



Richard Chapman*

NAMING OR SHAMING? PRESENTATIONS OF THE SELF IN SPECIALISED WEBLOG DISCOURSE

1. Introduction to the study

Identity, a post-modern issue par excellence, is perhaps even more open to debate and transformation in on-line contexts, where opportunities to claim and refute personas are ever-present. This identity is invariably mediated by language, explicitly employing names or identifiers, but also hinting at much more. During research into linguistic behaviour on discussion forums, a somewhat striking observation concerned the highly pragmatic nature of naming practices. The ostensibly simple, often unpunctuated pieces of text at the head of contributions suggested much more than mere nicknames, solely carrying out the basic task of differentiating one speaker-writer from another. A whole complex of practices appeared to be present in the data: we were able to observe the self presenting itself, and adding a narrative of justification to this representation. While initial conclusions must of necessity be tentative, data suggest that naming practices on-line are highly pragmatic and full of social meaning, and so worthy of greater attention.

2. The discussion forum corpus

A corpus of around 100,000 tokens was assembled over a single day in summer 2014 by collecting contributions to a weblog discussion relating to the Israel-Palestine conflict. The discussion was part of an “as it happened” page on *The Guardian* website on 4th August 2014 entitled, “Gaza crisis: Israel and Hamas agree to 72-hour ceasefire – as it happened.” The aims of the study were, firstly, to observe linguistic behaviour in English on an electronic forum over a very limited time period, involving a highly specific and controversial topic of discussion, and, secondly, to explore the methodological issues concerned with the creation of small corpora and how they should be interrogated. It was hoped to exploit the apparent ‘naturalness’ of the contributions collected in the corpus and, at the same time, benefit from the relative ease of their collection: using simple word-processing tools it is now possible to bring together significant quantities of already digitalised linguistic data quickly and with a reasonable degree of accuracy. Data were to be analysed both quantitatively (more usual approach in modern corpus linguistics) and qualitatively, but the intention from the outset was to take a hands-on approach to any data, attempting to contextualise them in a more thorough way than is perhaps common procedure with corpora. An assumption of the research was that increased emphasis on co-textual and contextual information would significantly enlighten the interrogation of corpus data and might suggest a valuable addition to the array of linguistic methods of analysis currently at our disposal in this approach to the investigation of language.

Both collection and analysis of data proved to be more complex and challenging than first envisaged.¹ Nevertheless, data revealed many interesting features deserving higher-level linguistic analysis (e.g. highly pragmatic uses of discourse markers, wide-ranging references, register variations and intertextual items), and some striking peculiarities emerged even on preliminary reading. The most immediate and suggestive element to be observed was the employment of names, both for apparently onomastic purposes, – that is, the “meaning of a denotative name as a name” (Nicolaisen, quoted in Alia 2009, 8) – and as pragmatic, discourse-level, instruments of expression. A re-reading of the data was thus undertaken, with the view to attempting to establish the general characteristics of naming behaviour on the weblog, and perhaps to arrive at a predictive

* Richard Chapman is Researcher and Lecturer in English Language in the Department of Humanities of the University of Ferrara, Italy. A first degree in history from Cambridge gives his linguistic research a cultural and anthropological feel, while extensive experience in teacher training results in a pragmatic approach to language. Publications include coursebooks for English language learners (most recently, *Flash on English Advanced*, ELI 2017) and studies in developments in language from sociolinguistic, textual and pragmatic points of view. Recent publications include work on computer-assisted language testing and on the pragmatics of language tests. Current interests are corpus linguistics and English as a lingua franca.

¹ See Chapman, “Corpus Linguistic Analysis: How far can we go?” (2018)



theoretical understanding of how people choose and use their linguistic signs of digital identity in discussion forums.

3. Naming oneself on a weblog

Textually, weblog names are immediately noticeable by their prominent position, detached from the individual contribution, but linked closely to it. This gives an opportunity for interplay between the main text and the author's digital handle, an ancillary text item which acts both as identifier of the contributor and, often, as a form of header for the message itself. Simple examples of this found in the data include, "peacenowandforever;" "Whereisthejustice;" "TheSaracen:" each handle giving an idea of the political position of the contributor, or her/his origins. Any comment will almost inevitably be read and interpreted in relation to its provenance as presented by the name used. Indeed, we could even go so far as to suggest that contributors' names virtually impose an immediate context for the message below them, giving, as we shall see, a strong clue as to the pragmatic intentions of the author, and thus overcoming the problem of lack of social, situational contextual elements (e.g. phonological indicators, bodily movements, hierarchical considerations and the immediate physical location of a conversation, among others) that notoriously bedevils internet communication.² All discussion forums have rules regarding the possible letters and numbers or symbols to be used in combination to make up the identifiers, but these are usually quite broad (Hagström 2012, 86), and for us the only immediately significant thing to remember is the obligation to use a unique name: the onomastic function of these names as identifying linguistic forms is thus clear. This might be in contrast to everyday life, where names are often far from unique, at least in their simplest or most used forms.

