The SocWoC corpus: compiling and exploiting ESP material for undergraduate social workers

Jane Helen Johnson

Abstract

Successful social work is based on communication (Clark, 2000; Pierson & Thomas, 2000; Thompson, 2010), yet there is little consensus in the field as to what exactly constitutes proper communication (Richards, Ruch, & Trevithick, 2005; Trevithick et al., 2004) and where models of good communication might be found. It is therefore not surprising that there is little material available for teaching the language required for social services students studying English for Specific Purposes (ESP) at undergraduate level in European universities (Kornbeck, 2003, 2008). This chapter describes the creation of a specialised corpus of material related to social work to meet the needs of undergraduate students using L2 English in Italy. It focuses on specific needs of this population (such as the language of social work and recent migration patterns), then describes the selection and preparation of source material for a monitor corpus in this domain. The chapter then shows how the resulting corpus can be used by teachers to select samples and analyse texts for classroom discussion. In conclusion, it is suggested that such material will contribute to language development through awareness-raising in the classroom as regards the effects of different language choices in a discourse context, while also mentioning areas for further development.

Keywords: SocWoC corpus, undergraduate social workers, ESP material.

1. University of Bologna, Bologna, Italy; janehelen.johnson@unibo.it

1. Introduction

Growing specialisation and the need for awareness of how language works within particular domains has led to an increasing demand in Europe for courses in ESP at tertiary level. In the social services field, the lack of material for teaching ESP to social services students at undergraduate level has long been recognised (Kornbeck, 2003, 2008). Since it is widely acknowledged that appropriate language and discourse are fundamental for successful social services work (Thompson, 2010), this is no small problem. In order to fill this gap, it is essential to have a clear view of what is meant by ‘appropriate language and discourse’ in this field and be able to relate this to the type of language competence students need to do social work in their local context, since “[s]ystematic knowledge about language and practical awareness of how it works is fundamental to the process of building mature communities” (Montgomery, 1995, p. 251).

While social work students in Italy might have once perceived the English requirement of their degree course a mere formality, the current migrant situation in Europe has led to an urgent need for active focus on English in context. Specialised corpora can help reveal what language and discourse is ‘appropriate’ or at least recurring. In an ESP framework, corpus linguistics has mainly been used to focus on the behaviour of lexical items more or less in isolation, or at least only as regards collocational patterns. Boulton (2012) suggested that most researchers focused on smaller amounts of phraseology and collocations rather than discursive resources across the text. A more holistic approach focuses on discourse as it appears in the contexts of the different genres in a corpus.

This chapter shows how one such corpus was developed for an ESP course with social work students in an Italian university. First, a needs analysis of these students is presented, and then genres appropriate to social work are identified. Following this, the contents of the Social Work Corpus are described, including a detailed exploration of how one sub-corpus could be exploited for classroom use. Finally, suggestions for further development are made.
2. Needs analysis

Degrees in social work at Italian universities were aligned with European standards after the Bologna Declaration (1999), though there is often little compulsory social work theory and practice content (Frost, Höjer, & Campanini, 2013). Social work placement, often shadowing the staff of local authorities, is an important part of the course and social work graduates in Italy need to apply their university education to interpreting situations and intervening within their specific cultural and ethnic contexts (Campanini & Frost, 2004). We should note here that the current refugee crisis has led to some major changes in an Italian social worker’s mandate. Italy is often the first point of arrival for refugees fleeing from conflict areas, with a huge increase in the number of unaccompanied minors, as well as incoming migrant workers, particularly from African nations. Figures for 2015-16 (Menonna, 2016) show the main countries of origin to be Eritrea, Nigeria, Gambia and Somalia, with a sharp rise in numbers from Bangladesh: all countries where English is the main foreign language. Besides these new concerns, ongoing social work in Italy deals increasingly with the additional issues of second-generation immigrants (Simone, 2016).

