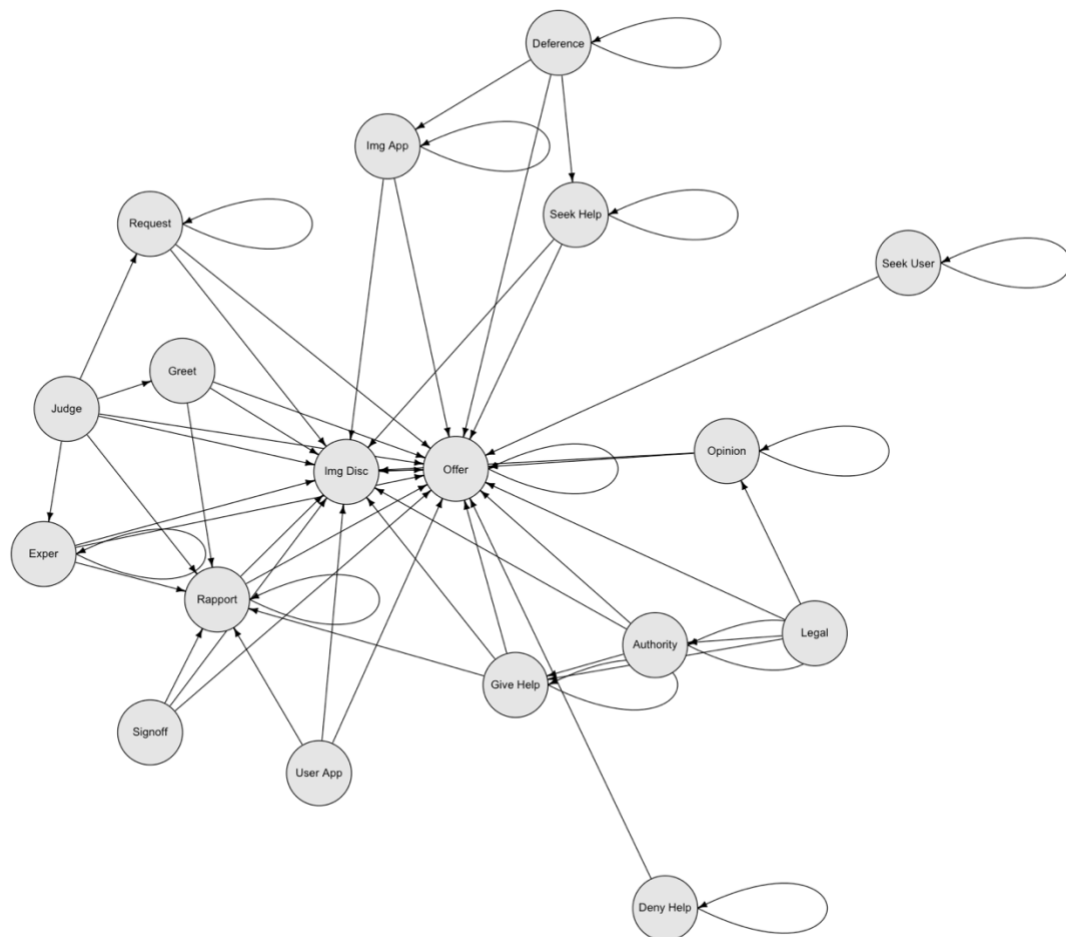

Figure 2. Move transitions of suspected offenders who used the *Offering IIOC* move.

Again, one of the first things we notice are the main attractors which tend to follow a wide range of different moves. Because we are searching only for user instances of *Offering IIOC* here, it is not surprising that this move follows the most others. Nevertheless, the centrality of this move is still notable, especially because only two moves in our entire system do not regularly precede an offer, *Meeting Planning* and *LEA Accusation*, which are the two least frequent moves overall. They are not included in figure 2 because they do not reach the transition probability threshold of 0.1 with any other moves in these user samples. Every other move in our system regularly occurs before picture offers (i.e. they tend to be followed by image offers at least 10% of the time) across these 1,313 user instances.

The main interpretation of this result is that *Offering IIOC* can be used extremely freely - essentially after any other move observed in our data. This makes sense as we know that obtaining IIOC is the main goal for all users, and so the *Offering IIOC* move is unlikely to draw complaints wherever it may feature in a user's string of moves. It is also impossible to predict on the basis of a user's previous string.

In addition to *Offering IIOC,* other common attractor moves for these user instances include *Image Discussion* and *General Rapport*. This indicates that behaviours associated with offering and discussing images and building rapport are particularly prominent for those who share imagery and that the deployment of these moves is relatively unconstrained. The bidirectional transitions between *Offering IIOC* and *Image Discussion* are also notable as this demonstrates that these users frequently discuss their own uploaded images, not just those of other users. This usually occurs in the form of some commentary on the image in question, either immediately before or after it is posted. Like the group as a whole, image offerers also engage in move loops of various kinds. Moves which tend not to appear in users' consecutive turns include *Law Discussion*, *Providing Assistance*, *Judging Character*, *Greeting*, *Sign Off*, *User Appreciation* and *Image Discussion*. Surprisingly, image offerers were also seen to use the *Requesting IIOC* move, and even loops of this move. It is possible that those who offer IIOC are socially sanctioned to request material to a greater degree than those who do not.

