

‘Psy’ expert evidence in the family courts: The potential for corpus-assisted analysis

Lauren Devine, Stephen Parker, Leigh Harrington & Nadia Makouar

Centre for Language and Law, Aston Institute for Forensic Linguistics
Aston University, UK

https://doi.org/10.21747/21833745/lanlaw/9_1a5

Abstract. *This article introduces corpus-assisted linguistic methods as an exploratory means of analysing expert psychologists’ reports used in public family law (child protection) cases. Analysis of this dataset is a new application for corpus linguistics (CL) and the primary purpose of this article is to explore viability and potential for its future research using CL as a core method. For this study we have created and analysed a 25 single-text-type specialised written corpus consisting of 25 expert psychologists’ reports (the Psychology Report Corpus “PRC-25”). The reports are a random sample selected from a population of all psychologists’ reports held in Cafcass files over a 10-year period, representing the first corpus of its kind in a currently under-researched area. Our study uses both an inductive (data-driven) approach to identify significant themes and topics in the reports, and a deductive (legal-intuitive) approach to explore psychologists’ use of legally significant terms, especially risk of and significant harm. We also explore the possibility for using this new methodological protocol to triangulate analysis of a larger and representative corpus of expert psychologists’ reports, and the possibilities for corpus-driven analysis of the genre of written expert evidence text types more generally.*

Keywords: *Expert evidence, Psychologists’ reports, Family justice system, Corpus linguistics, Child protection.*

Resumo. *Este artigo apresenta métodos linguísticos baseados em corpora como forma exploratória de analisar relatórios periciais de psicólogos utilizados em casos de direito da família (proteção de menores). A análise deste dataset constitui uma nova aplicação da linguística de corpus (CL), pelo que o objetivo primordial deste artigo é explorar a viabilidade e o potencial para a sua investigação futura utilizando linguística de corpus como principal método. Para este estudo, criámos e analisámos 25 corpora escritos especializados com um único tipo de texto, constituídos por 25 relatórios periciais de psicólogos (o Psychology Report Corpus “PRC-25”). Os relatórios constituem uma amostra aleatória selecionada a partir de*

uma população de todos os relatórios de psicólogos existentes nos arquivos Cafcass por um período de 10 anos, representando o primeiro corpus do género numa área atualmente sub-investigada. Este estudo utiliza, quer uma abordagem indutiva (baseada em dados) para identificar temas e tópicos relevantes nos relatórios, quer uma abordagem dedutiva (legal-intuitiva) para explorar a utilização de termos juridicamente relevantes por parte dos psicólogos, especialmente riscos e danos relevantes. Também exploramos a possibilidade de utilizar este novo protocolo metodológico para triangular a análise de um corpus maior e mais representativo de relatórios periciais de psicólogos, e as possibilidades de análise baseada em corpora do género de textos prova pericial, em geral.

Palavras-chave: Prova pericial, Relatórios de psicólogos, Sistema judicial de família, Linguística de corpus, Proteção de menores.

Introduction

This study was conducted as a pilot, extending a project investigating why Section 31 care order applications in England are consistently high.¹ Interviews conducted with lawyers and the judiciary at this start of the project revealed themes and trends of concern to those involved in decision making in the family courts. In these interviews, psychologists' expert reports were highlighted as an area of concern, particularly by the judiciary who felt more clarity in the use of legal language would assist them in making decisions. This supports findings in previous research, carried out for the Ministry of Justice's Family Justice Council by a team of inter-rating forensic and clinical psychologists using non-corpus methods to conduct largely qualitative research. This research found numerous issues of concern in the reports, including unsuitable and inconsistently used psychological tests, unsupported inferences drawn by the expert psychologists, misreporting of allegations as fact, and unnecessarily pejorative personal comments (Ireland 2012). New standards for expert witnesses in Children's Proceedings in the family courts were drafted as a response to these (and other) concerns to assist with the future focus and relevance of expert psychological evidence in child protection proceedings (Ministry of Justice and Family Justice Council 2013). For the analysis reported here, a pilot corpus, the PRC-25, was built and used to identify and examine the use of selected legal terms in psychologists' reports, as the judicial interviews indicated that the use of legal terms by experts was still an area of concern.

Expert psychologists' reports play a significant role in the court's decision about whether a child remains with their family, or whether parental responsibility is transferred to the State (Children Act 1989, Section 31). Other than Ireland (2012), there is no substantive empirical research into the content, quality, or consistency of these reports as forms of expert evidence, and Ireland's study was not representative. This article explores the potential application of corpus-assisted methods to "problem-solving" practical issues arising in forensic courtroom interactions, such as evaluating the focus and relevance of the expert psychologists' reports which can be applied to larger, representative corpora.

Corpus-assisted methods can be used to provide a quantitative overview of corpora and to contextualise qualitative results. Primarily a computational methodology using statistical procedures to examine large, machine-readable (or electronic) linguistic datasets (corpora), corpus linguistics (CL) is an established methodology in numerous

areas of linguistic interest across written and spoken contexts, including forensic linguistics (Gillings 2022), political discourse (Taylor 2022), news media (Marchi 2022), language evaluation (Scott 2010), and online informal legal advice in child protection cases (Tkacukova 2020). In conceptualising the paper, the authors also drew on the growing body of CL literature which considers and encourages the potential of corpus linguistics in new areas (Brezina and McEnery 2015; Brezina 2018), selected linguistics literature (Lebart *et al.* 1997; Partington 1998; Sinclair 1999; Cutting 2007; Baker *et al.* 2008), and relevant law and related documentation (Children Act 1989; HM Government 2018; Family Procedure Rules, Part 25 2020). As so little previous linguistic research on this data exists, established methodological protocols for their systematic analysis do not exist. Unless corpus assisted methods are used, time and resource limitations mean research leans towards small-scale qualitative studies conducted on an individual-report level, which is the limitation reported by Ireland (2012). As such, CL offers a promising avenue for the analysis of this type of forensic expert evidence and platform from which detailed linguistic analysis can be undertaken. A further limitation of previous research drawing on expert witness reports has been gaining access to and extracting meaningful large-volume data for analysis as the data were created for litigation, not research purposes. Although the detailed work involved in creating a corpus cannot be underestimated, a corpus of such data can be re-used to answer a variety of research questions.

Transparently reported, CL findings are reliable and replicable, offering insights into salient, as well as more subtle linguistic patterns in a corpus. As such, Charteris-Black and Seale (2010) assert that these tools can provide ideal avenues into the exploration of new corpora in principled ways based on statistical saliency rather than exclusively on researcher intuition. However, Baker *et al.* (2008: 274) observe that most corpus analysis does nonetheless require human input at various stages, whether deciding which texts to include when creating a corpus or when interpreting the findings of quantitative tests. Existing CL research points to the analytical synergies achieved by supplementing qualitative approaches with quantitative ones, the most suitable approach depending on the specific research questions (Baker *et al.* 2008; Biber *et al.* 1998).

The Psychology Report Corpus-25 (PRC-25, Table 1 below) is a specialised single text-type written English corpus created by the authors of this paper, comprising 25 expert psychologists’ reports used as expert evidence in child protection cases between 2011-2021. The corpus contains 235,111 running words, 11,763 types (the number of distinct words) and 9,436 lemmas (the form of words under which they are registered in a dictionary, for example, “parents”, “parenting” and “parental” belong to the same lemma “parent”). The average length of the reports is 11,048 words (longest = 31,799, shortest = 1,311).

Name	Language	Texts	Tokens (Total word count)	Additional information
Psychology Report Corpus (PRC-25)	English	25	235,111	Types: 11,763 Lemmas: 9,436

Table 1. The PRC-25

To create the PRC-25, the authors accessed the whole database of expert psychologists' reports compiled across England and Wales between 2011-2021, held by the Children and Family Court Advisory and Support Service (Cafcass).² The total population of reports was >10,000, ranging from 4 to 120 pages. Each report was extracted by the authors from the Cafcass files and organised into an encrypted database. The pilot corpus was created from a random sample selected using the Excel RAND() function. The sample is too small to be representative of all psychologist reports, but it is sufficiently sizeable that its 235,111 words would be difficult to analyse using manual, qualitative methods alone. Analysis was undertaken using two open access CL tools, #LancsBox v.6 (Brezina *et al.* 2020) and IRaMuTeQ (Loubère and Ratinaud 2014), a data analysis programme using R software and python coding language used for multi-dimensional text and questionnaire analysis. The PRC-25 was also compared with a larger, reference corpus, selected to represent contemporary general English for the same time period. This comparison was carried out to determine how the language used in expert psychologists' reports may be distinct from that of everyday English by statistically comparing the frequency of words in the PRC-25 to their frequency in a general reference corpus Stubbs (2010). Reference corpora are typically very large corpora that represent a chosen language across a variety of contexts within a defined timeframe. For this comparison the BNC Baby was selected, a four-million-word balanced subset of the British National Corpus, 2014 edition designed to be representative of present-day English-use. The BNC Baby 2014 is compiled of written and spoken texts in British English, including newspapers, fiction, academic writing, informal speech, and e-language.