As regards English language competence, admission to the degree course requires incoming students to have a B1 level in the Common European Framework of Reference for languages (CEFR). Social work undergraduates – in this author’s institution at least – tend to have a knowledge of general rather than specialised English and generally weak listening and production skills. While course attendance is not compulsory, students at Bologna University usually attend general English language classes prior to the ESP course with the aim of reaching a B2 level, focussing particularly on grammar and reading, since some of the core degree programme literature is delivered in English (Frost, Höjer, Campanini, & Kuhlberg, 2015). In their ESP course, therefore, students will need particular practice in speaking, listening and writing skills, while consolidating reading skills. Specialised terminology is best introduced in context, and language functions typical of social work also need to be presented and practised. The next section will discuss the importance of selecting appropriate learning material (Nesi, 2015).
3. Language and discourse for social work

Teachers need to evaluate the appropriateness of language and discourse, but this is challenging if the field is unfamiliar to them. The ideal informants are the social work practitioners themselves, given their familiarity with the required language use in social work practice, which the literature tends to discuss under the umbrella term ‘communication’. Communication is mainly described here as synonymous with spoken interaction, which has always provided the basis for successful social work (e.g. Clark, 2000; Cross, 1974; Day, 1972; Pierson & Thomas, 2000; Thompson, 2010; Trevithick et al., 2004).

It is no surprise then that much literature in social work studies has focussed on the communication skills that the social work student needs to acquire (e.g. Lishman, 2009; Richards et al., 2005; Shulman, 2006; Thompson, 2010; Trevithick, 2005; Trevithick et al., 2004; Woodcock Ross, 2011), although there is little consensus as to what exactly constitutes appropriate communication (Richards et al., 2005; Trevithick et al., 2004) and where good models might be found. To be relevant to our purpose here, we shall focus solely on references to communication as spoken or written discourse, leaving aside non-verbal aspects such as body language and dress. We will divide references from social work literature into two types: one concerning general language references and the other genre-based references.

3.1. General language references

Social workers require both production and reception skills: “all workers need to develop appropriate communication skills both for face-to-face and for written communications” (Pierson & Thomas, 2000, p. 95). Oral skills are of fundamental importance in social work (Juhila, Mäkitalo, & Noordegraaf, 2014b), since “social workers spend more time in interviewing than in any other single activity. It is the most important, most frequently employed, social work skill” (Kadushin, 1990, p. 3).

Issues of register-awareness and the connotations of language emerge from practitioners’ recommendations. For example, social workers need to “consider
the context in which they are required to speak and to write, and to ensure that they develop a style that is appropriate and relevant for their audience” (Pierson & Thomas, 2000, p. 95), since clients are “not best helped by workers who use obscure, inaccurate, deceptive or demeaning language” (Clark, 2000, p. 181).

The pragmatic function of discourse is especially important, since “[p]rofessionals are less effective on their clients’ behalf if they cannot communicate precisely and persuasively” (Clark, 2000, p. 181). Lishman (2009) lists the many different fields in which a social worker has to operate while giving an idea of the pragmatic functions of communication with which students need to be familiar:

“Effective communication […] includes, for example, providing basic care, giving advice, making assessments, providing care packages, counselling, writing reports, acting as advocates for service users, and working in interdisciplinary settings with health, education, housing and criminal justice” (p. 1).

This citation gives examples of both spoken and written discourse, as well as contextualising reading, writing, speaking, and listening skills in the professional life of a social worker. The importance of written communication skills in social work cannot be denied (Trevithick et al., 2004, p. 19). Lishman (2009) lists written material which may be required of the social worker, including letters, reports, and records for service users, agencies, and social workers.

Though the above authors were referring to a native-speaker environment, for the purpose of this chapter we may easily extend these examples to cover non-native contexts.

3.2. Genre-based references

A number of genres are described as particularly relevant to practitioners. Though dialogues between service users and social workers are important,
they inevitably refer to, draw on, and follow other background material, since “all kinds of written reports and interprofessional meetings support that talk to complete a trajectory” (Juhila et al., 2014b, p. 9).

Such written reports include case recordings, defined as “brief narrative accounts of contacts with and about clients, summaries, and plans of work covering a period of time” (Payne, 2000, p. 44), case studies (describing the issues at stake), supervisor commentaries, and case narratives, in which the social worker notes down his/her thoughts during the interview and describes the rationale behind the intervention. In relation to the refugee crisis, being able to read and compile material intended for cross-border use would be a major advantage for Italian social workers.