Perhaps surprisingly, we observed a higher number of user instances in which no image offers occur (2,568), although this is because we are looking at only short strings in this instance (i.e. two or more consecutive user turns); had we looked at only users with longer turns (e.g. at least 10), then non-image providers are very much in the minority. This shows that users are typically expected in this chatroom to contribute IIOC as well as consume it.
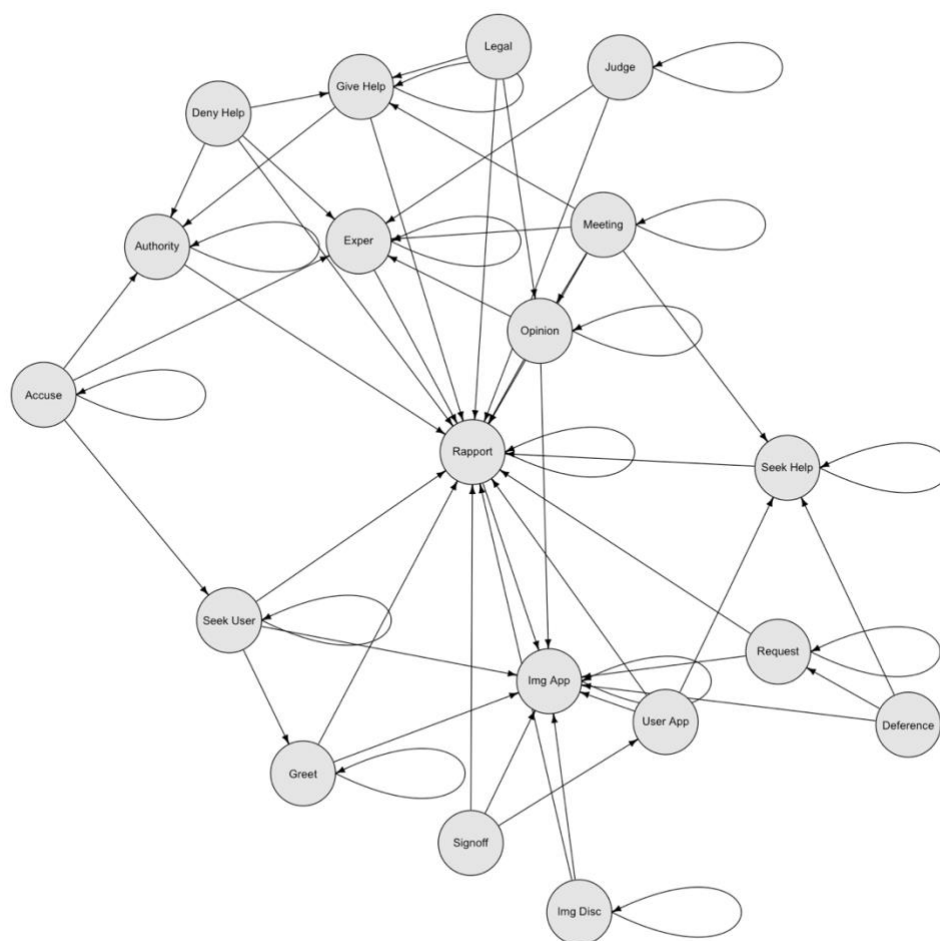
Figure 3. Move transitions of suspected offenders who did not use the *Offering IIOC* move.

Alternatively, the rhetorical structure for non-image offerers, as visualised by the Markov chain presented in figure 3, is notably different to the rhetorical structure we observed for image offerers. Not only is there necessarily a lack of the offer move, which was central in the previous model, but the most significant attractor moves to surface for these users are *General Rapport* and *Image Appreciation*, and to a lesser extent, *Describing Experience*. For a group who essentially take on the role of image consumer rather than image offerer, these moves and transitions are understandable, but it is important to confirm that these patterns are supported by the data. In particular, it points to the importance of commenting on images, expressing appreciation and building general rapport in this community if one does not engage directly in picture sharing.

From a practical standpoint, this is important information that can inform the behaviour of UCs who cannot share indecent images online. For example, it appears that these users are characterised by the prominent use of moves depicting rapport-building, expressions of gratitude and the sharing of experience, which facilitates their successful participation in this community. The frequent use of experience moves appears to be especially notable: if users do not share images, it appears that they must at least share experiences in order to demonstrate their engagement with these illegal activities, presumably in part to establish their credibility with community members. This type of information could certainly inform UCs' performance as

genuine CSA offenders online. It is also reassuring from an investigative perspective that such a large number of users seem to participate in this community without sharing imagery, although as we noted above, unfortunately this number drops dramatically if we focus on more active users.

Additionally, it is notable that these users also make requests for imagery, and that these are typically followed by further image requests, *Image Appreciation* or *General Rapport*. Crucially, based on the coded transcripts, we know that direct requests for imagery are often met with the *Exerting Authority* move, typically in the form of reprimands and reminders that the purpose of the chatroom is not to request images. It is easy to imagine that users who request imagery, particularly without engaging in the more socially positive behaviours related to rapport and appreciation, might not be so successful in their engagements with this community.

