THE LANGUAGE OF SUICIDE NOTES

by

JESS JANN SHAPERO

A thesis submitted to
The University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

Department of English
The University of Birmingham
January 2011

UNIVERSITY OF BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

Abstract

This thesis reports a study of a corpus of 286 suicide notes collected from the Birmingham Coroner's Office, with additional findings from 33 real and 33 fabricated notes from Los Angeles.

Following some background regarding how suicide notes are treated by Coroners' Courts and other courts in the U.K, the thesis compares topics used in real and fabricated suicide notes. Although there is considerable overlap between the two categories, they can be partially distinguished by some features that are more likely to occur in one category than the other. For example, dates, indications of author identity and trivia are more likely to occur in real notes than fabricated ones.

The thesis then concentrates on fake notes and scrutinises instances of atypical language or phraseology and contextually inappropriate content. It is found that these oddities are far more frequent in the fake notes than in the genuine ones.

Finally the thesis focuses on the corpus of genuine notes from the Birmingham Coroner's Office, using an automatic semantic tagger. The findings are that suicide notes contain significant proportions of items indicating affection, the future and their authors' kin. In addition, the notes include significant proportions of pronouns, names, negatives, intensifiers, maximum quantity terms, and discourse markers.

Dedication

To my mother, Rivelyn Shapero, 8th July, 1922 - 31st May, 2008.

And to my friends.

Acknowledgements

Firstly I wish to express my thanks and appreciation to Susan Hunston, my supervisor, without whom this thesis could not have been completed. I also want to express my thanks and appreciation to Malcolm Coulthard, my original supervisor, without whom this thesis would not have been started. Thanks are also owing to Marilyn Washbrook and to Michelle Cullen for their help with various administrative concerns over the years.

My thanks also go to the following people who helped me in various ways, and to various extents: Dr Richard Michael Whittington (The Coroner), Jenny Goodwin, Lynn Stead, Chris Pushon, Delphine Guest, and all the other staff at the Coroner's Office, Newton Street, Birmingham, B4-6NE for allowing me to see their database and their files; Rivelyn Shapero F.S.B.T. (Fellow of the Society of Business Teachers) for translating some Pitman's shorthand; and Ramesh Krishnamurthy, for his help with corpora size issues. And, for their various contributions, I want to thank my personal communicants who include: Carl Chinn - from Birmingham University; Paul Matthews – from King's College London; and, of course, Susan Hunston and Malcolm Coulthard. I apologise to anyone I may have, inadvertently, failed to mention in these acknowledgements.

This thesis had a previous life. It took on its present form after I changed supervisors. The original thesis was quite different, and during its time I was aided and abetted by various people whom I promised to mention in my acknowledgements. I now keep those promises. However, I shall not mention any specific contributions because they are no longer relevant to the thesis in its present form, and so might lead to confusion.

I thank Lexie Don, Oliver Mason, Antoon Leenaars, Edwin Shneidman, Someya Yasumasa, Waterstones bookshop, and David Campbell of the L.A. County Department of Coroner Operations Bureau. I also thank my personal communicants who include: John Olsson; Steve Thorne and Marion Thain - both from Birmingham University; Malcolm Ingram; and Stephanie Ernst. Again, I apologise to anyone I have left out.

The Language of Suicide Notes

Contents	Page

List of Figures

List of Tables

1. INTRODUCTION

In this chapter I begin by discussing the purpose of this thesis. I explain both my original aims and why they were thwarted, and my revised, current aims. I then indicate the importance of my work, and pose some questions for the forensic linguistics and corpus linguistics communities. Finally, I introduce my research questions, and then outline the other chapters in this thesis.

My interest in the subject of suicide notes arose when I became aware of cases where forensic linguists were discussing whether a suicide note was genuine or falsified. My intention in planning this thesis, then, was to look at the authorship of suicide notes, and to describe suicide notes in such a way that genuine ones could be distinguished from fake ones.

An example of such a case, with a suicide note whose authenticity was disputed, is R v. Gilfoyle (2000). On June 4, 1992 Paula Gilfoyle was found hanged in her garage; and a suicide note was discovered. A few days later, Paula's friends told the police that she had told them that her husband, Eddie, had persuaded her to write suicide notes in order to help him with his hospital course-work. Consequently, Eddie was arrested, tried, and found guilty of killing Paula. At the time of writing, several appeals later, he is still in prison. So, according to the courts, Mr Gilfoyle persuaded his pregnant wife to write a suicide note and then killed her.