Of course, the obligation to uniqueness is an opportunity for invention, creativity and self-advertisement, and so observing the names created may shed light on the participants and their origins, and the effects these identifiers might have, or be intended to have, on other users. In contrast to face-to-face linguistic interaction, these names operate largely without any other personal or social markers. If we use a person's name in a conversation, or present ourselves in public, we know there are various other sources of information open to the recipient: our facial expressions and appearance, various potential signs of social status, accent and other indicators of affiliation, and even clothing. Thus names often act in a supportive (or conflicting) non-verbal environment, presumably making their communicative task easier (Hagström 2012, 83). Instead, the digital names we observed were effectively operating on their own, carrying what information they could into the echo chamber of on-line discussion with little other than their written form and potential semantic and intertextual elements to aid them in the task. The only obvious exception to this is the use of miniature pictures in connection with the identifier. Examples in our corpus included familiar cartoon characters, images of politicians in burlesque style (see "CameronKnows" below, section 4) and ingenious montages (such as "MeandYou" which used two faces mixed into one). These are invariably presented in a rather small format on-screen (they may be photographs, sketches or other graphic images), next to the name of the contributor, and tend not to be altered once chosen. Digital names cannot usually be modified either,³ again in contrast to more common naming practices in which, depending on contextual needs or pragmatic considerations, original forms may be shortened or titles given or discarded (e.g. "Mr Smith," "Smith," "Smithee" etc.). Intonation will, of course, be absent too, eliminating a further level of linguistic variation. It was initially hoped that this linguistic 'isolation' might allow the examination of language features and behaviour in a way that would be particularly fruitful, being shorn of several complicating elements and so enabling the researcher to focus on a limited number of lexical or grammatical characteristics.

4. A brief description of the data

The original corpus of approximately 100,000 tokens was reduced to a total of just over 87,000 after preliminary adjustments were made, eliminating items that occurred repeatedly as elements framing the presentation of each contribution (e.g. "like;" "reply;" "report to moderator" etc.). There were assessed as being distinct from

² Emoticons are the simplest and most explicit attempt to solve the problem of pragmatic context deficit online, giving an obvious indication of humour, irony or such like to replace the lack of visual or tonal cues. For a pragmatic-oriented study of the use of emoticons, among the extensive literature (both journalistic and academic), see Dresner and Herring, "Functions of the Nonverbal in CMC: Emoticons and Illocutionary Force" in *Communication Theory* Volume 20, Issue 3, 2010.

³ On the Guardian 'As it Happened' blog the rules only allow one alteration, and this only occasionally.



the actual linguistic behaviour under observation and it was decided that they would inevitably distort frequency calculations. However, it is worth noting that even before quantitative analysis had begun in earnest, some subjective judgements thus needed to be made (Dörnyei 2007, 204-6). In our case these entailed the exclusion of some terms occurring as frames, and contributions that included verbatim repeats of previous comments.

Total word count of original corpus	106,870
Total tokens (after cleaning up data)	87,000
Total contributors identified in corpus	328

Table 1. Basic information about the discussion blog corpus

In all, 328 individual contributors' names were found, of which 114 (35%) appeared with an accompanying personal image. All examples provided in this paper are taken from these 328 digital identifiers, found in the August 2014 Israel-Palestine corpus described here. Quite often, though by no means in the majority of cases, the image illustrated the semantic content of the identifier, and occasionally the image and the name interacted in a humorous or ironic way: for example, the contributor "CameronKnows" had chosen a photograph of the former British Prime Minister with an elongated, Pinocchio-style nose. It is interesting to note that an ostensibly genuine photograph of the blogger her- or himself was used as an accompanying image in less than 20 instances (and of these many may not be images of the real person at all). This lack of personal photographic representation highlights the absence of real visual or socio-contextual 'help' that the identifying language had at its disposal, and so how much the names had to work on their own.