For the analysis, an exploratory two-pronged approach was adopted, drawing on the disciplinary expertise of lawyers and linguists, all with knowledge of both disciplines. The aim of this exploratory analysis was to identify the frequency and general use of pre-selected legal terms used in the reports. These are important terms because the judiciary need experts reports to be "focussed and relevant" to tie their psychological evaluations to the relevant legal thresholds and criteria (Munby 2011). The first prong was a deductive legal-intuitive approach, informed by prior legal knowledge. Concordance and collocation tools (explained in Section 3.1 below) were used in this approach to analyse pre-selected legal terms. The second prong of the analysis was an inductive, data-driven approach. Drawing on no prior contextual knowledge, the main themes and topics in the PRC-25 were explored using keyword analysis. A hierarchical cluster analysis was also undertaken as part of the second approach. Overall, the aim was to consider whether the two-pronged inductive and deductive approaches might provide an effective means of methodological triangulation (Egbert and Baker 2020).

The expert psychologists' reports are intended to be documents to assist the court in their decision making, and as such are not primarily intended to comment on legal terms. However, to be able to contextualise the reports, some reference to the psychologists' findings in relation to the appropriate legal threshold terms might reasonably be expected. In the reports, the psychologists give their opinion in relation to parenting capacity, presenting an entextualised version of the verbal interviews between the expert psychologist and the respondent. This includes the psychologists' interpretation of other reported data, such as accounts from third parties. The role of predictive expert evidence based on the psychologists' estimates of *future* risk plays a significant part in the threshold question for the judiciary. This is a challenging task for both expert

and judge, as protracted social work intervention into private family life creates lengthy and often highly contested versions of complex real-life scenarios, which are not easily translated into items of evidence that will assist judicial decision making. To translate the lived reality of family experience into evidence enabling a court to adjudicate on whether a child should be removed from their family, there are procedural and structural measures in place intended to assist, set out in Practice Direction 12A (Ministry of Justice 2021a) of the Public Law Outline (PLO) and Practice Direction 25A (Ministry of Justice 2021b). It is in the context of this framework that the expert psychologists' reports are prepared and given in evidence.

Some initial observations were made about the reports during the corpus creation process. The authors did not conduct an analysis of register conventions in the PRC-25 but noted that the reports should be presented in the format required by PD25A and adhered to the conventions of the British Psychological Society's Guidance British Psychological Society (2016). The methodological approaches used in the reports were not standardised across the sample, but broadly included (1) collecting antecedents of the parents, including a history of their own childhoods prepared from verbal recollection by the parents; (2) interrogation of the parents in relation to the matters the Local Authority complains of in relation to their parenting ability (but not necessarily presented to the parent as an allegation which requires an answer); (3) assessment of the parents' psychological state; (4) assessment of the parents' willingness and ability to comply with Local Authority requirements; and (5) the concluding opinion of the psychologist about what is likely to happen in the future.

There was a lack of methodological transparency and rigour across the corpus in relation to the selection criteria for the multiple psychological tests used in the reports and no explanation of reliability, or how the conclusions in the reports were informed by the results. There was a lack of standardisation in relation to the final presentation of the individual reports. All the reports typically contained a 'past/present/future' narrative trajectory comprising: 1) the 'histories'; 2) the present situation; and 3) a prediction of what will happen next, but there was no particular order or structure to how this information was presented, and a wide variation in length. For example, one report contained a long continuous section titled "Interview with X parent", whereas another breaks the interview down into multiple themed sections, such as "Early Years", "Attitudes towards the task of parenting", and "Relationship with parents".

Finally, there were insufficiently transparent protocols for reporting the verbal interviews and explaining how inferences are drawn from them. This included gaps in explaining the methods of converting the spoken interviews into written data. It was unclear, therefore, how the experts conducted the interviews themselves, and no information is provided regarding the experts' preparation for the interview, or how they were conducted, recorded, transcribed, or interpreted. Many of these observations are consistent with the findings of Ireland's (2012) study.

Given these observations, it was hoped that corpus-assisted analysis would not only reveal the legal terms in the reports that are germane to the decision-making process. The analysis also considered how they were explicitly used, and whether precise terms were used consistently, as these terms are understood to have a consistent and specific meaning in the legal process.

The legal framework

The reports form a small but important part of a complex process of law and legal process and must be considered in context. The standard of proof in child protection cases is significantly lower than in the criminal courts but nonetheless has significant and long-lasting effects on children and their families (see for example Broadhurst and Mason 2013a). Case statistics are compiled each year, showing a consistently high level (see Figure 1). Experts' reports are important and necessary to assist judicial decision making, but the latest major government review of expert evidence in general found that poor expert evidence is problematic, particularly in relation to the difficulties in communication between expert witnesses and legal professionals (Law Commission 2009, 2011). Although the Law Commission specifically included cases involving child abuse (albeit in the criminal courts), expert psychologists' evidence in relation to child protection cases remained problematic (Ireland 2012). Specific guidance for psychologists giving expert evidence in the family courts exists, but no substantive evaluation of any improvements in psychologists' evidence has taken place since it was published (British Psychological Society 2016).

There are two substantive sub-fields of law involved in child protection cases: Public family law and expert evidence. Public family law operates as a sub-field of family law but is arguably also a sub-set of public law. Expert evidence operates as a sub-field of the broader law of evidence. The positioning of public family law as a sub-field of family law rather than a sub-field of public law, or even a re-named field of its own, raises complex issues of State power and individual vulnerability insufficiently recognised in the family law process (Devine 2017). The litigator in Public family law cases is the State, referred to as the applicant (Children Act 1989), the party defending their position is the party(ies) with parental responsibility for a child (generally the parent(s), almost always the mother), and the victim (alleged victim until a judgment is given) is the child(ren). Regardless of age, the child is simply the subject of the proceedings and cannot halt or dictate their progress or outcome (Children Act, 1989).

Drawing both sub-fields together in this investigation draws attention to the paucity of empirical research in either. This is both a consequence of general difficulties of gaining access to the data needed for analysis and the difficulties in manually analysing large volumes of courtroom data which has not previously been collated for research purposes. Exploring CL methods that will enable swifter consideration of larger volumes of data is hoped to benefit studies of larger specialist corpora of similar data types.

Public family law (child protection) cases

Child protection cases are the mechanism by which children considered to be "at risk of significant harm" can be removed into State care following an application by the Local Authority under Section 31 Children Act 1989. Judgments in child protection cases are focussed on the child's welfare as the paramount consideration (Section 1(1) Children Act 1989). The threshold legal test is set out in *Re B (Care Proceedings: Appeal)* [2013] and see Copley and Lowe (2009) for comment on the earlier House of Lords decision).

Experts' reports provide substantive evidence in many cases and are often the sole source of information in relation to what is likely to happen to a child if he or she is not removed. As such, the reports primarily relate to the prediction of future risk to a child if the child is left with their family (or in their current place of residence). Predicting risk is

inherently risky, and normally requires expert forensic input, hence the long tradition of reliance on psychologists: Lady Justice Butler Sloss outlined the need for expert opinion evidence to assist the courts in general when children are involved:

“Many if not all family cases involving children feature expert opinion evidence ... In cases involving children, expert medical and psychiatric evidence from paediatricians and allied disciplines is often quite indispensable to the Court. As Parker L.C.J. said in *Director of Public Prosecutions v A & BC Chewing Gum Ltd.* [1968]1 Q.B. 159 at 165A, when dealing with children, the court needs 'all the help it can get.'” (Re M & R (Child Abuse: Evidence), 1996)

The use of expert psychologists' evidence in child protection proceedings is thus an important area for research. The prediction is intended to assist judges in deciding whether the risk of making no court order is greater than the risk of making an order. Following Ireland's study (2012), there is no recent substantive empirical research into their content, quality, or consistency, although there is detailed literature in relation to the language of courtroom interactions more generally (for example, Coulthard *et al.* (2016); Aldridge (2010)).

Psychologists' reports for expert evidence are generally commissioned to comment on two main areas: parental (or caregiver) and the child's functioning to establish whether there is a problem (which will have been articulated by the applicant Local Authority acting on behalf of the State), and to give predictive comment on what is likely to happen to the child in the future (risk prediction). Both areas are controversial as they necessarily involve elements of methodological uncertainty (Regehr *et al.* 2010; Camasso and Jagannathan 2000): psychologists often have little to assist them other than verbal recall and accounts of past events. In addition, the evidence-gathering process for the psychologists differs significantly from a clinical evaluation as the primary role of the psychologist as expert is not to offer therapeutic support to the interviewee. In the PRC-25, the respondents reported profound fear and trauma during their interviews in finding themselves facing child protection processes. In addition, most report having suffered historic abuse, and many were facing current domestic abuse situations. This reflects a well-established observation in published literature (Lindley 1994; Broadhurst and Mason 2013b). The full reports also include events involving third parties that could not be independently verified, for example, respondents are routinely asked for retrospective accounts of their own, often traumatic, childhood experiences (Broadhurst and Mason 2013a). In clinical practice, a psychologist is undertaking a therapeutic task which is not compatible with the aims of adversarial litigation where the subjects of their inquiry may be both in urgent need of psychological support (for general explanation about clinical judgment see, for example, Goldberg (1968)).

Child protection cases are heard in the Family Courts, where the standard of proof is the 'balance of probabilities' rather than the criminal standard of 'beyond reasonable doubt'. This means that for the Local Authority applicant to obtain a judgment in their favour, their case must be stronger than the respondent's case: more likely than not; more than 50/50. In contrast, for the *respondent* to obtain judgment in their favour their case does not, in theory, need to be stronger, it simply needs to be at least of equal merit to that of the applicant. If both arguments are equally compelling the applicant should fail. In practice, most cases brought by Local Authorities result in an order being made even if it stops short of a care order, which of itself raises interesting research questions

(Redhill and Roe 2021: 31). It could be that Local Authorities are simply particularly good at only litigating where their case is likely to succeed on its merits, but it could also be that the power imbalance and vulnerability of the respondent(s) leave them at significant disadvantage.