Overall, the most important difference between the two groups seems to exist the *Image Appreciation* and *Describing Experience* moves, both of which feature centrally for non-offerers, and only minimally for offerers. An interesting similarity is that both *Requesting IIOC* and *Exerting Authority* are used by both groups and appear to have roughly equal status (i.e. neither is more 'central' to either group), whereas it would be reasonable to expect these moves to discriminate between offerers and non-offerers in some way.

As well as providing information on the rhetorical structures around image offering, our dataset also provides an interesting opportunity to gather statistical information on offending behaviours of potential investigative value. According to Section 160 of the Criminal Justice Act (Home Office, 2003) it is against UK law to possess, take, make (including downloads), distribute or share IIOC (Home Office, n.d). From our coded data, we found that out of 3,881 user instances consisting of at least 2 turns, 1,313 (34%) used the *Offering IIOC* move and 1,421 (37%) use the *Image Appreciation* move, which shows users' positive reactions to the IIOC being shared, suggesting these users have clicked the offered links and thus become consumers of the material. Overall, 2,510 of the user instances in our dataset have used at least one of these two moves, providing evidence that a majority of these users (65%) have plausibly engaged in illegal behaviours. Furthermore, these numbers rise drastically if we focus on user instances characterised by longer turn sequences. For example, if we look at the 377 user instances with at least 10 turns, then 314 (83%) provide pictures, suggesting that over time, most highly active users will eventually engage in illegal activity online. Of course, these are just indicators of user offense rates rather than robust evidence, but they nevertheless provide a useful starting point for understanding the rates of criminal activity that are facilitated by these sorts of online communities and environments.

3.2.3   Authority figures and non-authority figures

Another important consideration is the comparative move sequences of users who employ the *Exerting Authority* move (which is used to challenge another user in some way), and those who do not. We informally refer to users of the move as 'authority figures' but recognise that its mere use does not necessarily indicate that a user holds the status of authority figure within the group. By isolating those who use this move, however, we are able to capture the move transitions of a group of users who are at least more likely to occupy this position. Authority figures are interesting from an investigative perspective and a useful target group for UC identity assumption because they likely have a greater influence on the community and its practices than non-authority figures; a small number of these users, often referred to as 'mods' (moderators) or 'admins' (administrators) by other users in our samples, hold official gatekeeping roles and work to monitor user behaviour and generally uphold the rules that

govern the chatroom. Also, it is plausible that authority figures have acquired such status at least in part through the demonstration of offending experience, including the provision of large amounts of IIOC, making this group an investigative priority.

Figures 4 and 5 illustrate typical move transitions for authority and non-authority figures. Notably there are relatively few users who make authority moves -- only 292 user instances out of the 3,881 user instances consisting of 2 or more moves (8%), as opposed to 3,589 user instances who do not use authority moves. Overall, in addition to the presence and absence of the *Exerting Authority* move, these two sets of users are quite different from each other. In particular, the Markov chain for authority figures is quite distinctive from the other chains presented thus far, whereas the Markov chain for non-authority figures is very similar to the Markov chain for the complete set of users as presented in figure 1. This is unremarkable, as these two chains are based largely on the same user sets, as there are so few authority figures. We therefore focus primarily on interpreting the Markov chain for authority figures.
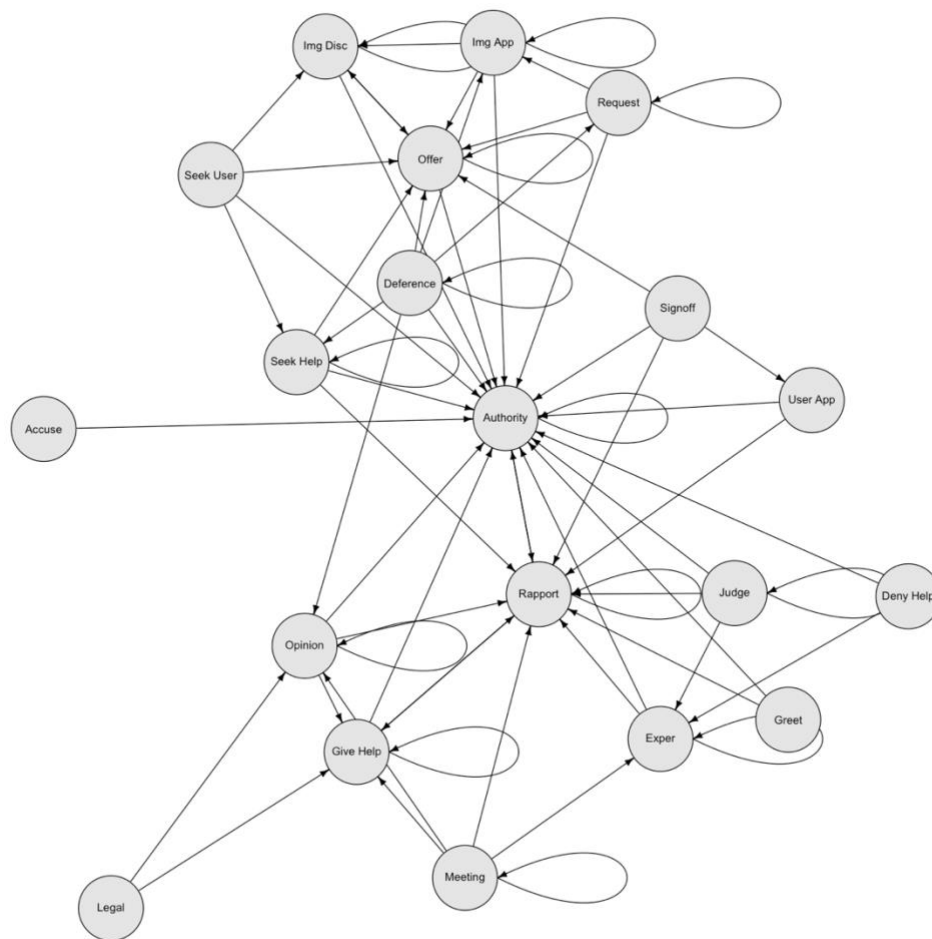