The most obvious feature of the identifiers observed was the lack of simple, 'ordinary,' apparently real names (e.g. "Leonard Kelly;" "Jessica McCallin" etc.), perhaps due to the need for individual identification and uniqueness mentioned earlier (each digital handle must be unique as an enrolment requirement, and namesakes are, of course, quite common in everyday life). Even with the most generous interpretation of 'ordinary names' only 41 out of 328 could be classed as such. Here it must be added that assessments of what count as familiar names are highly subjective and were problematical. Names that were identified as foreign to English-speaking naming practices added a further level of difficulty. For example, "YasminElReyah" and "Yaniv Baron-Gershfield" were categorised with a certain degree of confidence, but "sajidulghafoor" left more room for doubt as to interpretation. Whatever the challenges in observation, we can perhaps assume that the lack of recourse to the typical name-stock in English (on a forum where almost all contributions were in English) underlines the idea mentioned above, that the linguistic task of a digital identifier is somewhat harder than that of a name in real life: it has to perform its functions largely on its own and has to be different from all others.

This suggestion is further supported by the prevalence of the use of 'ordinary' naming vocabulary with a simple numerical addition. On 21 occasions a name and number sufficed to grant identity, and perhaps a certain anonymity (e.g. "Bert9000;" "marky226;" "JonM267" etc.). Six contributors combined a name and location ("PhilipinKrakow") and 16 used a place name or noun or adjective and a digital reference ("Leviathan2;" "reptile0000;" "Honest1"). Another solution was to use a foreign or seemingly-exotic reference with varying degrees of familiarity for an English-speaking audience ("laguerre;" "cavalonero;" "Bauhaus"). 8 contributors merely used an ID that seemed to have been assigned by the forum software (e.g. "ID098409") and 21 were deemed incomprehensible (usually a series of letters, such as "gibtardo"), but these are clearly open to interpretation, and it is worth mentioning that many apparently indecipherable names became less opaque on repeated reading.

But the most striking feature of contributor naming behaviour was probably the use of noun phrases (47 examples), and what the researcher classed as "full clauses" (at least 17 examples). Noun phrases included identifiers such as "Aesthetic Theory;" "MeandYou;" "dreamwatcher;" "WithoutMalice" among many others, and obviously opened up interesting questions as to the real function of names on discussion forums like this. These questions were thrown even more into relief with the occasional instances of full clauses being utilised, such as "Theremustbeadiffway" or "Whereisthejustice." These were often imperfectly formed, lacking punctuation and standard spelling, for example, but this takes little, if anything, away from their status as a clause, containing a subject and predicate, and carrying a recognisable pragmatic message. When we add another 20 names that were clearly attempted witticisms (e.g. "hotmale55" or "MrParadeRain"), thus also



operating largely on a clause level, we can observe an aspect of weblog naming that was not perhaps to be expected. The primary linguistic level of reference might be the clause rather than the proper noun or simple reference. Thus we have around 80 of our 328 names (roughly 24%) operating as predicates rather than names in the most obvious sense. And indeed, this figure is highly likely to be an underestimate, as interpretation was difficult and subjective, and the policy in describing and categorising data regarding noun phrases and clauses was to be conservative, not reading each digital handle as a clause unless the researcher was convinced to assign an expression to the predicate group on repeated later readings. The combination of a familiar name and a noun (e.g. “JOHNNYHEMISPHERE” or “SidsKitchen”) was not deemed to be a legitimate noun phrase, but might well be read as such, and certainly invites possible clause-level analyses to accommodate potential ironic interpretations.

One last observation should be made, concerning punctuation. Generally names were presented as a single word, even if they were a clearly recognisable stream of common words in English (e.g. “sonofabun”). This suggests another small layer of potential variation in the use or elimination of spaces between letters, but posed few serious problems in analysis of the data. There seemed to be a strong tendency (or even a perceived obligation?) to join usually distinct words together, perhaps emphasising an essential difference in this kind of naming. Only 11 identifiers (just over 3%) included spaces where traditional orthography would expect them, out of the 328 names found. It should be underlined here, however, that this small phenomenon has potentially serious consequences for computer-based corpus linguistic analyses, as the issue of what constitutes a token, and what might be three or four tokens has suddenly become more problematic on a theoretical level.