Courts follow the procedures in the Public Law Outline (PLO) including a timescale for cases to be heard within 26 weeks (Ministry of Justice, 2021). The PLO was intended to help reduce the number of children removed, but in practice has placed considerable pressure on litigants and the courts to conclude cases in a timescale that makes therapeutic intervention a short-term aim rather than a long-term gain. Lengthy therapeutic intervention is frequently a recommendation of the psychologists' reports, leaving the courts with a conclusion that the intervention cannot be carried out within the time frame and thus the child must be removed. Ironically, it is increasingly the case that cases fail to meet this criterion anyway; in 2019 (Q2) the average case duration was 35 weeks. In 2021 (Q4), the average had risen to 44 weeks (CAFCASS 2022), considerably increasing the load on the courts.

A further factor of concern, also considerably increasing the load on the courts, is an increase in the rate of child protection cases (see Figure 1). Since its national roll-out in 2014, the rate of child protection cases increased year on year until 2016/17. Since then, the rates have not increased but have remained unacceptably high (CAFCASS 2022).

Expert evidence in child protection cases: Psychologists' Reports

The courts increasingly rely on expert evidence to help them understand areas outside their expertise (Law Commission 2011). Child protection cases are exceptionally reliant as "*many if not all family cases involving children feature expert opinion evidence...*" (Re M & R (Child Abuse: Evidence), 1996). Normally witnesses are barred from giving their opinions and are only allowed to testify on factual events which they have witnessed, and to what they know of their own thoughts and actions as they relate to the case (Hollington v Hewthorn 1943). Expert witnesses in child protection cases (and elsewhere) are the exception, as experts are required to give their opinion by using their specialist expertise to help the court to understand facts and evidence, which non-experts in their field would not be able to reliably evaluate otherwise (Family Procedure Rules, Part 25, Section 4.1 2020).

Some types of witness evidence are more 'certain' to a known confidence limit. Physical or documentary evidence can be analysed and explained to the court through the lens of the expert's translation into lay language, or the science can be presented with a quantifiable confidence limit (Law Commission 2011). These types of evidence require explanation to the court, but it is unlikely that other experts in the same field would come to differing conclusions given the objective evidence underpinning their evidence in fields such as engineering or accounting, for example. However, in some areas of expertise, such as psychologists' expert evidence, the data is less certain as it is largely collected from subjective, verbal accounts of past events provided by the respondents.

Existing literature in other areas of forensic linguistics has already established that witness testimony merits linguistic examination (for summary, see Coulthard *et al.* 2016). Extensive literature in child protection research detailing the complex nature of risk prediction also suggests scope for linguistic analysis (Møller Jørgensen *et al.* 2021; Smeeton 2020; Parton 2011; Munro 2010; Gillingham 2006). The courts are tasked with



Figure 1. Child Protection Cases: 2014/15 – 2021/22 (CAFCASS, 2022)

making difficult decisions, which can have a significant impact on people's lives, and a wrong decision can be catastrophic. No feedback is given to the courts on outcomes of their decisions. For all these reasons, great emphasis is given in child protection cases to the importance of experts. In the same judgment referred to above, Lady Justice Butler-Sloss emphasises the importance of expert evidence in child protection cases, stating:

“... when the judge is of the opinion that the witness' expertise is still required to assist him to answer the ultimate questions (including, where appropriate, credibility) then the judge can safely and gratefully rely on such evidence, while never losing sight of the fact that the final decision is for him.” (Re M & R (Child Abuse: Evidence), 1996)

Courts are therefore wanting to be sure that they make the right decisions, looking to the expert to give them as much certainty as possible as to what that decision should be. The general rule for experts is that they can give their opinions but not comment on the ultimate issue (the decision that the court is there to make), on the basis that it is the court, not the expert, which needs to decide the matter at hand. However, and uniquely, in the family courts, the Family Procedure Rules (FPR) do allow experts to comment on the ultimate issue as they are trying to base a decision on a prediction of what might happen in the future. This gives the psychologists' reports significantly more influence over the court's decision than other types of expert evidence.

Methods: a two-pronged approach

For this exploratory study a two-pronged approach to the analysis of the PRC-25 was adopted. Both methods, the deductive legal-intuitive approach, and the inductive, data-driven approach, employ core corpus methods and offer different disciplinary perspectives which lend themselves to a comprehensive overview of a corpus. The tests used were concordance and collocation analysis for the deductive approach, and keyword analysis and hierarchical clustering for the inductive approach. These approaches collectively offer insights into linguistic patterns in the PRC-25 and provide a platform for additional, future qualitative analysis.

Deductive legal-intuitive approach

An advantage of corpus-assisted analysis of large volumes of legal texts and forensic materials such as those considered in the PRC-25 is its ability to respond to specific research questions. For this approach, the authors set out to test the potential of the PRC-25 to respond to pre-set areas of interest, in this case the use of specific legal terms. This type of analysis would respond well to questions asked by subject experts, such as those within the legal profession, the judiciary, psychologists and other experts, policy makers and academics. The selection criteria for these terms specified that they must be in either the primary and secondary legislation (Children Act 1989; HM Government, 2018) and that they must represent terms that experts are routinely instructed to assist the court with. The list for this latter criterion was identified from the written instructions to the psychologists (to which the authors have access as a contextual note attached to the original reports).

The terms, set out in Table 2, are: *best interests*; (HM Government 2018 p. 28 para. 58), *significant harm*; (HM Government 2018, p.28, para.58 (2)A) and Section 31(2)(a) Children Act 1989), *risk of*; (Children Act 1989, c. 41., Section 1(3)(e) England), *reasonable*;

(Children Act 1989, Section 47); *a concern*; (HM Government 2018, p.17 para. 17), and *opinion*; (Family Procedure Rules, Part 25, Section 4.1 2020). Table 2 lists their absolute and relative frequency and dispersion across the 25 texts. The asterisks, denoting wildcard searches on most corpus linguistic software, are included as these searches retrieve words directly matching the characters before the asterisk with any other combination of characters immediately after it. For example, searches for *reasonabl** retrieved instances of *reasonable* and *reasonably* from the PRC-25).

Legal term	Absolute Frequency in PRC-25	Relative Frequency (per 10k)	Reports appeared in PRC-25
<i>Best interest*</i>	20	0.85	11/25
<i>Significant harm</i>	2	0.09	2/25
<i>Risk of*</i>	75	3.19	14/25
<i>Reasonabl*</i>	33	1.4	12/25
<i>A concern*</i>	6	0.26	6/25
<i>Opinion</i>	159	6.763	22/25

Table 2. Pre-selected legal terms

This approach then combined the corpus linguistic tools of concordancing and collocation to analyse the set of pre-selected legal terms.

Concordancing is a tool central to all corpus analysis which enables users to display all examples of a search term or phrase, in its original context in the corpus, with a defined span of words (typically up to 20) to its left and right (Rayson 2015). A such, concordancing enables analysts to view large numbers of examples together in one place, and concordance lines around the search term can be sorted according to different criteria which help highlight patterns in the search term's use in the corpus (Baker 2006; Hunston 2010). Using the Key Word in Context (KWIC) function in the corpus software #Lancsbox, we generated concordance lines in the PRC-25 for each pre-selected legal term, enabling our lawyer-analysts to gauge their general use in the psychologist reports.

The concordance analysis was paired with a separate CL tool, collocation analysis, to identify the collocates of the terms *harm* (n=100, occurring in 13/25 texts) and *risk* (n=254, occurring in 14/25 texts). A collocate is a word that frequently co-occurs with another word in a way that is statistically significant. For example, a strong collocate of *fish* in British English is *chips* (Bailey 2019). *Harm* and *risk* were selected as they appeared the pre-selected legal term list. We undertook collocation analysis to complement concordance analysis. The latter is the most qualitative part of corpus linguistics, relying on the analyst to recognise linguistic patterns and explain their meanings and functions (Bailey 2019: 64). Collocation analysis highlights language usage that can be hard to pinpoint via (legal) intuition alone and is replicable, with the benefit of providing different analysts with the same, reliable results in a way that concordance analysis may not (Xiao 2015).

Inductive data-driven approach

The data-driven approach is more suited to furthering general understanding of this little understood genre by testing the ability of the corpus to respond to purely linguistic inquiries which assume no prior legal knowledge of the PRC-25. This analysis was

conducted using a keyword analysis of the PRC-25, using the BNC Baby 2014 as a reference corpus. Keyword analysis was also combined with hierarchical clustering analysis which visualised the data in a dendrogram. Together they offer an insight into the major themes and topics of the PRC-25.

Keyword analysis is a core corpus linguistic tool that is widely considered to provide a useful starting point for exploring specialised corpora Evison (2010). It provides a data-driven method for approaching corpora in a principled way, based on statistical saliency rather than following researcher intuition (legal or otherwise) alone (Charteris-Black and Seale 2010). Keywords are words that appear significantly more frequently in a specialised corpus, such as the PRC-25, than would be expected when compared with general language use, which is represented by a reference corpus (Baker 2006; Stubbs 2010). As such, keywords point to the “aboutness of a text or a corpus” (Scott and Tribble 2006) helping to identifying a corpora’s major themes (Scott 2010), and highlight other linguistic features worthy of further analysis (Bondi 2010). Keyword analysis, therefore, provides an ideal avenue for a data-driven approach to the PRC-25.