In addition to the suspect suicide note, several other texts were found, notably a letter which was apparently from Paula to Eddie admitting adultery, another letter ostensibly from Paula to Eddie which was recovered by the ESDA[1] process from its imprint on a notepad, and a handwritten note found "in a foot stool in the kitchen" (R. v. Gilfoyle, 2000, para.7). In the following extract from the Appeal Court hearing of R v. Gilfoyle I have made some words bold to emphasise the number of different texts with which the

---

[1] ESDA (Electro-Static Detection Apparatus) (described in Davis (1994)) is a process that can reveal impressions on paper made by writing on another piece of paper on top of it, even where several other sheets were between them.

court was dealing.

> "[6]… A letter which he said his wife had written to him... was referred to at trial as **the 'Nigel' letter**... After the deceased's death, other letters were found in notebooks in the house. One **typed letter** had been written about the end of October 1991, a day or two after the appellant had told his wife he had someone else... As a result of ESDA testing, another typed **'suicide' letter**, referred to as **the 'indented' letter** was revealed in a notebook. A handwriting expert said that there was   strong evidence that it had been written before March 1992, when some domestic accounts had been written in the same book... **Another note**, of unknown date, hand-written, and addressed 'To whom it may concern', was found in a foot stool in the kitchen. It said 'I Paula Gilfoyle am ending my life. I have taken my own life...' (R v. Gilfoyle, 2000, para.6).

Although the Gilfoyle note was deemed by the courts to be a simulated rather than genuine suicide note, it is still a subject of debate and disagreement amongst forensic linguists. This case is itself evidence that supports the premise that the study of suicide notes is worthwhile and important.

The original purpose of the thesis necessarily had to change to some extent because of the key problem that is the difficulty of obtaining a substantial corpus of non-authentic notes (see Chapters 3 and 4). Although I succeeded in acquiring a substantial corpus of genuine notes, the ethical and practical considerations in persuading and preparing people to write 'fake' notes are such that it was beyond the limitations of this study, and my thesis was therefore reconfigured to give a more detailed description of the notes themselves, comparing with simulated notes where this was possible, but otherwise independent from the 'real versus fake' issue. Further, I became more interested in defining the typical suicide note and less concerned with establishing the borderline between the fake and the real.

Thus, the main aim of this thesis is to investigate the language of suicide notes, and to find out what kinds of things suicidal people see as important to write about. I also want to establish some of the distinguishing features of genuine notes (where possible).

In this thesis I use two corpora of suicide notes which I refer to as the Birmingham corpus and the Shneidman corpus (see Chapter 4), and now I shall briefly elaborate on

them. I have collected a corpus of 286 suicide notes which I obtained from the Birmingham Coroner's Office (see Chapter 4). I refer to these as either the Birmingham corpus or the Birmingham notes. I divide these notes into two sub-corpora, one being 74 notes written by females and the other being 212 notes written by males. I also use a corpus of 66 American suicide notes built by Edwin Shneidman (Shneidman & Farberow, 1957e). The Shneidman corpus consists of two sub-corpora, one of which contains 33 genuine suicide notes, and one of which contains 33 simulated suicide notes (see Chapter 3).

For the forensic linguist the importance of my work is that I am able to establish what a typical suicide note looks like. I have also posed some questions with respect to forensic linguistics. Where I have found that the Birmingham notes are radically different from Shneidman's genuine notes I have raised questions about the consistency of texts across time (when they were written) and across space (where they were written) (see Chapter 5), and I have also raised questions about looking at small numbers of texts and words and looking at large numbers of them.

For the corpus linguist, first and foremost, I show an important application of corpus analysis: namely, its application to forensic cases. I also raise questions about how to deal with categories that are quite difficult to establish: for example, how to establish topic in a corpus. This is particularly tricky when the corpus is, on the one hand, entirely homogeneous, because the constituent texts are all suicide notes from the same city, yet on the other hand, the corpus is produced by writers who are very different from each other, in very different circumstances, and who are writing privately for private consumption. Some of the variation I found within my corpus is because of this dichotomy or juxtaposition. The note-writers are not people who are writing publicly in an established genre. They are people who are writing privately, and they may never have read a suicide note before. Swales (2007), referring to his 1996 paper in which he apparently coined the term, calls this kind of non-established genre "occluded". Occluded genres are "hidden and out of sight to all but a privileged few" (Swales, 2007, p.149), and so people producing texts in these genres seldom have access to example texts to guide them.