5. Analysis and discussion of naming data

The first, and rather unavoidable observation to be made was the recognition of the difficulty of analysing the data collected. Indeed, a first finding of the research could well be that naming behaviour in discussion forums is quite varied and defies simple categorisation. As many as fifteen classifications were used during readings (e.g. proper noun; traditional name; name with numerical element; noun phrase etc. See table 2), but it was repeatedly obvious that individual names could slip from one category into another very easily. It was a highly subjective thing to assign personal identifiers to categories, as became even more clear on later readings, when the same name might also fit into the new category under examination. An inclusive policy was eventually adopted, allowing, for example, a noun phrase also to be counted as a witticism, and meaning that categories were not exclusive. In other words, the same name might be assigned to more than one category. This also means that any percentages given indicate the observed presence of a particular feature as a proportion of the whole, but this does not mean that the total of per cent scores for the different categories will add up to a hundred. Indeed, numbers are expressed as per cent scores only rarely in this paper as they could be deceptive. But it was considered of dubious value to struggle to force a name that might well operate on various linguistic levels into one class, excluding it from another.⁴ An open categorisation policy like this was also intended to be less subject to judgements depending more on personal bias than objective criteria, an error of inclusion being considered perhaps less serious than one of exclusion.

All categories	328	100%
With pictures	114	35%

⁴ It is interesting to note a potential difficulty that all corpus linguistics may be subject to, namely the issue of POS tagging and parsing (McEnery, Xiao, Tono 2006, 75-6). This is invariably done in accordance with a rather traditional linguistic understanding of parts of speech (i.e. noun; pronoun; adjective etc.) which is often perfectly acceptable, but calls into question the oft-made claim of “corpus-driven research:” if language research is truly corpus-driven then the functional divisions of a sentence should perhaps be discovered through interrogation of the corpus rather than assumed a priori. For us, the more relevant issue is precisely the potential for a plurality of function in a particular name: in classifying a digital identifier as one thing, we might be ignoring or concealing another aspect of its functionality. An example of this from our corpus is the contributor known as “SELAVY;” a name that might be merely a one-off, strange-sounding handle aiming at originality, or could be an ironic take on the French expression, “C’*é* la vie.” There are clearly various potential communicative levels requiring analysis and a one-off or categorical tagging procedure might not do them justice.



Pictures interacting with names	3	1%
Full clauses	17	5%
Noun phrases	47	14%
'Usual' names	41	12.5%
Jokes	20	6%
Exotic references	18	5.5%
Name with digital reference	21	6%
Geographical place name	16	5%
Name with joke	5	1.5%
Name with place name	6	2%
Incomprehensible	21	6%
Proper nouns	8	2%
Interjections	1	1%
Simple ID number	8	2%

Table 2. Categories of weblog identifiers found in the corpus

It is also worth reminding ourselves at this juncture of the fundamental importance to this study of the procedure of data analysis employed. It was a process of complex and contextualised examination of each instance of a particular linguistic feature, rather than one of simply counting frequencies. It is suggested that the data largely defy an exclusively quantitative approach and emphasise the value of small corpora in their ability to occasionally throw up linguistic surprises in much the same way as large corpora do, but at the same time enabling detailed and context-based analysis of them. In other words, the approach is more ethnographic, following Dörnyei's recommendations (2007, 132-3) for understanding social processes and seeing the participants' perspectives, and perhaps generating initial hypotheses.

Another methodological safeguard was the iterative and repeated readings of data. It was found that assessments were modified and refined as readings for each category were carried out and as the data were tagged (Dörnyei 2007, 126-7; 143). Additionally, identifying language could be compared to the content of the contributions they introduced, granting a further layer of textual information to compare and contextualise assessments. Lastly, the software of the forum offered an external level of analysis in the form of the contributor profiles it made available to any user. When attempting to come to a conclusion as to the linguistic category or pragmatic intention of a particular digital identifier it was possible to click on the profile of the contributor and see various other comments she or he might have made in other discussions, allowing the researcher to form a slightly more informed, and certainly textually richer, idea of the category the name might be assigned to. An example in the data is "jakboot." an examination of the contributor's profile enabled the handle to be assigned to the category of humorous naming with some confidence, as the poster had very high (probably native) linguistic competence and the likely lexical associations of the screenname showed a good 'fit' with the points of view of the writer. Interestingly, this procedure also grants the researcher the opportunity to contextualise the contributor's behaviour not merely over larger spaces of text, but also over time, adding a diachronic contextualising element to a corpus approach that is inevitably largely synchronic on most occasions.