Using the corpus software #LancsBox, we have defined a keyword as a word that occurs at least 5 times in the PRC-25 and has a log ratio statistic above 8. Log-ratio, an effect-size statistic that emphasises absolute frequencies, indicates the relationship between word occurrences in a specialised corpus and a reference corpus (Pojanapunya and Todd 2016). A log-ratio was used as an effect-size statistic the PRC-25 is a small pilot corpus of the much larger database of reports. As such, the PRC-25 cannot be treated as representative of all reports. The log-ratio calculation generated an initial list of 223 key words. For the purposes of this first empirical exploration of the PRC-25, which aimed to determine the corpus’ main themes, it was decided to focus primarily on lexical keywords. Consequently, one mathematical symbol was removed, as were functional keywords and contractions (e.g. *it’s*), and 88 proper nouns to leave a final keyword list of 122 keywords. Concordance lines of each keyword were examined to determine their general usage in the PRC-25 and to group keywords into semantic categories. This process of semantic categorisation, explained in the findings section below, highlights the major themes of the PRC-25 (Scott 2010).

A data visualisation tool, the hierarchical cluster analysis, was used to generate a dendrogram, carried out to consider topics as a discourse category as they appeared across the whole of the PRC-25. These clusters, generated by IRaMuTeQ, highlight the semantic characteristics of prevalent topics in a corpus. This approach is based on Van Dijk (2000)’s consideration for semantic study of text. Topics, defined as semantic macrostructures, regulate the overall coherence of discourse, representing what speakers find most important (Van Dijk 2000: 234). There are many different types of cluster analysis (Everitt *et al.* 2011) but for CL analysis a hierarchical agglomerative cluster analysis is typically used (Gries 2013: 336ff). For this analysis, individual data points were extracted by the IRaMuTeQ procedure and joined together to create larger clusters. These clusters contained all the data points, displayed in a tree diagram (dendrogram) in Figure 2 (Brezina 2018).

The dendrogram enabled exploration of the PRC-25 by highlighting its internal characteristics, in comparison to the keyword analysis which compared the PRC-25 corpus to a reference corpus. The rationale of this exploratory and free assumption visual method is predicated on revealing themes that tend to be found in similar contexts

(Lebart *et al.* 1997). This analysis complemented the keyword analysis by identifying and classifying additional major thematic groups found across the reports where the clusters are likened to topics in the sense that they highlight the internal semantic organization of the corpus.

Since the role of psychologists in these specific cases is at the intersection of psychology and law, the identification of lexical fields could shed light on the content of these reports and, in particular, topics that psychologists find important. According to Van Dijk (2000: 234), topics are “[d]efined as semantic macrostructures and [...] represent what speakers find most important, they regulate overall coherence of discourse, how discourse is planned and globally controlled and understood, and what is best remembered by the recipients”. The hierarchical classification extracted “context units”, or “text segments”, from a lemmatized version of the corpus, each segment containing 40 occurrences/forms. Lemmatization is a process of regrouping words related to a same stem into a single lemma (Brezina 2018: 40). The principle of the procedure implies that the tool cuts out the context units, or text segments, from the lemmatized corpus, removing the inflectional word endings. The PRC-25 was lemmatized only for this clustering procedure step. Each text segment contained 40 occurrences/forms which were clustered according to their vocabularies and distributed according to the reduced form frequencies, bringing together forms found in similar contexts (Lebart *et al.* 1997) before being divided successively into clusters (dichotomously at each step). Thus, the more times that the number of common forms to two given segments is high, the more these two segments are considered close and likely to be grouped together in a same cluster.

Results: synergies and separations

Concordance analysis: pre-selected legal terms

The expression of the pre-selected legal terms in the concordance lines revealed that they were frequently used out of their correct legal context and were either not understood in the reports to have specific legal meaning or were used to imply and comment on a legal meaning, again out of context. The following extracts illustrate sample concordance lines with a brief narrative explanation:

Extract 1:

*Although I am very clear in my mind that it is not in [the child's] **best interests** to return to her parents, there is no easy solution to [the child's] alternative placement*

*best interest** (used 20 times in 11 of the reports) was used to support the expert's opinion, although *interests* rather than *best interests* is the appropriate legal test unless there is a conflict of interests between a child(ren) and others (HM Government, 2018). In the full reports, however, the use of *best interests* did not correlate with the required conflict of interest criteria and may have been used out of context.

Extract 2:

*The Local Authority had concerns that [the child] would be at risk of **significant harm** due to the controlling nature and domestic violence they believed to be a feature of [the mother's] relationship with [the boyfriend]*

Extract 3:

*The Local Authority believes that the children have suffered **significant harm** due to being inadequately parented. It is also the Local Authority's position that [the mother] has failed to meet the*

*significant harm** appeared with contextual accuracy but was in only 2/25 reports, which is surprising as it is the primary threshold legal test. Example 2 aligns with the legal term *risk of significant harm*. Example 3 also uses a legal term *suffered significant harm* but is expressed with a non-legal term *have suffered*. The legal test uses the present tense, *is suffering*, or future tense *is likely to suffer*. Although contextually relevant to the court, the past tense should not influence the court in its decision making unless there is additional direct evidence of present or future significant harm.

Extract 4:

*The HCR-20 is a standardised tool which has been shown in research to be an effective method of assessing the **risk of future violence**. However, it does not have a significant number of normed scores available; and therefore, to some extent, the*

Extract 5:

*[the father] in my opinion presents a low to medium **risk of engaging in future acts of violence or inappropriate behaviours** which may directly affect any child in his care*

*risk of** was frequently used (75 times in 14/25 reports) and reflects the risk assessment objective of the psychologists in conducting the reports. It is surprising that it does not appear in more reports as it is the primary focus of the psychologists for their reports. It was also of note in the reports that the way in which *risk of* is used does not give the court much assistance in how precisely the risk is quantified or precisely which research is drawn on to support findings.

Extract 6:

*If an individual were to score highly on either or both measures, compared with the normative sample, it is **reasonable** to conclude that the subject is dissimulating (faking good or lying).*

*reasonable** was used 33 times in 12/15 reports and was used to add weight to the psychologists' conclusions rather than in its legal context.

Extract 7:

*This history of sexual victimisation is **a concern**. Although the general research may or may not apply to her, as an individual,*

a concern was used 6 times in 6/25 reports. In all instances it supported a negative conclusion about the parent(s), indicating the psychologist was worried about something. In Example 7, the worry was about the prior victimisation of the mother.

Extract 8:

*the reader that I am not an expert in 'attachment theory', yet I have been asked to state my professional **opinion** on this issue. As a general rule, I think it is reasonable to conclude that a child's permanency of care*

opinion was used 159 times in 22/25 reports. Experts are the only witnesses in court allowed to give their opinion, based on facts within their knowledge and expertise. In this example, the expert has been asked to comment outside their expertise and is potentially in breach of their duty to the court, although to their credit they highlight the issue in the text.

Identifying the concordances of the selected legal terms revealed the immediate contextual use of the terms across the PRC-25. The following collocation analysis identified the terms with which *risk* and *harm* co-occur.

Collocation analysis: risk and harm

Table 3 and Table 4 present the collocates of *risk* and *harm* in the PRC-25 respectively. The collocates in the tables have been generated using Mutual Information, an effect size measure commonly used to calculate collocation in corpus research. A possible limitation of MI against other collocation statistics is that it can tend to prioritise low frequency items (Baker 2006). To negate this, a minimum frequency threshold of 10 and a MI score of 5 for *risk* and 3 for *harm* was set. For both terms the calculation took into consideration words 5 to the left and 5 to the right of the search word.

Index	Position	Collocate	Stat (MI)	Freq (coll.)	Freq (corpus)
1	L	pose	9.446482	13	17
2	R	offending	8.779058	13	27
3	L	increased	8.511579	12	30
4	R	factors	8.166082	17	54
5	R	name*	7.87931	16	62
6	R	harm	7.511579	20	100
7	R	harming	7.485583	11	56
8	L	associated	7.460048	11	57
9	R	sexual	6.887546	34	262
10	L	low	6.648395	21	191
11	R	future	6.485583	11	112
12	R	name*	6.203029	13	161
13	R	self	5.919762	14	211
14	M	significant	5.789112	12	198
15	L	level	5.572979	12	230
16	R	assessment	5.355829	25	557
17	L	name*	5.093266	11	294
18	L	any	5.058406	21	575

Table 3. *risk* collocations in PRC-25

There are three quantifiers of risk in Table 3: *low*, *level*, and *significant*, which may help the court to assess the harm level perceived by the psychologist if these levels are able to be quantified. Quantifying human risk of future behaviour is challenging; in previous linguistic studies concerning risk and quantification, risk was used non-quantitatively (Boholm 2018). The expected collocate for *harm* would be *significant*, based on the legal threshold criterion in the primary legislation *significant harm* (Children Act 1989, Section 47) but this does not appear in the table. This observation merits further

Index	Position	Collocate	Stat (MI)	Freq (coll.)	Freq (corpus)
1	L	deliberate	10.04805	21	46
2	L	self	8.981756	46	211
3	L	physical	8.09183	12	102
4	L	risk	7.512537	20	254
5	L	emotional	7.463799	22	289
6	L	sexual	6.605302	11	262
7	L	from	5.529375	19	954
8	L	of	4.494243	58	5968
9	R	in	3.683248	21	3791
10	R	the	3.523243	43	8673
11	R	is	3.375808	10	2234
12	R	with	3.3038	12	2818
13	L	and	3.277512	33	7892
14	R	a	3.225097	15	3720
15	L	to	3.144219	30	7869

Table 4. *harm* collocations in PRC-25

examination in the context of whether the language used in the experts' reports is sufficiently clear and quantifiable. Legal language is often left deliberately vague and open to interpretation, as terms such as *significant harm* and *risk of* illustrate.