Both Shulman (2006) and O’Hagan (1996) mention different fields with which competent social workers need to be familiar, since “the knowledge underpinning social work practice derives from many different sources” (O’Hagan, 1996, p. 8) and students must “become attuned to the ways in which society, government, courts and professional associations influence practice” (Shulman, 2006, p. xxi). They should keep up with legislative developments such as revisions to ministerial responsibilities with repercussions on health and social care at national and local level, both at the source – through dedicated governmental websites – and trickling down through the printed media to the service user. In particular, students need an awareness of:

“law, social policy, philosophy (ethics), sociology, social administration, organisational policies, procedures and guidelines, numerous theories, [and] differing social work methods” (O’Hagan, 1996, p. 8).

Thus we may interpret these recommendations as requiring access to relevant material, including legislation and best practice recommendations produced by national and local governing bodies with relevance to social care, legal documents, as well as healthcare and medically related material. In an Italian context, this would provide material for contrast with students’ own national situation, as well as raising awareness about other international applications.
Some practitioners (e.g. Pierson & Thomas, 2000) also draw attention to the need for familiarity with the academic genre, particularly the research article. This is due to increasing emphasis on a research culture in social welfare, in which “workers need to be able to communicate clearly with funders, research colleagues and research participants in order to produce high quality results and be able to disseminate their findings clearly and imaginatively in order to improve practice” (Pierson & Thomas, 2000, p. 95).

Information is highly interconnected in a number of specialised fields (c.f. ESP and Law, in Breeze, 2015) and social work is no exception. ESP students need to be aware of these connections and the dissemination of information across different genres. News articles constitute useful sources of information for the social worker as regards public reception and perception of social work legislation. News articles are a potential means to encourage students to have a proactive approach to their profession. Brawley (1997), for example, uses newspaper discourse to promote media advocacy among social worker students, “with a view to correcting distorted messages about vulnerable people given by the press, developing policies and services, and relaying important messages to large target audiences” (Trevithick et al., 2004, p. 30).

Social work students should therefore cultivate awareness of a wide range of genres, including reports and examples of social care and social work interaction, editorials and news reports, academic and legal genres, and government documents. While non-native speakers in an Italian context would certainly need to be able to read and compare such material with similar texts in Italian, current social work in Italy may also involve actual production of such material in English.

To sum up, communication as described by social work practitioners may be broken down in linguistic terms as being genre-based and macro skills-based, including familiarity with particular genres and relative genre constraints such as register, lexis and phraseology, the need for pragmatic awareness, and practice in all four macro skills, particularly those hitherto under-developed speaking skills. In particular, students must become aware of the linguistic nuances of
spoken language in social work interviews in English in order to communicate successfully in the various interactions at the heart of social workers’ professional practices. It is worth noting here that the students’ university curriculum does not include active focus on the language used in their first language. Required reading prior to placement (Zini & Miodini, 2004) does feature a section on ‘language’, but only as regards what to talk about rather than how to say it and why. The focus of this chapter, therefore may be considered doubly useful for Italian social work undergraduates, since it applies to social work universally and is particularly important in a context where English is not the students’ first language.

Given the connections between the different aspects of a social worker’s profession, the sources of interest to language teachers should be considered as a network of resources. Once collated into an electronic corpus, authentic texts constitute relevant and topical material on which to base a meaningful language course for ESP Social Work undergraduates, following an integrated language-teaching approach whereby the four macro skills are taught in conjunction with each other. Such a corpus is described in the next section.

4. Introducing SocWoC

The Social Work Corpus (SocWoC) is intended as a repository of teaching material in electronic form. It has high ‘face-validity’ (Flowerdew, 1993, p. 239), including all the material the students are exposed to during their ESP course and featuring examples of many of the different genres the students may encounter in their future profession.

Divided into various sub-corpora, SocWoC is a monitor corpus in that it may constantly be added to, either vertically (by adding additional texts of the same genre) or horizontally (by adding new genres). It may thus also be updated to reflect new legislation or topics in the current interest. While its present

---

2. Though SocWoC was designed by the author, compiling a monitor corpus lends itself to collaborative effort and any additional material contributing to enhance the utility of a social work corpus would be much appreciated, in exchange, naturally, for access to the whole corpus.
composition, on which this chapter is based, is designed to give a ‘snapshot’ of current practice, thus consisting of documents published during the period 2015-16, there are plans to add material dating from earlier years to each sub-corpus to facilitate broader linguistic analysis. Though mainly consisting of material from UK sources ‘written to be read’, it also features transcripts of spoken material, including dialogues from social workers in a US context to allow any differences resulting from language variety to emerge. A brief description of each sub-corpus is given in Table 1.