The user profile function of the software tells the researcher something more about the contributor, or about her/his behaviour, but actually still leaves us largely in the dark. Indeed, one of the fundamental aspects of analysing online linguistic behaviour that should not be forgotten is the sheer paucity of hard information at the researcher's disposal. On-line, our gender is often not given any indication by name (e.g. "imperium3" or "Laserprinter"), or we can even use an apparently gender-specific identifier to dissimulate (we cannot know whether "CatherineLouise55" is truly a female contributor or even assume that "Christopher3175" is male. This makes the identifiers we are examining even more 'stand-alone' items of language. Indeed, rereading naming data gave the researcher pause to contemplate the special irony of internet naming: contributors are invariably seeking to present an identity, and perhaps a highly pronounced one at that, while at the very same time aiming to achieve and exploit anonymity. Contributors want recognition on the forum and secrecy to have the freedom to act (politically and linguistically) as they wish. Names on weblogs have these two mirroring functions, identifying and obscuring in a way that is in sharp contrast to linguistic uses of naming in the 'real



world,' where it is only titles, extremely common names or sobriquets, perhaps, that give the user a shield of public concealment. Naturally, this might be of particular importance if one's political views are considered extreme and controversial, or if the intention is to deceive readers as to one's genuine position on a topic. Many times it was observed that a contribution was entirely or partially dedicated to debunking, not the content, but the provenance of a particular comment. Contributor's identifiers were called out as false, or their comments were labelled as intended to be divisive or misleading: "I'm an Israeli and a Socialist, and I hate Hamas. You are British and a socialist, and you love Hamas." "...and I'm a banana." Naming a contributor as a troll was not uncommon and usually accompanied by indications of annoyance. Although the year of our data collection, 2014, was before the controversy about 'fake news' had taken off, it is interesting to note that in our data it was already considered to be a problem, with statistics questioned and accusations of falsehood common and strongly expressed. The digital identifier "TonyPantsonFire" (probably an ironic reference to the childish proverbial accusation of dishonesty in English, "Liar, liar, your pants are on fire!") suggests awareness of this issue, even to the point of ironically claiming to be a liar, either to underline the immodest or dishonest claims of absolute truth by others, or to give sarcastic prominence to her/his own claims. "Liesandstats" is another digital name that depends purely on the proverbial expression, "there are lies, damn lies and statics" to have meaning and demands unpacking by readers. For our analysis, the significance of this behaviour is perhaps in the choice of identity and its marker as a means to establish credibility: in other words, the naming going on in forums is highly pragmatic, serving the function of staking a claim to authority or, at least, authenticity. Rather than "doing things with words," our contributors are doing things with names. Names have been seen as speech acts before (Alia 2009, 10-12), and in our study this aspect of naming seems even more apparent.

The pragmatics of weblog naming became clearer and clearer during data analysis. As we have said, the corpus contained numerous examples of full clauses, and even noun phrases could often be interpreted as predicates of simple verbs: "RecantedYank" implies an elided "I am a" preceding it, and might constitute an attempt to claim both national origins and a definite geopolitical perspective. In the same way, the complete clause, "Capitalismdoeswork" might establish a neoliberal outlook, and the noun phrase, "yourmiddleclassfarce" again nails political colours to the mast unequivocally. Of course, irony cannot be excluded and the choice of a handle may not imply complete identity between the semantic content of the name and the social, political or geographical character of the real person. Complications in interpretation abound, as we can be sure neither of the integrity nor of the seriousness of these speech acts, but this is of lesser importance here: the significant aspect of the data is the observation of explicit pragmatic elements in the digital identifiers we collected. The intended meanings of names could only be established with some certainty by reading the individual contributions in the blog, and, if possible, blog activity over an extended period.

Intertextuality and reference are powerful pragmatic tools, emphasising as they can both deep cultural knowledge and subtle appreciation of nuances of social experience. They can be both intense generators of group solidarity, and effective means of identifying and excluding lames (Spolsky 1998, 31-43; Baker 1992, 190-196). Not surprisingly, intertextual elements were quite common in our data and seemed to perform clear pragmatic functions. "Crumpsalllass" makes direct reference to an area of Greater Manchester (Crumpsall), but one that is unlikely to be known by many participants; "badbagpuss" references a British children's TV show of the 1970s, marking the contributor out both for location and generation; "andyoldlabour" indicates political affiliation and awareness of subtle differences in British politics; "nnnnnineteen" is presumably a reference to a popular hit of the 1980s (Paul Hardcastle 1985⁵), which was critical of the Vietnam War. Again, detailed, almost private knowledge is on show here, and the pragmatic content is undeniable. In the data it was also noted how, in general, the sociolinguistic effects of such intertextual referencing were mostly exclusive, creating lames (Spolsky 1998, 36), rather than aiming to create a community in the way that has

⁵ Paul Hardcastle, *19*, Chrysalis Records, 1985. An innovative single using sampling of voice commentary to reflect on the Vietnam War, it surprisingly reached number one in the UK charts and in many other countries. It is perhaps interesting to observe (using the profile function offered by the weblog) that the poster using this handle seems to be highly critical of Hamas, but also of the British and US 'establishment,' and claimed on one occasion to be a UKIP voter. Clearly the choice of identifier is a complex mix of personal experience, associations and potential political messages, and is not open to simple categorisations.



been described in literature previously (Hagström 2012, 89). This may well be something to be expected in a forum that is essentially competitive, hoping to enable debate on a controversial topic, rather than a site intending simply to entertain.