Figure 4. Move transitions of suspected offenders who used the *Exerting Authority* move.
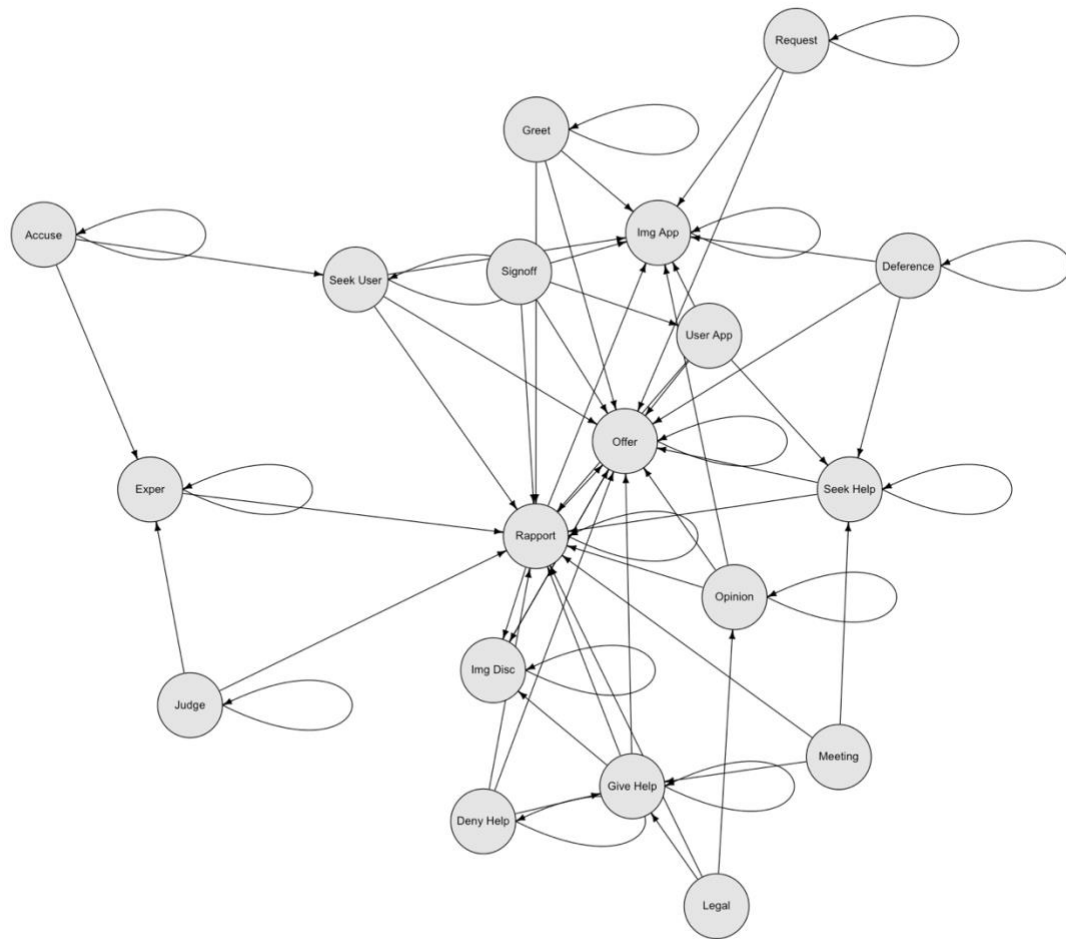
Figure 5. Move transitions of suspected offenders who did not use the *Exerting Authority* move.

In terms of similarities between the two groups, both *General Rapport* and *Offering IIOC* feature as central attractor moves, although interestingly, *Offering IIOC* tends to follow a greater range of moves for non-authority figures. This may indicate that non-authority figures need to a greater extent to establish credibility by offering images, whereas there is not so much pressure in this regard for those who adopt an authoritative position in the community. In fact, out of the 292 user instances that contain the *Exerting Authority* move, only 84 also contain an image offer (29%), whereas out of the 3,589 user instance that do not contain the *Exerting Authority* move, 1,229 contain offer moves (34%). This offers a potentially important insight for UCs: if they are unable to share pictures, an alternative way to establish an authentic offender persona may be to adopt the role of an authority figure. It seems highly likely, however, that this kind of role must be built up through extended interaction and relationship-building with and within the community, and involve the provision of IIOC, so this strategy is perhaps more useful for UCs attempting to assume the identities of specific known offenders rather than a more generic offender persona. Understanding the typical communicative strategies of authority figures can also assist in a more general investigative assessment of the community and its goals and users, especially in terms of those with the most influence who may pose the highest risk.