A further point is that the notes vary considerably in length. So my questions for corpus linguists essentially concern what techniques can be used in analysing this particular kind of corpus. Another point that should be of concern to both the forensic and the corpus linguistic communities is the importance of looking at phenomena in corpora on the basis of how diversely they are spread across the component texts (to avoid results that are skewed by occurring only in a few of the texts).

I now present some research questions prior to previewing the other chapters of the thesis. As mentioned above, in this thesis I use two corpora of suicide notes which I refer to as the Birmingham corpus and the Shneidman corpus. My main research questions are:

1. What factors are involved in building a corpus of suicide notes?

2. What are the characteristics of the Birmingham corpus (from a mainly qualitative viewpoint)?

3. Are there ways to distinguish between the Shneidman sub-corpora of genuine and simulated suicide notes, in addition to those used by other researchers?

4. What are suicide notes about (from a mainly quantitative point of view)?

5. Is the Birmingham corpus demonstrably more like the Shneidman genuine sub-corpus or the Shneidman simulated sub-corpus?

How I answer the above questions is outlined as follows.

1. To see what factors are involved in building a corpus of suicide notes I describe how I built my corpus and I discuss the difficulties involved in doing so (see Chapter 4).
2. To see what the characteristics of the Birmingham corpus are (from a mainly

qualitative viewpoint) I manually examine the corpus to see what topics appear to be prevalent in it, and then manually count the number of suicide notes that actually contain these topics (see Chapter 5).

3. To see whether there are ways to distinguish between the Shneidman sub-corpora of genuine and simulated suicide notes, in addition to those used by other researchers, I concentrate on the simulated suicide notes and look for features that they have in common with each other (see Chapter 5). The features are considered in the light of what some other researchers have discovered.

4. To see what suicide notes are about (from a mainly quantitative point of view) I look at some wordlists (see Chapter 6), and then use the WMatrix software application to find topics that are more quantitatively defined than those mentioned in point 2 above (see Chapter 7).

5. To see whether the Birmingham corpus is demonstrably more like the Shneidman genuine sub-corpus or the Shneidman simulated sub-corpus, in Chapter 5 I compare the results of various studies across the sub-corpora (see also Chapter 8).

The above list is a somewhat simplified view of how the research questions are answered. Some of the answers are spread over more than a single chapter. The following is a brief set of descriptions (including a brief reprise of some of the above) indicating what the various chapters of the thesis contain, and where the reader can find the answers to the research questions.

In Chapter 2 I give some legal background to the subject of suicide notes. I endeavour to define both suicide and suicide notes, and I look at how the Coroner's Court works when suicide notes are involved. I also present a short study of some legal cases to show how courts (other than Coroners' Courts) are influenced by the presence or absence of suicide notes.

In Chapter 3 I review some of the relevant literature on suicide notes, and consider the difficulties in collecting a corpus of fake, or simulated, suicide notes. As Chapter 3 will show, other researchers who have examined these notes have compared them against

various other corpora, considered them at various levels of discourse (from lexis upwards), and looked at them in terms of numerous categories, types of categories and mixtures of categories. Some have worked manually, some with computers. It would appear that little more could be done with the notes. However, (as Chapter 3 will also show) most studies are from researchers based in medical institutions or psychology departments and there is, at least, an opportunity for an approach that is from a linguistic perspective.

In Chapter 4 I present my data and describe the problems in acquiring and transcribing it. Chapter 5 presents a two-pronged linguistic approach to the problem of distinguishing falsified from genuine suicide notes. Firstly, an outline of the anatomy of a typical suicide text is derived from a set of genuine suicide notes, and comparisons are made with some fabricated suicide notes. Secondly, a set of simulated suicide notes is analysed, and the results of this are applied to some genuine notes. Thus, in Chapter 5, looking first at real suicide notes is the first prong in my "two-pronged" approach, and looking first at fake suicide notes is the second prong. In Chapter 6 I use word and keyword lists to see what suicide notes are about. In Chapter 7 I use some semantic tagging software, namely Wmatrix2 (Rayson, 2007), to further investigate what suicide notes are about, and what distinguishes them from non-suicide texts. Finally, in Chapter 8, I present my conclusions to this thesis.

Throughout this thesis quotations and extracts from suicide notes have kept their original spellings, indicated with "sic" where necessary. I have also endeavoured to represent the original spacing and layout of suicide notes, and their extracts, where this is appropriate and/or relevant.