The humour (another highly pragmatic feature of language) observed tended to make use of intertextual reference heavily as well. “PortaloMassacre,” complete with image of a temporary prefabricated structure, makes obvious reference to the Peterloo Massacre in St. Peter’s Field, Manchester in 1819, but depends on some knowledge of English history (it is a highly politicised memory of the brutal suppression of a political demonstration) in order to recognise the wit and divine the political import of the name. “MoboDuck” is possibly a sly reference to Herman Melville’s classic novel, and “Wombatman” reminds us of the famous cartoon character. Other cartoon references that may be quite international, but tend to be limited to one or two generations, included “Tintin’sdog” and “muttley79.” In a similar vein, oxymoron or unlikely combinations of words were used for comic effect, such as “fisherpricetuxedo,” “MysteryRabbit” or “GreatUncleEuphoria,” but again in analysis these were deemed to depend on sensitive cultural or linguistic awareness to be effective. While some humorous names were perhaps more internationally comprehensible (e.g. “KarmaGeddon” or “demorat”), the data suggest that humour almost always acts both as a strong sociolinguistic binding force for those ‘in the know’ and even more as a merciless gatekeeper for those who cannot ‘get the joke.’

Humour and linguistic deviation also have the pragmatic task of presenting the speaker: they give the contributor the opportunity to show her/his verbal dexterity and inventiveness. Examples of this have been mentioned previously (e.g. “CameronKnows” in Section 3, paragraph 2, above) and there were others: “psygone” (presumably a phonological pun on the city Saigon); “mcscotty” (an exaggerated Scottish reference); “paqman78” (referencing the old arcade game); “bAnDoLeRo” (exploiting upper and lower case letters to spell out initials), among others. This attempted humour or linguistic play indicated to the researcher another, rather surprising, aspect of the data. Names were essentially representational rather than referential in their linguistic workings. A name in society is primarily referential, referring to a precise individual (Alia 2009, 12-14). Instead, as we have seen, the context of situation is more ambiguous on internet discussion pages, since names are intended both to be recognised and often to conceal true identity, misleading the other readers, or merely giving little information away. This can be for reasons of personal safety or privacy, as well as for any attempts at deception.

Reading the data, it became apparent that the two-fold role of digital identifiers was, firstly, to act as a unique signifier, but then secondly, and more importantly, to be powerful speech acts, attempting to present or represent the contributor to the audience in the forum. A participant can, to a large extent, take on any persona s/he chooses, and so the choices themselves are linguistically significant for us as indicators of language use and pragmatic behaviour. We may observe and describe the repertoires (Bloomaert 2010, chapter 4) that are used (and so available to the self-identifying contributor) and see them as realisations of cultural experience, as well as of linguistic competence. Pennycook (2010) and Bloomaert (2010) have emphasised the limitations of these repertoires in a globalised community: they are not always interpreted correctly or easily, and people have greatly unequal access to repertoires for communicating. And we can recognise another potential limitation. Repertoires may intentionally be employed to emphasise difference and a lack of shared human experience, rather than to create bonds of solidarity. Thus, “bobsyouruncle1” depends upon idiomatic competence of a reasonably high level (Prodromou 2008); “Sodtheproles” is not just an aggressive statement of political attitudes, but also a semi-private reference to a proverbial British English expression; “Isanybodyoutthere” is perhaps more easily interpreted and appreciated by someone aware of the rock band Pink Floyd;⁶ “Meet Joe Black” depends on the film of the same name; “NorthMonkey” demands detailed cultural knowledge of the United Kingdom and “Lebesgue” (presumably referencing the French mathematician) is hardly transparent to all contributors. All of these references risk being missed or being misinterpreted, as representational language often is. But they are also indications of the essential feature of representational linguistic behaviour: it is invariably based on predicates rather than single items of lexis: “the constitution of meaning is a property of the undivided sentence.” (Ricoeur 1978, 183).

In the simplest sense, all digital identifiers, as we have seen, are clauses with the preceding “this is” or “I am” left out as an ellipsis. But more than this, the data reveal the sentence-level grammar required to interpret and analyse weblog names successfully (e.g. “guardiansystemssuck,” “bobsyouruncle1,” “riddlemestupid” etc.).