Another clear difference between the two groups concerns the *Image Appreciation* move, which is a far more central attractor for non-authority figures than for authority figures. It seems

reasonable to expect those more likely to occupy a lower group status to engage more in behaviours around expressing appreciation towards other users, much as we found for non-image-offerers. Again, this has important ramifications for UCs: whereas regularly expressing appreciation is an important part of acting authentically if one is to adopt the persona of a non-image-offerer, the opposite is true if one is to adopt the persona of an authority figure. Understanding these subtle differences in the different personae is potentially crucial for UCs to construct authentic offender identities in these communities.

Another less obvious difference is a stronger tendency for move sequences to lead to *Providing Assistance* in authority figures, compared with a stronger tendency for move sequences to lead to *Seeking Assistance* in non-authority figures. This result is consistent with the assumption that higher-status authority figures have more experience in using this chatroom, and are therefore better equipped to help other users, whereas non-authority figures with less experience would be more likely to need and request some form of help. This finding echoes Chiang (2018), who noted that around a quarter of her sample of 'newbie' suspected offenders used a move identified as *Seeking support* in their introductory forum posts to existing offending communities, whereas no move pertaining to the provision of support or assistance was observed.

Overall, we see clear differences between authority figures and non-authority figures in this community, which we believe offer some practical guidance for UCs attempting to adopt these different offender personas online. This approach, however, is just one method of testing the possible differences between high and low-status users, and empirical testing with data featuring users of 'known' status is needed in order to make any strong claims regarding this issue. Further work needs to be undertaken to truly understand the different linguistic behaviours of high and low status CSA offenders, but our analysis has demonstrated one way that assumptions about user behaviour might be operationalised and tested.

9. Conclusions and recommendations for law enforcement

It has been demonstrated that linguistic analysis can positively impact law-enforcement practices in the online undercover context (see Grant & MacLeod, 2016; 2018; MacLeod & Grant, 2017). In this exploratory study, we have aimed to build on this work by addressing some important questions around how CSA offenders interact with each other online. While we have only just scratched the surface, we observe a number of practical implications our findings could have for agents working undercover in online CSA-focused environments.

By describing the rhetorical strategies used by suspected offenders through a novel combination of move analysis and Markov models, we have captured a sense of the interactional goals they strive to achieve, as well as some of the ways they go about doing so. Quantifying the overall move frequencies has demonstrated that most of the interactional work taking place in this chatroom involves just a small number of moves. The sobering conclusion is that the most common and most freely employed move — *Offering IIOC* — is entirely unavailable to UCs. However, we have also seen some of the rhetorical structures employed by those users who did not offer imagery, and our descriptions of these move patterns can be used as a basis for UCs aiming to construct discourse in this environment authentically as non-image sharing offenders. Additionally, it is important to note that the remaining 19 moves are available to UCs, and while many of these will certainly prove demanding in various ways, some of the most common seem fairly straightforward to employ (i.e. those associated with greeting the group, rapport-building and expressing appreciation). We have also identified some far rarer moves such as *Meeting Planning* and *LEA Accusation*, which could be extremely important in the investigative context. At the very least, they warrant closer investigation in the future if we are to better understand both the rhetorical steps that can lead to offline meetings between

offenders, and the linguistic triggers that raise suspicion around chatroom users' potential involvement in law-enforcement.

Looking at the common moves of the chatroom users in general gives an idea of which moves tend to follow others, and, perhaps most importantly, which moves tend to feature as the centralised attractors of the interactions. From figure 1, we saw that offers of imagery, expressions of appreciation and rapport-building work are perhaps the most fundamental and freely used moves in this particular chatroom. Understanding the most integral moves and move sequences could provide UCs with a basic awareness of the nature of the chatroom, its functions, and its users' linguistic behaviours.

One of our most promising findings regarding undercover work is that there were many user instances in which no IIOC offers were made, meaning there are users of this forum who do not actively share IIOC. It would make sense, then, that UK UCs tasked with performing as CSA offenders would sit best within this group of non-image sharing offenders, and focus their efforts on emulating these particular moves and move sequences. Our findings indicate that important moves for this group include *General Rapport*, *Image Appreciation* and *Describing Experience*.

Finally, we explored a method of investigating the differences between possible authority and non-authority figures. This analysis was merely a test of our assumptions as to the likely linguistic behaviours of the two user types. Having a sense of how authority figures interact could help UCs not only to identify and prioritise the most prolific offenders, but also to learn the linguistic behaviours (besides *Offering IIOC*) that contribute to users' successful ongoing community participation, and even the maintenance of high status and respect within the community. Further to this, understanding and emulating the way newer or less experienced users participate in these interactions could be equally useful because UCs, who are tasked with performing as CSA offenders without actually sharing any of the relevant experience, may find it a simpler, less cognitively and emotionally demanding task to approach such an online community as an inexperienced but 'interested' party, rather than one with substantive experience of abusing. Better understanding of the differences between authoritative, experienced users and non-authoritative, inexperienced users could be useful for both offender prioritisation and online identity assumption in investigative police work, and in future work we plan to test the issue with users of known status, i.e. those identifying as and identified by others as chatroom moderators and administrators.