Keyword analysis: semantic categories

The keyword analysis and semantic categorisation of the keywords in the PRC-25 identified eleven main semantic categories (excluding "Other" below), some of which contain sub-groups (Table 5). The keywords in each category are listed with the highest keyness score first and the lowest last. Keywords marked with * appear in more than one semantic category (for instance, *self-worth* is both the focus of a some of the psychological tests that are explained, and it is also discussed in relation to the evaluation of an individual).

These semantic categories provide an overview of the main themes and concerns of the PRC-25. Firstly, the largest category (containing 88 keywords) was Proper nouns. Manually anonymising (replacing names and places with placeholders) the PRC-25 in full before running initial corpus procedures presented an extremely labour-intensive and time-consuming task. The keyword analysis was therefore run using the original reports, permanently redacting all identifying information retrospectively. Unsurprisingly, the names of individuals, organisations, and places were shown to be key in the PRC-25 as they were very unlikely to appear in the reference corpus, the BNC Baby 2014. This does, however, confirm the assumption that individual behaviours, rather than situations giving rise to child protection cases, are a major concern of these reports. This was checked by reading the full reports which confirmed the focus was generally on parent and child behaviours, even where concerns had been noted or raised about the general circumstances of the family, or poor professional responses, which were generally dismissed as irrelevant and not pursued. This is consistent with the role of the parent as respondent in the case. This is reflected in the related semantic category, Generic possessives (see also the Relations subcategory in Other). Examples of the

Semantic Categories	(Subgroups)	Keywords
Proper nouns (names and places)	-Redacted-	-Redacted-
Psychological testing	Test types (technical detail)	Wechsler, WASI, I.Q., CTQ, Achenbach, PDS, Tully, BDI-II, Bene, DSM-IV, MCMI, IIP, Millon, Centile, PSS, CISS, MCMI-III, Paulhus, Hickman, IQ, Visuo, CSRPI, DAP-IQ, STEEM, Maccoby, IASC, B/G, ASI, BDI, ABAS, Trichotech*, BAI
	Characteristic of test	self-reported, self-report, well-established, non-offending
	Focus of test	self-care, self-deception*, self-awareness*, self-direction, self-sacrificing, self-worth*, self-concept*, non-verbal, non-support, relatedness*, psychopathology*, malingering*
	Measure types	sub-tests, subscale, subtest, subscales, subtests, sub-test
Evaluation of subject(s) (psychological and behavioural description)		avoidant, symptomatology*, parented, life-story*, pre-contemplation*, psychopathology*, absconding, recidivism, self-harm, impulsivity, cooperative, minimisation, self-worth*, self-concept, relatedness*, self-sacrificing, malingering*, self-regulation, nurturance, self-deception*, self-awareness*
Generic possessives		mother's, child's, father's, person's, parent's, individual's, authority's, mum's, client's, grandmother's
Negative verbs		didn't, don't, doesn't, can't, wasn't, won't, hasn't, couldn't, wouldn't
Psychologist's credentials		Psychol, DFES, BPS British Psychological Society (2016), Clin., Solon, Psychopathology*, Symptomatology*, Multiagency
Inter-/intra-report references (i.e. cross-referencing, structure)	CAFCASS reports themselves	addendum, A/G, CAFCASS, parent-child*, pre-contemplation, affidavits
	Other relevant forms	C13, UAT, Trichotech*
Programmes		DAP (Integrated Domestic Abuse Programme), DBT (Dialectical Behavioural Therapy), Life-story*
Substances		citalopra, amphetamines
Parenting		parenting
Other	Relations	great-grandmother, great-grandfather, half-brother
	Dates	Dob, d.o.b., year's
	Miscellaneous	keyworkers, parent-child, reunification

Table 5. Semantic categories of lexical keywords in the PRC-25

excluded family circumstances included housing, health, poverty, parental vulnerability and/or victimisation and instances of reported professional failures to mitigate family needs.

The second semantic category, psychological testing, points to a second major concern of the reports. This category comprises mainly of the different tests that psychologists have used during their assessments. The sheer number of different tests named points to the non-standardisation of different psychologists' approaches. This reflects the fact that numerous tests are employed by psychologists when conducting clinical evaluations, but the statistical certainty of the outcome of the tests will not necessarily be understood by the courts (Law Commission 2011). As Extract 9 below illustrates, several reports do offer explanatory information about some of the tests used (as indicated by this semantic category's sub-categories, for example Characteristic of test, Focus of test, Measure types).

Extract 9:

[name] was psychometrically using the Wechsler Scale of Intelligence, designed for quick and accurately testing of an individual's intellectual functioning from the age of 6 onwards

Several of the psychological testing keywords, such as *relatedness* and *self-worth*, also appear in the group of keywords concerning the psychologists' actual evaluation of the subject/parent in the Report itself. Therefore, unsurprisingly, descriptions of individuals' behaviour (adjectives: *cooperative*, *self-sacrificing*, verbs: *absconding*), and psychological states or traits (adjectives: *avoidant*, nouns: *impulsivity*) is another major focus of the PRC-25. It is also relatively unsurprising that the keyword analysis highlighted a semantic category whereby psychologists outline their professional credentials. For instance, the keyword *BPS British Psychological Society (2016)*(n=6) appears when psychologists outline their qualifications:

Extract 10:

I am a member of the BPS Division and Neuropsychology and I have completed further training on the use of

Given that texts in PRC-25 take report form, and the complexity of related documents that psychologists may draw on in their assessments, the semantic category of Inter-/intra-report references, including report section titles and other documents is also unsurprising. Lastly, although the semantic category Parenting contains only one keyword, as would be expected in reports concerning parenting capacity, *parenting* (n=240) is a significant theme. Alphabetically sorted concordance lines immediately to the left and the right of the node *parenting* highlight several sub-themes of parenting referred to in the PRC-25. For instance, three sub-themes include parents' ability (i.e., *parenting ability*, *capacity*, *competence*, *skills*), measurable standards of parenting (i.e., *parenting assessment*, *performance*, *satisfaction*, *level*, *standard*, *basic*, *poor*, *positive*), and approaches to caring for children (i.e., *parenting behaviours*, *role*, *styles*). The sub-themes consist of subjective interpretations of behaviours and situations and their importance or otherwise. Future analysis should investigate the process by which the spoken and written materials from which the reports are constructed are entextualised into a re-constructed expert's narrative. From this initial analysis it was observed that the parent-respondents emphasised situations as being the issue of importance, whereas the psychologists emphasised behaviour as the primary focus.

Hierarchical clustering analysis: semantic categories

The results of the hierarchical clustering analysis match with some of the semantic categories in the keyword analysis. This especially applies to the psychological terminology category which matches with the institutional language cluster. Figure 2 shows that the hierarchical analysis revealed four clusters:

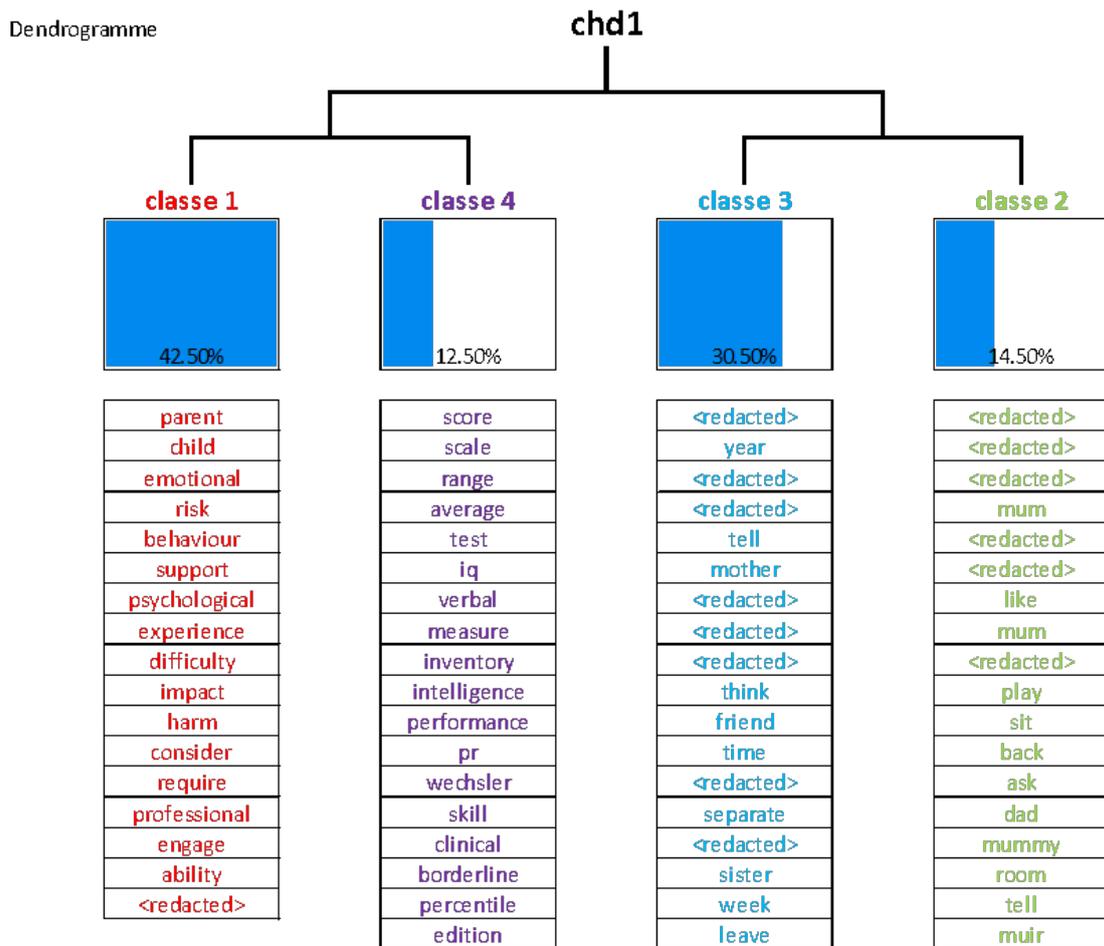


Figure 2. Hierarchical clusters in the PRC-25

Clusters 3 and 2 (blue and green respectively) represent one discourse type in the corpus. This discourse concerns the psychologists' reporting and reformulating the interviews with the parents and children in question and the contents of those interactions. Clusters 2 and 3 contain similar words, which will be explored in more detail in the larger study. As Figure 2 shows, these discourse clusters therefore comprise primarily of proper nouns, such as names and places (redacted here for anonymity), nouns relating to family members and familial relations (mother, dad, sister, friend), and a set of reporting clauses (tell, think, ask). As such, these clusters appear to represent those aspects of the interviews that the psychologist is observing and considers are worthy of reporting. Clusters 1 (red) and 4 (purple) represent a different type of discourse in the corpus: the professional and institutional-legal discourse used by the psychologists. From the small sample in the PRC-25 it was difficult to make general observations about whether they

were linked topically, semantically and/or organizationally in the reports themselves. The clusters did not directly correspond with sections of the reports, as the reports themselves did not follow a set structure and some did not contain clearly defined sections at all, as highlighted in the introduction where general observations were made about the whole reports within the PRC-25.

Cluster 4 (purple) groups the terms relating to quantitative assessment of psychological aptitudes and illustrates the scientific and statistical language (e.g., nouns such as score, measure, test, and percentile). It shows the specificity of this theme that dominates a wide part of the corpus and supports the results observed by the keyword analysis above. Finally, Cluster 1 represents a discourse of assessment developed in a legal and institutional language (e.g., nouns and adjectives such as *risk*, *psychological*, *impact*, *harm*, and verbs such as *consider*, *require*, *engage*).

The analysis included brief focus on the legal and institutional discourse identified by the hierarchical analysis. Table 6 shows the 20 first terms appearing in Cluster 1 (legal and institutional language) ranked on the basis of their Chi² score.

Term	Eff. s.t	Eff. total	Percentage	Chi2
parent	355	513	69.2	166.63
child	559	929	60.17	145.66
emotional	196	249	78.71	140.54
risk	136	155	87.74	133.9
behaviour	264	385	68.57	115.91
support	169	219	77.17	112.56
psychological	141	173	81.5	111.47
experience	173	233	74.25	100.72
difficulty	223	324	68.83	98.17
impact	84	90	93.33	96.85
harm	122	151	80.79	93.37
consider	151	201	75.12	91.14
require	85	96	88.54	84.85
professional	88	102	86.27	81.59
engage	117	149	78.52	81.49
ability	173	248	69.76	79.26
therapeutic	66	70	94.29	77.86
future	77	90	85.56	69.48
issue	114	151	75.5	69.31
attachment	97	123	78.86	68.17

Table 6. First 20 terms appearing in Cluster 1: legal and institutional language

The table shows also other values such as the number of text segments of this cluster containing the form at least once (Eff. s.t), the total number of text segments in the entire clustered corpus containing the form at least once (Eff. Total) and the percentage of Eff. ST compared to the Eff Total. The Chi² value expresses the level of intensity that links the term to the given cluster. It indicates the thematic affiliation of the terms and highlights the proximity and entwinement of the terms relating to psychological vocabulary and those relating to a legal vocabulary. A strong proximity

is observed between the psychologists' observation/conclusion vocabulary (e.g., *risk, ability, emotional, behaviour, experience, difficulty, impact*) and individual denomination (*parent(s), child(ren)*) or verbs (e.g., *to parent, to consider, to require*). Extracts 11 and 12 are two characteristic examples, meaning they have a high score, that contain the most frequent words associated with Class 1 such as *parent, behaviour, risk*, which are denoted in italics.

Extract 11:

in light of this history therefore child protection professionals have concerns that B may not have the ability to protect S from future sexual harm if M is considered to pose a potential risk in this regard

Extract 12:

given the level of concern regarding ST's parenting and associated concerns about her children's behaviour, alongside ST's difficulties engaging with professionals, I do not believe she has the ability now or in the foreseeable future to meet her children's care needs

In these two examples, the psychologists use logical connectors (e.g., *in light of, therefore, given the level of concerns*) to justify their arguments regarding causes and consequences or effects in these cases. In these utterances, they seek to justify their observations (*I do not believe*, Extract 12) and translate them into language intended to be accessible to the court.

Discussion

Both the legal-intuitive and the data-driven approach yielded observations made possible within a relatively short analytic time frame by drawing on corpus assisted methods. The analysis illustrates potential to significantly broaden and enrich the scope of legal analysis by supplementing qualitative analysis with quantitative measures, revealing patterns of language use that are hard to identify by (legal) intuition alone Xiao (2015) or that may belie expectations of a corpus. Searching for pre-selected terms, however, proved useful, subject to the noted methodological limitations, occurring at the intersection of quantitative and qualitative analysis.

The authors found the dual approach most useful in relation to the concordance analysis. Such analysis requires a significant input from the analyst, and in the absence of objectively correct practices (which do not exist for this type of data) the organisation and analysis of concordance lines are generally determined by the individual researcher. The legal-intuitive analysis brings value to concordance analysis via contextual legal knowledge, however, basing analysis solely on this knowledge may also equally eclipse other important and/or unexpected linguistic patterns relevant to the analysis of pre-selected legal terms. Using further methods of analysis such as collocation analysis deepens the enquiry. As well as bringing salient but hard to detect linguistic patterns into relief, collocation analysis is also replicable, with the benefit of providing different analysts with the same, reliable results in a way that concordance analysis may not. It presents an essential tool for the analysis of large legal corpora.

Overall, the selected legal terms were expressed less frequently than might reasonably be expected given their prominence in the primary and secondary legislation, and there was a general lack of clarity in the reports to demonstrate experts sufficiently

understand how best to use legal terms to assist the court. There was uncertainty in relation to the way in which they are used and/or omitted in the reports, leaving it unclear whether experts either deliberately use legal terms out of context or are unaware the terms they are using have a specific meaning when used in expert evidence. A data-driven approach would help to identify if, or which, terms are used instead of explicit legal terms by the psychologists and whether the report authors are discussing risk in more implicit ways, or not discussing risk at all.

The single most significant omission in relation to the legal terms is the lack of reference to the terms denoting the legal threshold justifying the legal proceedings, *significant harm*, which only appears in 2 of the 25 reports. This is surprising as it seems a reasonable assumption that salient, threshold legal terms would feature highly in the word frequency list as the purpose of the reports includes providing certainty to the courts (PD12A). The two instances can be contrasted with, for example, the use of *mother*, appearing 580 times in the PRC-25. Although specific analysis of gender bias (or otherwise) is outside the scope of this paper, emphasis on “mother” rather than “father” or “parent” reflects historic research findings arguing gender bias in child protection with over-focus on mothers (Appell 1997; Edleson 1998). Although other phrases can be used to convey the concepts of *significant* and *harm*, in legal materials the precise phrase is imperative and if the reports simply do not engage with this term it is hard to see how they are going to be useful to the courts.

There was also a widespread use of generally unquantified phrases throughout the PRC-25, reflected most starkly in the use of legal terms. Consequently, the terms used in the reports do not necessarily serve to clarify the issues at hand. Our analysis of *harm* and *risk* illustrate this; we found the terms were used without specific reference to how they should be quantified, which reflects their use in general English rather than specialised expert evidence (Boholm 2018). Although there were three statistically meaningful instances where *risk* was quantified (*low*, *level*, and *significant*), little information was provided as to how that assessment was made. In legislation, some terms are left deliberately vague, enabling judges to interpret them (Li 2017). If expert evidence is unable to establish the thresholds between, for example, *harm*, *significant harm*, *no risk*, *some risk*, or *high risk*, further research on this point is indicated. These are the very issues where the court needs most guidance, hence the experts' ability to comment on the ultimate issue is a unique position. If experts are using legal terms with inherent uncertainty such as *harm*, *risk*, *reasonable* or *significant*, particularly where they are paired with terms expressing uncertainty, they are unlikely to provide the certainty the courts may hope for.

Perhaps unsurprisingly, the analysis of the semantic categories shows an over-representation of the terms relating to the vocabulary of psychological evaluation, rather than legal certainty. This was illustrated in the hierarchical clustering analysis which highlighted the coexistence of four predominant themes in the corpus. Two of the clusters have similarities to those identified by the analysis of semantic categories (vocabulary of assessment, and that relating to proper names and family relations). Another identifies a more complex discourse which mixes both the evaluation discourse and that relating to its translation into a conclusion.