Table 1. List of sub-corpora within SocWoC

<table>
<thead>
<tr>
<th>Sub-corpus</th>
<th>Description</th>
<th>Number of words</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAT</td>
<td>training materials</td>
<td>75,237</td>
</tr>
<tr>
<td>ACAD</td>
<td>academic papers</td>
<td>172,011</td>
</tr>
<tr>
<td>NEWS</td>
<td>newspaper articles</td>
<td>156,441</td>
</tr>
<tr>
<td>GOV</td>
<td>government guidelines</td>
<td>80,692</td>
</tr>
<tr>
<td>GLO</td>
<td>glossary of social work terms</td>
<td>15,534</td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td>499,915</td>
</tr>
</tbody>
</table>

With the exception of most of the MAT texts, the corpus material was downloaded from the internet and stored in separate files without incorporating metadata. Material for the MAT sub-corpus was difficult to access since social work does not have much tradition of being observed and recorded and social work interviews are confidential encounters. For this reason, selected social work practice books (Bisman, 2014; Woodcock Ross, 2011) were used as well as training material available online from reliable sources such as the SCIE website (www.scie.org).

The contents of the five sub-corpora in their current state are described as follows:

- MAT contains training material intended for classroom use by social work students in both the UK and the US in the form of case studies, case narratives, supervisor commentaries, and the transcriptions of social work interviews.
Chapter 8

- **ACAD** contains academic papers from the field of social work. It currently features some 30 articles published in the *British Journal of Social Work*, representing the British Association of Social Workers.

- **NEWS** features articles appearing in the Guardian and the Daily Telegraph: quality UK newspapers from both sides of the political spectrum. Texts containing the node words *social work* were collected.

- **GOV** contains a selection of government-sponsored examples of good practice in social care available online. Given SocWoC’s predominantly British focus, at least in this initial stage, sources included the UK’s Department for Education, the Department of Health, and government-funded institutions such as the Social Care Institute for Excellence.

- **GLO** contains social work terms with their corresponding definitions from glossaries in social work textbooks and websites, such as local government sites, and the online resources of institutions offering social work degree courses.

Metadata was included to associate each dialogue with the corresponding case studies and case narratives, while manual markups of the beginning and end of both social worker and service user turns was added to enable semi-automatic extraction of the discourse of the different categories of speaker. In the next section, we discuss possible uses for SocWoC by describing some initial corpus linguistic findings from one sub-corpus and giving some brief suggestions of how these might be exploited in the ESP social services classroom. We shall conclude by considering the corpus as a whole.

5. **Exploiting SocWoC**

Without the knowledge of or ability to practise effective communication, “the social worker is unlikely to use encounters with users of services or colleagues
for purposeful communication” (Lishman, 2009, p. 207). This is particularly true for speakers of other languages. Thus the teacher needs to focus both on awareness-raising activities and on language practice. In order to identify important language features, the teacher can use corpus linguistics techniques to query SocWoC in order to prepare teaching materials. Corpus linguistics is certainly nothing new in the field of ESP (Biber, Reppen, & Friginal, 2010, p. 559). It has been exploited in an ESP and English for Academic Purposes (EAP) framework (e.g. Boulton, Carter-Thomas, & Rowley-Jolivet, 2012) particularly for investigating genre and vocabulary. Examples of research into the possibilities offered by these techniques in relation to language teaching include Krishnamurthy and Kosem (2007), Hyland and Tse (2007), Ghadessy, Henry, and Roseberry (2001), Scott and Tribble (2006), O’Keeffe, McCarthy, and Carter (2007), Breeze (2015), and Boulton (2016).