By exploring these issues, we have further demonstrated the utility of linguistic and computational methods of analysis in the online investigative context, and begun to address the current gap in our understanding of offender-offender interactions. The fact that this chatroom exists around the exchange of IIOC perhaps makes it operationally one of the most difficult online environments for UCs to participate in. We hope that our research will provide a basis for the development of training materials which can be delivered to law-enforcement agencies, particularly in the areas of online identity assumption and offender prioritisation.

Specifically, on the basis of this research we tentatively offer the following take-home messages for undercover law-enforcement agents posing online as child sexual offenders in image exchange chatrooms and fora:

- High-frequency available moves, e.g. *Greetings*, *Image Appreciation*, *General Rapport*, *Image Discussion* may be the easiest way to 'blend in' to the crowd. In particular, *Image Appreciation* and *General Rapport* are used most freely, i.e. these moves can follow most others, and are especially important for non-image offerers.

- *Requesting IIOC* is likely to be an unsuccessful move in general, but particularly for non-image offerers, likely to provoke reprimand, especially when used without more socially positive moves like *General Rapport* and *Image Appreciation*.
- *Describing Experience* may be a useful strategy for demonstrating credibility in lieu of offering imagery.
- Few users employ *Exerting Authority* moves, but authority figures may be under less pressure to provide imagery than non-authority figures.
- *Providing Assistance* was more strongly associated with authority figures, and *Seeking Assistance* was more strongly associated with non-authority figures.
- Overall, users of this forum approach the exchange of IIOC as a friendly, discursive and social practice, not as a straightforward business-like transaction.

We have presented a novel approach to examining typical rhetorical structures of CSA-focused image exchange chatroom interactions on the dark web. Our methods have enabled us to test assumptions regarding how individuals use this sort of chatroom, and better understand users' linguistic behaviours, rather than relying on assumptions alone. Building on this work in the future, we aim to conduct more in-depth analyses regarding the questions proposed in this paper, and to continue investigating dark web fora in CSA and other domains. We also aim to expand on our analytical methods, enabling us to focus on other aspects of the language, such as specific lexical items, interactional patterns, and matters of turn-taking and topic control. This wider range of methods will enable us to build a fuller picture of the linguistic behaviours of online offenders, and make more confident and more detailed recommendations for undercover agents in this and other dark web domains.

References

al-Khateeb HM, Epiphaniou G. How technology can mitigate and counteract cyber-stalking and online grooming. Computer Fraud and Security. 2016;1:14-18.

Black PJ, Wollis M, Woodworth M, Hancock JT. A linguistic analysis of grooming strategies of online child sex offenders: Implications for our understanding of predatory sexual behavior in an increasingly computer-mediated world. Child Abuse & Neglect. 2015;44:140–9.

Bremner S, Phung B. Learning From the Experts: An Analysis of Résumé Writers' Self-Presentation on LinkedIn. IEEE Transactions on Professional Communication. 2015;58(4):367–80.

Briggs P, Simon WT, Simonsen S. An Exploratory Study of Internet-Initiated Sexual Offenses and the Chat Room Sex Offender: Has the Internet Enabled a New Typology of Sex Offender? Sexual Abuse: A Journal of Research and Treatment. 2011;23(1):72–91.

Cardei C, Rebedea T. Detecting sexual predators in chats using behavioral features and imbalanced learning. Natural Language Engineering. 2017;23(4):589–616.

Chiang E, Grant T. Online grooming: moves and strategies. Language and Law/Linguagem e Diereito. 2017;4(1):103-141.

Chiang E, Grant T. Deceptive Identity Performance: Offender Moves and Multiple Identities in Online Child Abuse Conversations. Applied Linguistics. 2018;40(4):675-698.

Chiang E. Rhetorical moves and identity performance in online child sexual abuse interactions. PhD. Thesis, Aston University. 2018. Available at:https://publications.aston.ac.uk/id/eprint/39062/1/Chiang_E._2018_Redacted.pdf (Accessed Oct. 2018).

Craven S, Brown S, Gilchrist E. Current Responses to Sexual Grooming: Implication for Prevention. The Howard Journal of Criminal Justice. 2007;46(1):60–71.

Csárdi G. igraph: Network Analysis and Visualization. R package version 1.2.4. Available at: https://cran.r-project.org/web/packages/igraph/index.html. (Accessed March 2019).

Davidson J, Gottschalk P. Characteristics of the Internet for criminal child sexual abuse by online groomers. Criminal Justice Studies. 2011;24(1):23–36.

Finkelhor D. Child sexual abuse. Free Press; 1984.

Gagniuc PA. Markov chains: from theory to implementation and experimentation. Wiley Blackwell; 2017.

Grant T, Macleod N. Assuming Identities Online: Experimental Linguistics Applied to the Policing of Online Paedophile Activity. Applied Linguistics. 2016;37(1):50–70.

Grant T, Macleod N. Resources and constraints in linguistic identity performance: a theory of authorship. Language and Law/Linguagem e Direito. 2018;5(1):80-96.