⁶ “Is There Anybody Out There?” *The Wall*, Pink Floyd, Columbia Records, 1979



Jokes, partisan declarations or judgemental propositions are all possible only at the clause or sentence level and, strangely, the repeatedly observed habit of employing non-standard punctuation to digital names only served to emphasise their non-lexical nature. Thus “curiousaltruistic” seems to tell us what the contributor is like, much more than Goffman’s “untested stereotypes” (Goffman 1969, 13) would with names in daily life, and “evenuglierthanyou” initiates a whole strand of thought and ‘feels’ like a full clause. The habitual presentation of digital names as single strings of letters may look a little peculiar, but actually the researcher found this aspect of them an aid to understanding: Bakhtin’s observation of the necessity to “establish the principal and clear-cut boundaries of the utterance” (Bakhtin 1986, 112) as the first step in understanding seems to be realised in current weblog naming and again spotlights the integrity of the message contained within a name.

6. Towards some conclusions

If our observations based on the data are correct, then it is clear that fuller linguistic analysis of weblog naming will need to be at the level of discourse. Sentence-level analysis in many ways presupposes discourse (unless we really wish to analyse the make-up of a sentence discretely), and in highly political forums such as this Critical Discourse Analysis will probably be an effective tool for deepening our understanding of the choices people make in attempting to identify themselves (or mask themselves) in weblogs. This is made even more explicit in cases where there is a clear, strong political point of view contained within the name itself (e.g. “neomarxist” or “trueblueozzy” in our data). We require a good deal of external information to begin to examine the name as a part of discourse (the discussion topic and historical factors; the stage in the discussion when the contributor makes her/his entrance; previous messages and opinions expressed etc.). Only then can we disambiguate the terms and judge whether the contributors reveal a Marxist influence in ideology, or can be recognised as a conservative Australian, as the two identifiers suggest respectively. This is all the more the case when we observe a clear (and sometimes ironic) interplay between the name and the point of view expressed in the accompanying post. Here we are in tune with Fairclough’s well-known opinion (1995/2010) that “we should not look at language but at discourse when we consider sociolinguistic globalisation processes.” In a sense, in observing naming behaviour online this is exactly what we are trying to do. ‘Diachronic contextualisation’⁷ (through the exploitation of individual profiles) greatly assists this approach, as does the identification of progressive relatedness or cross correlations over broader amounts of text.⁸ The idea that ‘it’s what you say that counts,’ and not what you are called, has been questioned in sociology and sociolinguistics (Hagström 2012, 83). However, in these data the problem hardly exists, as we can see the claim of identity as a part of the message; the contribution and contributor melding into one, both linguistically (to some extent) and in discourse terms. The overall speech act is perhaps as much a question of self-presentation as of debating a particular point. And the interrelationship and intertextual interplay between the digital name and the contribution, along with accusations aimed more at the messenger than the message, underline this.

Name-calling and attempts to unmask trolls and dishonest debaters only emphasise even more clearly how identity works both in the header and the body of each post, and how each header is also a part of the message. Examples from the data include: “Actually it is clear that Dief is a Tel Aviv agent;” “And this is your fifth comment since joining in May [...] says it all;”

‘Sunofabun’ stated in a post yesterday that he/she is an Israeli [...] which explains the ‘we.’
Contributors are identified by the opinions they claim to hold and are sometimes shamed for doing

⁷ This expression is used to indicate a method of placing each utterance in a context that includes its immediate co-text, the general context of the debate taking place and the world-historical topic in a broader sense, and, lastly, the comment seen in relation to other comments posted by the same contributor over time, as presented in her/his profile or in other blogs, or their internet or social media activity over time.

⁸ The author would like to suggest that progressive relatedness tends to be greatly underestimated or even ignored in large-scale corpus linguistic studies: usually only very limited co-text is presented in published findings and this is hardly consonant with a discourse-level analysis of language behaviour. Discourse analysis has the significant advantage of trying to appreciate frequencies or relations over paragraphs or complete texts (or even several, distinct texts within a particular theme or interaction), and this should be employed along with more fine-tuned token or type counts focussing merely on the text within the corpus.



so, again underlining the coalescence between persona and point of view. When your name is questioned in daily life it is to be questioned as a person. (Hagström 2012 82)