Three observations can be made about the way these psychological features of the Reports were constructed. Firstly, as noted, there was a focus on individual behaviours

and not family circumstance or the causes of those behaviours. As such, these were not problem-solving reports, which would be constructed with the aim of offering clinical input. Secondly, there was a high frequency of references to psychological testing without specific explanation to assist the court in interpreting the meaning of the tests and the results. Thirdly, in the reports, lengthy therapeutic intervention was a recommendation in many of them. Given the timeframe of the Public Law Outline is 26 weeks (albeit frequently exceeded in practice) the authors suggest this may create an impossible dilemma for judges. A question for future research may be whether a case will almost inevitably result in a child's removal in reports where a recommendation is made for therapeutic input beyond the 26-week timescale the courts can allow.

Two structural observations about the methodology of the reports emerged; firstly, there was a lack of standardization across the reports and there was a lack of transparency. In the small sample in the PRC-25 this posed problems for analysis of the reports as individual whole documents. Regardless of how the reports were organized, it was observed that there were frequent mentions of psychological tests, but there was no overt method of extracting how the conclusions in the reports had been drawn from the results of the tests, or discussion of their reliability. Instead, logical connectors were used to link events with narratives and the test results to generate the inferences and draw conclusions. Logical connectors also featured in what was observed to be a (re)construction of the lived experience of the subject(s) of the reports. Used to join two units of language, logical connectors attempt to justify the inferential links made in the reports. The marked use of logical connectors in the reports illustrates precisely why the reports may pose a problem for the court in interpreting their contents in the context of applying threshold tests to the appropriate standard of proof. To clarify this point, psychologists are permitted to comment on the ultimate issue. In the non-representative sample studied here, this did not happen, and the reports did not provide the level of clarity hoped for following the 2013 review Ministry of Justice and Family Justice Council.

This raises two structural issues for further examination: firstly, the question of focus on behaviours rather than situations merits further examination. Secondly, the process of (re)construction of the source material for the reports, including the verbal interviews, requires future analysis to investigate the process of (re)construction of the narratives and their entextualisation. Both these issues concern the level of clarity (or otherwise) in the protocol for reporting interviews, and how inferences can be drawn from them, including the methods of converting the spoken interviews into written data. Of note, from the reports, it is unclear how the experts have conducted the actual interview with the parents. No information is provided regarding the experts' preparation for the interview, how they carried out the interview itself (for example, following a predetermined set of questions), and significantly, how they went about making a record of the interview (for example, synchronously or asynchronously). It is known from linguistic research in forensic and legal contexts (Richardson *et al.* 2022; Andrus 2011; Bucholtz 2009; Fraser 2003) that interviews themselves as pieces of forensic evidence and the processes through which spoken discourse is captured in a written format is a fraught and complicated matter. In this case, the inferences that are made from these reported interviews may be particularly problematic in the conclusions that

psychologists make. This is certainly an aspect of the PRC-25 that requires further attention, combining linguistic and legal expertise.

Overall, the approaches used in this study were found to be complementary, providing overlapping findings which were triangulated to some extent. This demonstrates the synergy between the deductive and inductive elements of the study, confirming the initial hypothesis that this is a promising avenue for exploring expert evidence in a larger corpus. The combination of keyword analysis and hierarchical clustering analysis highlights important observations about salient language in the PRC-25, its main themes, and focusses, and areas for further analysis. Both approaches yielded promising results for the potential of corpus linguistics for developing a methodology to answer specific legal (and other) questions, as well as the potential to interrogate the data to generate new hypotheses. More can be established in terms of this type of report as a genre, including investigation of the register/genre distinctions, and the lexical bundles (phrases) highlighting conventions for the text type. Further detailed linguistic analysis of the reports on a representative sample using CL methods will provide clarity on both the issues raised by this study and will generate further insights into the language used in the reports. Further analysis of a representative sample of the reports will also illuminate the process of entextualisation behind their creation.

Conclusions

This paper demonstrates that the two approaches are complementary, providing overlapping findings, but also highlighting and mitigating mutual methodological and practical shortcomings which can be resolved during and after the creation of a larger and representative corpus. In combining the two approaches to maximise both the legal knowledge, which is vital to the contextualisation of linguistic data, and a principled linguistic approach to the analysis of large datasets, interesting indicative results were generated about the use of legal language in the PRC-25. This demonstrates the synergy and potential for corpus-assisted analysis of written and spoken legal materials.

The PRC-25 offers promising and indicative findings which justify scaling up to a larger study creating a large corpus of psychologists' reports and protocol for development across other areas of expert evidence in a wider range of legal proceedings. The synergy between the approaches is typified by the analytical journey of the terms *risk* and *harm* throughout this paper. Firstly, they were pre-selected legal terms of interest in the corpus. Concordance analysis enabled each use of these terms and what we already knew to be their related phrases (eg *risk of, significant harm*), amongst others, to be viewed in their original context. This provided a convenient platform for a legal-intuitive analysis to draw on. Secondly, collocation analysis also brought unexpected associations with these words into relief where they may have been otherwise missed by legal analysis alone. Lastly, keyword and hierarchical analysis designed to enable us to gain insights from the PRC-25 and its main themes and concerns, also identified *risk* and *harm* as salient terms in this corpus, thereby confirming our selection of them for more detailed analysis.

A limitation of previous research with expert evidence transcripts has been the difficulty of accessing the data. To overcome this limitation for future research, the next phase of this research is the preparation of a large and representative corpus for researcher re-use. Although it poses new challenges, the effort required for this

task is justified to enable a deeper and reliable consideration of this type of evidence. The raw reports in the database require considerable work to render them useable for research purposes and anonymisation of individual reports presents a labour-intensive task justifying an automated solution as part of the corpus creation process. A process such as encoding in XML-TEI is envisaged which will allow analysis with corpus tools which recognise this encoding (e.g Lungen (2017)). Despite these challenges, once created, the corpus can be re-used efficiently to answer a variety of research questions.

Acknowledgements

Enquiries to l.devine@aston.ac.uk. The authors would like to thank Ms Poppy Cooke for her careful work creating the database from which the PRC-25 was extracted. We would also like to thank HHJ Stephen Wildblood QC for his insightful observations directing us towards this field of research. This work was funded by the Nuffield Foundation under grant number JUS/43090 for the project *Theorising the Care Cases Crisis*, and *Mind your Language* funded by Research England under an Expanding Excellence in England (E3) Award. Also, thanks to the reviewers for their helpful comments on an earlier draft.

Notes

¹*Theorising the Care Cases Crisis* (Grant number JUS/43090) and *Mind your Language*, funded by Research England under an Expanding Excellence in England (E3) Award.

²The research has been approved by the President of the Family Division of England and Wales. Data has been accessed under a Cafcass Data Processing Agreement has received two ethical approvals from Cafcass and Aston University respectively.

References

- Aldridge, (2010). Vulnerable witnesses in the criminal justice system. In M. Coulthard and A. Johnson, Eds., *The Routledge Handbook in Applied Linguistics*. Abingdon. Routledge.
- Andrus, J. (2011). Beyond texts in context: Recontextualization and the co-production of texts and contexts in the legal discourse, excited utterance exception to hearsay. *Discourse & Society*, 22(2), 115–136.
- Appell, A. (1997). Protecting children or punishing mothers: Gender, race, and class in the child protection system [an essay], 48 s. *C. L. Rev.* 577.
- Bailey, A. (2019). *Discourses of dementia in the British press: a corpus linguistic approach*. University of Nottingham.
- Baker, P. (2006). *Using corpora in discourse analysis*. London: Continuum.
- Baker, P., Gabrielatos, C., Khosravini, M., Krzyzanowski, M., McEnery, T. and Wodak, R. (2008). A useful methodological synergy? combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the uk press. *Discourse and Society*, 19, 273–306.
- Biber, D., Conrad, S. and Reppen, R. (1998). *Corpus Linguistics. Investigating language and use*. Cambridge University Press.
- Boholm, M. (2018). Risk and quantification: a linguistic study. *Risk Analysis*, 29(ue 6), 1243–1261.
- Bondi, M. (2010). Perspectives on keywords and keyness. In M. Bondi and M. Scott, Eds., *Keyness in Texts*. Amsterdam: John Benjamins, 1–18.
- Brezina, V. (2018). *Statistics in Corpus Linguistics: A Practical Guide*. Cambridge: Cambridge University Press.