In what follows I provide suggestions for teachers to exploit SocWoC with students, although it is of course also possible to give learners themselves direct access to the data in line with ‘data-driven learning’ (e.g. Johns, 1991; Boulton, 2016). Given space restrictions, I will focus here on an analysis of the MAT sub-corpus alone, more specifically the part containing the transcripts of social work interviews.

5.1. Mining the MAT sub-corpus for classroom use

Preliminary corpus investigation highlighted a variety of elements meriting further attention in the classroom. The examples given below are mainly concerned with pragmatics, such as preferred choice of process according to speaker turn, the use of interrogatives, the discussion of ‘feelings’ in what are often extremely distressing situations, the importance of figurative language and paraphrasing, as well as issues regarding language variety. A final example is a semantic analysis.

Since the dialogues in MAT are tagged for speaker turns, wordlists may be run for either speaker. Social worker turns, for example, contain a predominance of mental processes (want, need, like, think):
(1) I am glad you are sharing this with me. You have told me in previous sessions that at various times in your life you have had these feelings. What do you think is going on now that you are considering killing yourself?

(2) What I want to do, first of all, is explain to you why we are getting together, if you like. [...] We need to do an assessment today, aspects of your life and things that might affect your offending behaviour that we can work on together. Do you know anything about the Youth Offending Team?

These mental processes often appear in interrogative forms as in Examples 1 and 2. Interrogatives are used not only to establish facts (e.g. Do you know anything about the Youth Offending Team?), but also to help the social worker to “put feelings into words” and “reach for feelings”: an important part of the social work interview (Shulman, 2006, p. 136), as in Example 1 (What do you think is going on now that you are considering killing yourself?). The different functions of interrogatives in the social work interview lend themselves to class discussion, while question types and their functions (e.g. Richards et al., 2005; Trevithick, 2005) can be further investigated in the classroom using SocWoC.

Example 3 illustrates the utility of figurative language for discussing feelings in social work talk.
Searching the corpus for *like* reveals the same function fulfilled by phrases such as *sounds like, is it like*, as in Example 4:

(4) It sounds like it's quite painful ... and that when you're walking, checking over your shoulder, is it like you're running away from something?

Classroom exploitation could focus on classifying the pragmatic functions of such figurative language in both speaker turns. The frequency of mental processes used in the interrogative form, interrogatives generally, and phrases acting as explanations all relate to *facework* and politeness (*Goffman, 1955*), which are central issues in social work interviews (*Juhila et al., 2014b*).

Wordlists show a high frequency of reporting verbs, whose different functions in social work have been described by *Juhila, Jokinen, and Saario (2014a)*. Reported speech also plays an important part in *paraphrasing*, a fundamental communication skill (*Shulman, 2006; Trevithick, 2005*). An example is shown in Example 5:

(5) Just take a minute and think about how you are feeling. You have said before that you have trouble being aware of and articulating your feelings.

In Example 5 the service user’s words (signalled by ‘you have said before’) are paraphrased by the social worker. This enables him/her to refer to the service user’s problems with discussing feelings, and thus move towards a possible solution. In class, students could look at examples of paraphrasing marked by reporting speech and practise transforming the service user’s words in different contexts.

Since both UK and US English is represented in the material in MAT, differences between these two *language varieties* may be highlighted. A frequent phrase in the British dialogues was *at the moment*. While this phrase was not found in the US social worker turns, the same meaning emerged in the frequent phrase
right now. Similarly differences in verb colligation emerged: *talk with* versus *talk to*, for example. Such investigations in the classroom make it possible to promote awareness-raising as regards language variety.

Finally, an investigation into the **key semantic fields** in the dialogues provides opportunities for focussing on ESP phraseology. Using WMatrix (*Rayson, 2008*), the dialogues were compared with the spoken section of the British National Corpus. One of the top semantic domains was ‘worry’. While this in itself is not unexpected, given the nature of most social work interviews, intuition alone might not be able to supply all the vocabulary related to this field (examples in Table 2).