Gunawan FE, Ashianti L, Candra S, Soewito B. Detecting online child grooming conversation. 2016 11th International Conference on Knowledge, Information and Creativity Support Systems (KICSS).

Kloess JA, Hamilton-Giachritsis CE, Beech AR. Offense Processes of Online Sexual Grooming and Abuse of Children Via Internet Communication Platforms. Sexual Abuse. 2017;31(1):73–96.

Liu J, Deng L. A genre analysis of web-based crowdfunding discourse. Asian ESP Journal. 2016;12(2):171-202.

MacLeod N, Grant T. "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. 2017;4(2):157-175.

Marshall WL, Barbaree HE. An Integrated Theory of the Etiology of Sexual Offending. Handbook of Sexual Assault. 1990;257–75.

Martellozzo E. Policing online child sexual abuse - the British experience. European Journal of policing Studies. 2015;3(1):32-52.

Mcmanus MA, Almond L, Cubbon B, Boulton L, Mears I. Exploring the Online Communicative Themes of Child Sex Offenders. Journal of Investigative Psychology and Offender Profiling. 2016;13(2):166–79.

Mitchell, K. J., Wolak, J., Finkelhor, D. & Jones, L. (2012). Investigators using the Internet to apprehend sex offenders: findings from the Second National Juvenile Online Victimization Study. *Police Practice and Research, 13* (3), 267-281.

Moreno AI, Swales JM. Strengthening move analysis methodology towards bridging the function-form gap. English for Specific Purposes. 2018;50:40–63.

Home Office. Indecent images of children: guidance for young people. n.d. Available at: https://www.gov.uk/government/publications/indecent-images-of-children-guidance-for-young-people/indecent-images-of-children-guidance-for-young-people (Accessed Feb 2019).

Home Office. Criminal Justice Act 2003. Available from: https://www.legislation.gov.uk/ukpga/2003/44/contents (Accessed Feb 2019).

O'Connell R. A Typology of Child Cybersexploitation and Online Grooming Practices. Cyberspace Research Unit, University of Central Lancashire, 2003. Available at: http://image.guardian.co.uk/sys-files/Society/documents/2003/07/17/Groomingreport.pdf (Accessed Oct. 2018).

Pranoto H, Gunawan FE, Soewito B. Logistic Models for Classifying Online Grooming Conversation. Procedia Computer Science. 2015;59:357–65.

Skalicky S. Was this analysis helpful? A genre analysis of the Amazon.com discourse community and its "most helpful" product reviews. Discourse, Context & Media. 2013;2(2):84–93.

Spedicato GA, Kang TS, Yalamanchi SB, Thoralf M, Yadav D, Cordón I, Jain V, Giorgino T. Markovchain: Easy Handling Discrete Time Markov Chains. R Package version 0.6.9.14. Available at: https://cran.r-project.org/web/packages/markovchain/index.html (Accessed March 2019).

Stemler, S. E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research & Evaluation, 9*, (4).

Swales JM. Aspects of article introductions. Aston ESP Research Reports No. 1, Language Studies Unit, Aston University; 1981.

Swales JM. Genre Analysis: English in academic and research settings. Cambridge University Press; 1990.

Westlake BG, Bouchard M. Liking and hyperlinking: Community detection in online child sexual exploitation networks. Social Science Research. 2016;59-36.

Appendix: table illustrating full move system

Move	Abbreviated name	Frequency	Function(s)	Common strategies	Textual examples
Offering IIOC	Offer	30.19%	To offer or provide IIOC to chatroom users (in general or specific user)	Providing links to IIOC	"*link*"
				Offering to provide IIOC links	"Want to see something?"
Greeting	Greet	15.84%	To indicate user's presence in chatroom	Generic greeting terms (and variants)	"hey", "hi", "hello"
			To initiate interaction	Greeting term + gif or emoji	"Hey all :-)"
Image Appreciation	Img App	12.89%	To simultaneously discuss imagery and express appreciation/gratitude	Complimenting images/victims	"*subject name* is perfect"
				General positive evaluations of image/victim	"Hot", "nice", "cute"
				Expressing sexual pleasure/physical arousal derived from image consumption	"mmmm"
General Rapport	Rapport	8.81%	To facilitate the building of positive social relationships and smooth/cooperative exchange of IIOC	Well-wishing	"Hope you're all well", "stay safe"
				Politeness terms	"Sorry, no links of her.."
				Positive minimal responses	"cool", "lol", "ok haha"
			To sustain continued interaction/conversation flow	Discussing topics outside of IIOC/CSA	
				Supporting sexual narrative with questions/comments	"wow" "did she react"
				Indicating/explaining temporary absences	"brb" "damn i timed out"
Image Discussion	Img Dis	7.44%	To describe or opine about images being exchanged in immediate interaction	Describing content/subject(s) of images	"red head"
				Expressing preferences for particular images/victims	"I'd have number 4"
				Providing meta-info regarding image posting	"last one for today *link*"
Seeking Assistance	Seek Help	4.04%	To obtain help, support or advice (e.g. technical	Inquiring about image content/subject	"anyone know this girl's name?"