When your opinions are called into question on a forum the effect is the same, if not more intensely so, and often a contributor's origins and identity are aggressively attacked, rather than the actual beliefs expressed. "The act of naming is an act of power," (Guenther, quoted in Hagström 2012, 90) but it requires affirmation because, being a speech act, it requires both illocutionary force and perlocutionary effect, and so the recipients have a part to play in this highly politicised game ('politicised' not merely because of the topic of discussion, but much more because of the social interactions in action). Bourdieu (quoted in Alia 2009, 6) claimed that everyone aspires to naming, and this process of naming was inseparable from other political phenomena, so perhaps it is not surprising to find power relations being enacted in the naming used on weblogs. It is other participants who accept or question the persona presented (the accusation of being a troll is the ultimate insult in speech act terms, revealing that 'the mask has slipped'). Levi-Strauss (1996) claimed that the power to name was steeped in political power, and so in naming oneself online there are the elements of this struggle. A word might be "a microcosm of human consciousness," but we should adjust this concept for internet interactions, suggesting that a digital name is a microcosm of the presentation of human consciousness; the representation of a person for a particular communicative need at a particular juncture. Of course, this representation will not necessarily be perceived as we intend by our recipients. As we found in our data, many references ran a high risk of falling away unnoticed or being simply impenetrable for many readers. Any conception of a name is based on personal experience and, especially, cultural knowledge, and what seems striking to one community might even be a common naming practice in another already.⁹ Of course the speech acts of contributors, and attempts at representing the self, require a response, and are essentially of a dialogic character. They initiate a conversation, whether it be friendly or hostile, in a way that our first names or family names usually do not. Digital identifiers do not have the obligation to documentation: instead they are acts within a postmodern version of Bruner's 'folk psychology' (Bruner, 137-8). It is hardly surprising that headers and messages were seen to be working in concert, as they are constituent parts of this "exercise in narrative and storytelling." And "the Self as narrator not only recounts but justifies" (Bruner 121), and so, again, we have a highly pragmatic performance that can only be fully appreciated at the level of discourse. If an identity is challenged, a narrative is disbelieved, and so a justification is rejected, and a contributor shamed.

Works cited

- Alia, Valerie. *Names and Nunavut: Culture and Identity in the Inuit Homeland*. New York: Berghahn Books, 2009.
- Baker, Mona. *In Other Words. A Coursebook on Translation*. Abingdon: Routledge, 1992.
- Bakhtin, Mikhail M. *Speech Genres and Other Late Essays*. Austin: University of Texas Press, 1986.
- Blommaert, Jan. *The Sociolinguistics of Globalisation*. Cambridge: CUP, 2010.
- Bruner, Jerome S. *Acts of Meaning*. Cambridge: Harvard University Press, 1990.
- Chapman, Richard. "Corpus Linguistic Analysis: How far can we go?" *Computer Supported Qualitative Research, 2nd International Symposium on Qualitative Research, ISQR 2017*. Eds. A.P. Costa et al. Springer, 2018. 42-47.
- Dörnyei, Zoltan. *Research Methods in Applied Linguistics*. Oxford: OUP, 2007.
- Dresner, Eli and Susan Herring. "Functions of the Nonverbal in CMC: Emoticons and Illocutionary Force" *Communication Theory* Volume 20, Issue 3 (2010): 249-268.
- Fairclough, Norman. *Critical Discourse Analysis*. Harlow: Longman, 1995/2010.
- Goffman, Erving. *The Presentation of Self in Everyday Life*. London: Penguin, 1969.
- Hagström, Charlotte. "Naming Me, Naming You. Personal Names, Online Signatures and Cultural Meaning" *Names and Identities*. Eds. Botolv Helleland, Christian-Emil Ore and Solveig Wikstrøm. Oslo: Oslo Studies in Language 4.2 (2012).
- Lévi-Strauss, Claude. *The Savage Mind*. Oxford: Oxford University Press, 1996.

⁹ Here we might mention naming in China, where common names can be lexically identical to words indicating affiliation or opinions (Tian and Wu 2007), rendering this act on weblogs much less impactful.



- McEnery, T., R. Xiao and Y. Tono. *Corpus-based Language Studies*. Abingdon: Routledge, 2006.
- Melville, Herman. *Moby Dick or The Whale*. New York: Harper, 1851.
- Pennycook, Alastair. *Language as a Local Practice*. Abingdon: Routledge, 2010.
- Prodromou, Luke. *English as a Lingua Franca*. London: Continuum, 2008.
- Ricoeur, Paul. *The Rule of Metaphor*. Abingdon: Routledge, 1978/2003.
- Spolsky, Bernard. *Sociolinguistics*. Oxford: OUP, 1998.
- Tian, Robert G. and Yan Wu. "Crafting self-identity In a virtual community: Chinese internet users and their political sense form." *Multicultural Education and Technology Journal* 1.4 (2007).