Devine, L. *et al.* - 'Psy' expert evidence in the family courts

Language and Law / Linguagem e Direito, Vol. 9(1), 2022, p. 92-119

- Brezina, V. and McEnery, T. (2015). Collocations in context. *International Journal of Corpus Linguistics*, 20(2), 139–173.
- Brezina, V., Weill-Tessier, P. and McEnery, T. (2020). Lancsbox 5.x and 6.x. <http://corpora.lancs.ac.uk/lancsbox>.
- British Psychological Society, (2016). Psychologists as expert witnesses in the family courts in England and Wales, standards, competencies, and expectations. *Leicester: The British Psychological Society and the Family Justice Council*.
- Broadhurst, K. and Mason, C. (2013a). Child removal as the gateway to further adversity: Birth mother accounts of the immediate and enduring collateral consequences of child removal. *Qualitative Social Work*, 19(1), 15–37.
- Broadhurst, K. and Mason, C. (2013b). Maternal outcasts: raising the profile of women who are vulnerable to successive, compulsory removals of their children – a plea for preventative action. *Journal of Social Welfare and Family Law*, 35(3), 291–304.
- Bucholtz, M. (2009). Captured on tape: Professional hearing and competing contextualizations in the criminal justice system. *Text & Talk*, 29, 503–23.
- CAFCASS, (2022). Website: Annual summaries, table of LA rates of care application per 10,000 population. tab: Public care application of excel spreadsheet. <https://www.cafcass.gov.uk/download/6250/>.
- Camasso, M. and Jagannathan, R. (2000). Modeling the reliability and predictive validity of risk assessment in child protective services. *Children and Youth Services Review*, 873–896.
- Charteris-Black, J. and Seale, C. (2010). *Gender and the language of illness*. Hampshire: Palgrave Macmillan.
- Cobley, C. and Lowe, N. (2009). Interpreting the threshold criteria under section 31(2) of the children act 1989.
- Coulthard, M., Johnson, A. and Wright, D. (2016). *An introduction to forensic linguistics: Language in evidence*. London: Routledge, 2nd ed.
- J. Cutting, Ed. (2007). *Vague Language Explored*. Basingstoke: Palgrave Macmillan.
- Devine, L. (2017). *The Limits of State Power and Private Rights: Exploring the limits of safeguarding referrals and assessments*. London: Routledge.
- Edleson, J. (1998). Responsible mothers and invisible men: Child protection in the case of adult domestic violence. *Journal of Interpersonal Violence*, 13(2), 294–298.
- Egbert, J. and Baker, P. (2020). *Using Corpus Methods to Triangulate Linguistic Analysis*. New York: Routledge.
- Everitt, B., Landau, S., Leese, M. and Stahl, D. (2011). *Cluster Analysis*. Wiley, 5th ed.
- Evison, J. (2010). What are the basics of analysing a corpus? In A. O’Keeffe and M. McCarthy, Eds., *The Routledge Handbook of Corpus Linguistics*. Oxon: Routledge, 122–135.
- EWCA, (1996). Re M & R (Child Abuse: Evidence) [1996] EWCA Civ 1317; [1996] 2 FLR. 195.
- Fraser, H. (2003). Issues in transcription: Factors affecting the reliability of transcripts as evidence in legal cases. *International Journal of Speech, Language and the Law—Forensic Linguistics*, 10, 203–26.
- Gillingham, P. (2006). Risk assessment in child protection: Problem rather than solution? *Australian Social Work*, 59(1), 86–98.
- Gillings, M. (2022). How to use corpus linguistics in forensic linguistics. In A. O’Keeffe and M. McCarthy, Eds., *The Routledge Handbook of Corpus Linguistics*. London: Routledge, 2nd ed., 589–601.

- Goldberg, L. (1968). Simple models or simple processes? some research on clinical judgments. *American Psychologist*, 23(7), 483–496.
- Gries, S. (2013). 50-something years of work on collocations. what is or should be next? *International Journal of Corpus Linguistics*, 18(ue 1), 137–166.
- HM Government, (2018). Working together to safeguard children. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/942454/Working_together_to_safeguard_children_inter_agency_guidance.pdf.
- Hunston, S. (2010). How can a corpus be used to explore patterns? In A. O’Keeffe and M. McCarthy, Eds., *The Routledge Handbook of Corpus Linguistics*. Oxon: Routledge, 152–166.
- Ireland, J. (2012). *Evaluating Expert Witness Psychological Reports: Exploring Quality*. University of Central Lancashire. <http://www.fixcas.com/news/2012/IrelandJane.pdf>.
- KB, (1943). *Hollington v hewthorn* [1943] kb 587, para 595.
- Law Commission, (2009). The admissibility of expert evidence in criminal proceedings in England and Wales, cp190. https://s3-eu-west-2.amazonaws.com/lawcom-prod-storage-11jsxou24uy7q/uploads/2015/03/cp190_Expert_Evidence_Consultation.pdf.
- Law Commission, (2011). Expert evidence in criminal proceedings report, LC325. https://s3-eu-west-2.amazonaws.com/lawcom-prod-storage-11jsxou24uy7q/uploads/2015/03/lc325_Expert_Evidence_Report.pdf.
- Lebart, L., Salem, A. and Berry, L. (1997). *Exploring Textual Data*. Springer Science & Business Media.
- Li, S. (2017). A corpus-based study of vague language in legislative texts: Strategic use of vague terms. *English for Specific Purposes*, 45, 98–109.
- Lindley, B. (1994). *Families in Court: A qualitative study of families’ experiences of the court process in care and supervision proceedings under the Children Act 1989*. London: Family Rights Group.
- Loubère, L. and Ratinaud, P. (2014). Documentation iramuteq 0.6 alpha 3 version 0.1., iramuteq. http://iramuteq.org/documentation/fichiers/documentation_19_02_2014.pdf.
- Lüngen, H. (2017). Anonymisation of the Dortmund chat corpus 2.1. In E. Stemle and C. Wigham, Eds., *Proceedings of the 5th Conference on CMC and Social Media Corpora for the Humanities (cmccorpora17)*, 21–24, https://ids-pub.bsz-bw.de/frontdoor/deliver/index/docId/6639/file/Luengen_Beisswenger_Herzberg_Pichler_Anonymisation_2017.pdf: Eurac Research.
- Marchi, A. (2022). Corpus linguistics in the study of news media. In A. O’Keeffe and M. McCarthy, Eds., *The Routledge Handbook of Corpus Linguistics*. London: Routledge, 2nd ed., 576–588.
- Ministry of Justice, (2020). Family procedure rules. part 25. experts and assessors. https://www.justice.gov.uk/courts/procedure-rules/family/parts/part_25. Last updated 6th.
- Ministry of Justice, (2021a). Practice direction 12a - care, supervision and other part 4 proceedings: Guide to case management, last updated.
- Ministry of Justice, (2021b). Practice direction 25a – interim injunctions. https://www.justice.gov.uk/courts/procedure-rules/civil/rules/part25/pd_part25a.
- Ministry of Justice and Family Justice Council, (2013). *Response: Standards for expert witnesses in Children’s Proceedings in the family courts*. London: HM Government.
- Munby, S. J. (2011). President of the family division. <https://www.judiciary.uk/announcements/tighter-rules-introduced-on-expert-evidence-family-cases/>.

- Munro, E. (2010). Conflating risks: Implications for accurate risk prediction in child welfare services. *Health, Risk & Society*, 12(2), 119–130.
- Møller Jørgensen, A., Webb, C., Keddell, E. and Ballantyne, N. (2021). Three roads to Rome? comparative policy analysis of predictive tools in child protection services in aotearoa new zealand, england, & denmark, *nordic social work research*.
- Partington, A. (1998). *Patterns and meanings*. Amsterdam: John Benjamins.
- Parton, P. (2011). Child protection and safeguarding in England: Changing and competing conceptions of risk and their implications for social work. *The British Journal of Social Work*, 41(ue 5), 854–875.
- Pojanapunya, P. and Todd, R. (2016). Log-likelihood and odds ratio: Keynes statistics for different purposes of keyword analysis'. In *Corpus Linguistics and Linguistic Theory [Preprint]*.
- Rayson, P. (2015). Computational tools and methods for corpus compilation and analysis. In D. Biber and R. Reppen, Eds., *The Cambridge Handbook of Corpus Linguistics*. Cambridge: Cambridge University Press, 32–49.
- Redhill, J. and Roe, A. (2021). What do we know about children in the family justice system? nuffield family justice observatory. https://www.nuffieldfjo.org.uk/wp-content/uploads/2021/09/nfjo_infoG_explanatory_note_2021206.pdf.
- Regehr, C., Bogo, M., Shlonsky, A. and LeBlanc, V. (2010). Confidence and professional judgment in assessing children's risk of abuse. *Research on Social Work Practice*, 20(6), 621–628.
- Richardson, E., Haworth, K. and Deamer, F. (2022). For the record: Questioning transcription processes in legal contexts. *Applied Linguistics*, 1–22.
- Scott, M. (2010). Problems in investigating keyness, or clearing the undergrowth and marking out trails. In M. Bondi and M. Scott, Eds., *Keyness in Texts*. Amsterdam: John Benjamins, 43–58.
- Scott, M. and Tribble, C. (2006). *Textual Patterns: Key words and corpus analysis in language education*. Amsterdam: John Benjamins.
- Sinclair, J. (1999). Concordance tasks. the tuscan word centre. <http://www.twc.it/happen.html>.
- Smeeton, J. (2020). A murky business": A phenomenological ontology of risk in child protection social work. *Qualitative Social Work*, 19(2), 284–300.
- Stubbs, M. (2010). Three concepts of keywords. In M. Bondi and M. Scott, Eds., *Keyness in Texts*. Amsterdam: John Benjamins, 19–42.
- Taylor, C. (2022). Corpus linguistics in the study of political discourse: recent directions. In A. O'Keeffe and M. McCarthy, Eds., *The Routledge Handbook of Corpus Linguistics*. London: Routledge, 2nd ed., 602–614.
- Tkacukova, T. (2020). Changing landscape of advice provision: Online forums and social media run by mckenzie friends. *Child and Family Law Quarterly*, 4, 397–420.
- UK Parliament, (1989). Children act 1989 c.41.
- UKSC, (2013). Re B (Care Proceedings: Appeal) [2013] UKSC 33.
- Van Dijk, T. (2000). On the analysis of parliamentary debates on immigration. In M. Reisigl and R. Wodak, Eds., *The semiotics of racism. Approaches to Critical Discourse Analysis*. 85–104.
- Xiao, R. (2015). Collocation. In D. Biber, Ed., *The Cambridge Handbook of English Corpus Linguistics*. Cambridge: Cambridge University Press, 106–124.