The citation references used in this thesis for individual suicide notes, and extracts from suicide notes, from the Birmingham corpus are codes, which I have devised as follows (see Figure 1.1 below and Chapter 4, Section 4.5). "M" or "F" indicates whether the note's author was male or female, respectively, two numbers indicate the inquest year (which is usually the year in which the author died), a further three numbers are the inquest number, and a final alphabetic character, if any, is used to distinguish multiple

notes written by a single author. An example is "M98666b" as shown in Figure 1.1 below.
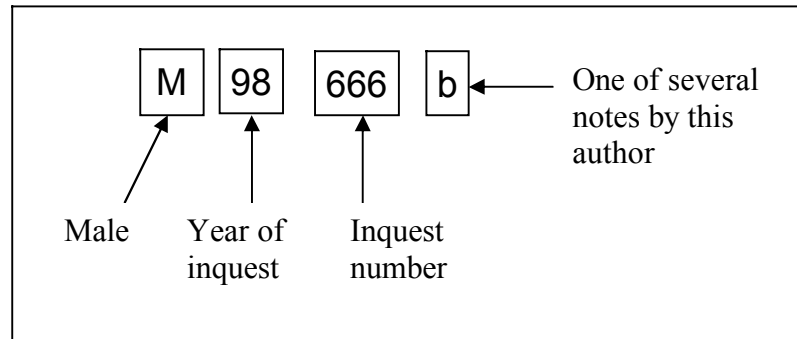
Figure 1.1. Birmingham Note Reference Example

The citation references used in this thesis for individual suicide notes, and extracts from suicide notes, from the Shneidman corpus are codes consisting of either "Sim" or "Gen", indicating the sub-corpus to which the text belongs (see Figure 1.2 below and Chapter 4, Section 4.7), and a two digit number followed by a single alphabetic letter replicating the codes given to the notes by Shneidman (Shneidman & Farberow, 1957e). An example is "Gen33a" as shown in Figure 1.2 below.
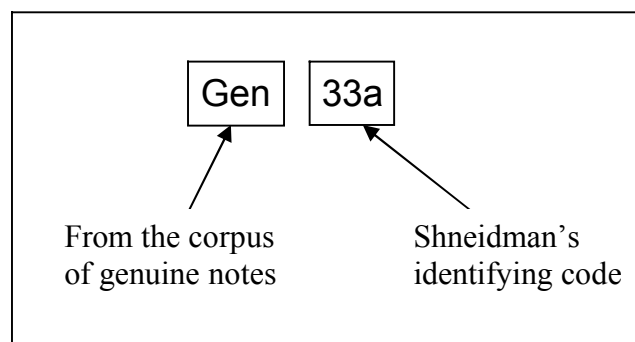
Figure 1.2. Shneidman Note Reference Example

2. LEGAL BACKGROUND

In this chapter I want to provide some background for my thesis, and so I look at how suicide notes are used by the courts. Courts, particularly Coroners' Courts, are usually where the official classification of 'suicide' is applied to the case of someone who has taken their own life, and so it seems appropriate to look at how this evolves. Firstly, I give a brief account of the legal processes involved in arriving at a verdict of suicide, and what a Coroner's Court does. This is a short account of the legal system in England and Wales (but not Scotland) from 2000 to 2009 as it relates to suicide, and how suicide notes are dealt with in that system. Secondly, I take a look at why none of the definitions of 'suicide', or of 'suicide note', are simple or as straightforward as they may seem to be. Thirdly (and finally), I present a study of some reports of court proceedings in which suicide notes played a part. In this study I look at what various non-Coroners' Courts have said about suicide notes and see how they contribute towards verdicts. Since this chapter is concerned with the legal system, it is worth mentioning that since the Suicide Act 1961 in England and Wales it is not a criminal offence to kill one's self.

2.1. The Coroner's Court

I am now going to turn my attention to the Coroner's Court system. I wanted to provide a description of Coroners' Courts because they are the source of much of my data (see Chapter 4), and are usually the first, and often only, courts where suicide notes may make an appearance.

A main reference source of legal rules for coroners is Jervis on the Office and Duties of Coroners, named after Sir John Jervis who wrote the First Edition in 1829. I will be referring to the tenth (1986) edition by Matthews and Foreman. According to Lord Justice Glidewell (1986, p.v), "For well over a century, [the book written by] Jervis was the leading work on 'The Office and Duties of Coroners.'". Birmingham City's coroner uses the current edition of Jervis (Coroner's Office, personal communications). It would seem, because there have been so few editions since the first one, that there have been comparatively few changes in the law as it affects coroners over this period of

time. But although changes have occurred, the office of coroner remains largely an oddity (see below). Matthews and Foreman (1986) is the source of much of the information in this Section.[2] Unless otherwise indicated, even when the term 'U.K.' is used, the details given apply to the situation in England and Wales only (and exclude Scotland).