				issues, accessing IIOC/victim, moral issues)	Inquiring about image access/location methods	"can anyone help me find vids?"
					Inquiring about safety and security of IIOC exchange	"is *site name* safe to use?"
					Reporting moral issues/struggles associated with CSA offending	"there are 2 girls making it hard to resist in my life"
Requesting IIOC	Request	3.40%	To request/express interest in receiving IIOC from chatroom users	Directly requesting IIOC (general or specific)	"any *subject name*?"
				Indirectly requesting IIOC (general or specific)	"who here likes to share?"
				Accepting offered imagery	"yes please"
Providing Assistance	Give Help	2.92%	To provide help, support or advice (e.g. technical issues, accessing IIOC/victim, moral issues)	Explaining rules and norms of chatroom use/ technical aspects of IIOC exchange	"We don't give personal info"
				Suggesting image access/location methods	"See the 'Links' button."
				Offering advice/encouragement about offending behaviour	"don't do it. Stay here and look only."
Exerting Authority	Authority	2.43%	To demonstrate authority/status over other user(s) To challenge another user's opinion, assertion, belief, practice, behaviour. To be uncooperative.	Instructing/commanding/warning user(s) about community rules and practices	"no personal info!!"
				Disagreeing/arguing with user	"some are shy and you need to respect that"
				Addressing user(s) as variant of 'newbie'	"@*username* I bet you dont know the vids"
Expressing Opinion/Preference	Opinion	2.21%	To express opinions/preferences and debate on non-image related issues To express wishes, desires and fantasies	Expressing preferences for particular victim or offender attributes/types	"9-12 perfect", "i Love pedomums"
				Expressing moral stances on issues of CSA and IIOC	" I don't exactly condone the production of CP "
				Expressing wishes, intentions or fantasies	"I'd like to experience

					something like that."
				Predicting/describing future offending experiences or lifestyles	"I'll probably end up being a pedo mom one day :/"
Describing Experience	Experie nce	2.18%	To share previous and current sexual/abusive experiences and events (real or imagined) To demonstrate (in)experience, (in)competence or (in)expertise regarding CSA offending, chatroom use and image exchange practices To share personal information regarding life and activities	Reporting/describing previous or current experiences of CSA/IIOC	"she hugs me"
				Reporting planned abuse events	"Anyone else going to enjoy a preteen today?"
				Describing current home/family life	"my daughter is 8, 9 next month"
				Reporting aspects of own identity	"im a dude btw"
Seeking User/Interact ion Type	Seek User	1.94%	To initiate interaction with user of specified characteristics, interests or experience	Seeking individual user for private chat	"looking to chat "
				Seeking user of particular ethnicity/language background	"any spanish speakers here?"
				Seeking user of particular interests/experience	"any active dads here?"
User Appreciation	User App	1.58%	To indicate appreciation of/gratitude towards individual users or whole group (directed at users rather than imagery)	Thanking users	"TY @*username*"
				Complimenting users	"you have such great taste"
				Complimenting whole group/online environment in general	"love this place"
Sign Off	Leave	0.86%	To indicate imminent departure from chatroom	Generic sign-off terms (and variants)	"bye" "good nite"
				Well-wishing	"TC SS" (take care, stay safe)
				Explaining departure	"gotta start to work soon …"
Judging Character	Judge	0.38%	To assess the interests/history/experien	Inquiring about previous online CSA experience	"first time u are here *username* lol"

				ces/tendencies of another user	Inquiring about personal information, home and family life	"you have a daughter?"
					Inquiring about opinions/preferences	"you seem to like em young"
Showing Deference	Defer	0.26%	To defer to/comply with other users'/offenders' judgment or instruction/command To express extreme admiration/respect for another user	Complying with instructions or commands	"ok thanks"
				Apologising for breaching rules	"sorry mate"
				Expressing eagerness for input from a particular user	"ohhh, waiting in excitement"
Law Discussion	Legal	0.16%	To discuss legalities around CSA/IIOC offending and law enforcement agencies/practices	Discussing personal experiences of LEA interaction	"didn't they have to say how they got you?"
				Discussing others' legal cases and posting links to relevant news articles	"tell that to this guy… fox4kc.com/2016/02/18/platte-co-man-to-spend-135-years-in-prison-for-raping-child-creating-child-porn/"
				Warning users of suspected LEA-run sites	"*site name* is LEA honeytrap"
Denying Assistance	Deny Help	0.06%	To refuse or resist offering/providing help, support or advice	Directing user to cease request	"well you need not ask here anymore thank you"
				Deliberately unhelpful/joke responses	"have you tried NetFlix.. plenty of vids there"
Meeting Planning	Meet	0.04%	To arrange/organise an offline meeting with another user	Requesting contact/location details	"PM me your infos if you want to meet up."
				Providing contact/location details	"i live in *location* by the way :)"
				Discussing details of meeting	"Should we set a date and time?"
LEA Accusation	Accuse	0.03%	To accuse another user of being associated with law-enforcement agency/operation	Direct accusations	"*Username* is a cop!"
				Inquiries/suggestions of user involvement with LEA	"so LEA ? :)"
				Declaring/suggesting particular sites/webtools are	"trap!"

				involved in entrapment methods	
Unclear		2.33%			

Table 3. Moves observed in online offender-offender interactions around the exchange of IIOC.