**Table 2. Concordances exemplifying the semantic domain ‘worry’**

<table>
<thead>
<tr>
<th>I act like I don’t</th>
<th>care.</th>
<th>But then when I ‘m in</th>
</tr>
</thead>
<tbody>
<tr>
<td>mom to hit you. We have serious</td>
<td><strong>concerns</strong></td>
<td>about your bruises</td>
</tr>
<tr>
<td>you have said before that you have</td>
<td><strong>trouble</strong></td>
<td>being aware of and</td>
</tr>
<tr>
<td>Doug, you seem a little</td>
<td><strong>anxious</strong></td>
<td>about me being here.</td>
</tr>
<tr>
<td>Graeme had presented a lot of mental</td>
<td><strong>distress</strong></td>
<td>, and had been hungry, cold</td>
</tr>
<tr>
<td>Yoko is under a lot of</td>
<td><strong>stress</strong></td>
<td>right now</td>
</tr>
<tr>
<td>That list of things that</td>
<td><strong>bother</strong></td>
<td>you included feeling unloved</td>
</tr>
</tbody>
</table>

Concordances from this semantic domain may be used to teach a range of sometimes very sensitive vocabulary choices in context, including collocations (*serious concerns*), colligation (*anxious about*), specialist combinations (*present + mental distress*) and phraseology (*under + [quantity] + stress*). These are just some examples of the types of language items that could merit further linguistic investigation by/for students in one sub-corpus of SocWoC. The final sub-section will examine discourse in context across the different parts of SocWoC.

### 5.2. Classroom focus across SocWoC

The social work profession is people-focussed. One way of identifying key participants is to compare wordlists from each sub-corpus with a general reference corpus such as the BNC:
One activity could be to investigate the differences and similarities between the items in these lists. Why do we find refugees only in one part of SocWoC? Which participants appear in all parts of SocWoC? Investigation could continue into a comparison of the items, thus allowing focus on, for example, preferred or excluded terms in particular genres. Corpus tools such as SketchEngine (Kilgarriff, Rychly, Smrz, & Tugwell, 2004) may be used to generate a word profile to focus on the search word in context. The most frequent processes associated with children, for example, were found to be different across the sub-corpora, as was the positioning of children as subject or object. This sort of investigation is particularly important for non-UK based social work students, since the context where these students expect to practise may be different from the UK context. In the family-oriented Italian system, families are the main actors in solving people’s problems whereas the state plays a subordinate role (Frost et al., 2013, p. 341). This is slightly different to the more individualistic cultural models in the UK and US and students must be aware of how people’s needs are dealt with in the different contexts.

We can, of course also look outside the corpus to gain further information regarding, for example, current problematic issues. With reference to the investigation of children outlined above, we might decide to follow the trajectory of children’s social work through the discourse of different genres
in relation to legislative items such as the recent Children and Social Work Bill and its implications for social work in the UK, including both attitudes to and receptions of such measures.

6. Conclusion

Italian social work undergraduates were given little theory and practice as regards the social work interview in their own first language, so that focus on the social work interview during lessons was more than just ‘language comparison’ but actually became a way to teach the functions of language in this context. Thus, we might consider that the way forward for the ESP teacher is increasingly moving towards combining language with content teaching, in order to be able to fully exploit the nuances and functions of the language. This requires the ESP teacher to be ever more prepared. In this chapter, we have described how a specialised corpus can be created and can inform the development of material for use in the classroom activities of an ESP course. The chapter focussed on the needs of a particular subset of ESP students, and suggested that information from social work literature could be used to guide the compilation of a suitable corpus of classroom material. In line with the student-centred approach of task-based language learning (e.g. Nunan, 2004), corpus linguistics techniques may then be applied in order to exploit this material in the classroom.

Further development and subsequent exploitation of SocWoC will need to take into account other factors important for the Italian social work context, particularly the need for texts to be appropriate for present-day social work in Italy, given the recent influx of immigrants using English as a lingua franca (Simone, 2016). Future work will thus involve refining the needs analysis, for example through pre-service and stakeholder surveys (e.g. Chovancová, 2014). Further development would also require fully integrating the findings briefly described here within the design of a comprehensive course for social work students, while it would also be important to be able to assess outcomes in terms of language awareness in students on completing such a course.
References


https://doi.org/10.1017/CBO9780511667336

Jessica Kingsley Publishers.

https://doi.org/10.1017/CBO9780511497650


https://doi.org/10.1075/ijcl.13.4.06ray

https://doi.org/10.1080/02615470500096928

https://doi.org/10.1075/scl.22