The Coroner's Court proceeds by the process of inquisition rather than accusation. Law courts in England and Wales (unlike, for example, in France) are usually based on the adversarial system, but the Coroner's Court is a rare exception in the U.K. and is based on the inquisitorial system. Matthews and Foreman (1986, p.149) describe the Coroner's Court as "an (inferior) court of record" whose "inquest verdict is not binding on any person trying to raise the same issue in subsequent litigation" (ibid., p.149). And as Lord Justice Glidewell says, "its verdict is not final, in the sense that it determines neither criminal guilt nor civil liability." (1986, p.v). (I discuss inquests in Section 2.1.2 below.)

According to the Coroners' Society web site (Anon, n.d., p.2), "just over one third of all deaths in England and Wales" are reported to coroners. Therefore, in nearly two thirds of cases of death, there is no legal requirement to report a death to the coroner. The web site notes that the law may well change in the future to cover all deaths.

2.1.1. The Coroner

To qualify to be a coroner (or deputy coroner) requires one to be a solicitor, a barrister, or a medical practitioner of at least five years "standing" (Matthews and Foreman, 1986, p.16 & p.41). According to Matthews and Foreman, in 1986 there were "153 coroners in England and Wales (all men save one)" (ibid., p.16), most of whom were solicitors (ibid., p16). They also state that about 90% of these coroners worked only part-time as coroners (ibid., p17). The period during which this thesis was written covered the tenure of more than one Birmingham coroner: the first (1976-2001) had been a medical

---

[2] Some of the quotations I have attributed to Matthews and Foreman seem to have been copied by them from various statutes, and, although properly attributed by them, because there were numerous such sources and because they were not put in quotation marks, I had originally assumed that the wordings were theirs.

practitioner; the second (appointed in November 2001) had been a solicitor; both were full time coroners (Coroner's Office, personal communication). The "Birmingham" suicide notes used in this thesis (see Chapter 4) were collected under the first coroner. The "Birmingham City Coroner" referred to in this thesis is actually the district coroner for "Birmingham and Solihull", this being a "district" within the county of "The West Midlands" (Allen, 2000, p.14).

According to Matthews and Foreman (1986),

> "The coroner is put upon enquiry when he is informed that the dead body of a person is (or in some cases was) lying within his jurisdiction and there is reasonable cause to suspect that such person has died either a violent or unnatural death, or has died a sudden death of which the cause is unknown, or that such person has died in prison, or in such place or under such circumstances as to require an inquest in pursuance of any Act." (Matthews and Foreman, 1986, p.47).

Regarding "violent or unnatural death" (Matthews and Foreman, 1986, p.47), as mentioned in Section 2.2.1 below, they state,

> "There is no definition of what is natural or unnatural and the coroner clearly has to use medical evidence and opinion together with the current values of our society to make a judgment on this issue." (ibid., p.112)

Regarding "sudden death of which the cause is unknown", Matthews and Foreman (1986, p.47) point out that the death in question must be both unanticipated and of unknown cause. Concerning the latter, they suggest that either the "terminal cause" or the "underlying condition which is the real cause of death" may suffice as the unknown (ibid., p.63). In fact, if the coroner thinks that a post-mortem might elucidate the cause, he/she can order one. If this confirms his/her suspicions that there was nothing to warrant an inquest, he/she can decide not to hold one (ibid., p.63-64).

According to Matthews and Foreman (1986) if the deceased died, or is suspected to have died, in prison, then there must be an inquest; but if the deceased died otherwise in police custody the coroner is only advised to hold an inquest. However, if there is an inquest in such a case, there must be a jury (ibid., p64). In the latter scenario the

coroner acts as the judge, but the process is still inquisitorial (Paul Matthews, personal communication).

Coroners are aided by Coroner's Officers. "The coroner's officer is commonly a serving police constable of considerable experience who is specially and permanently detailed for this duty" (Matthews and Foreman, 1986, p.39), but in recent times this officer is more likely to be drawn from "police civilian staff" (ibid., p.40). The coroner's officer's duties include visiting the scene containing the body, arranging for its removal, "searching for evidence..., interviewing potential witnesses,... dealing with lawyers and others representing interested parties", notifying the relevant people about any post-mortems and/or inquests, summoning witnesses and any jurors, and performing as "a mixture of usher and clerk of the court" (ibid., p.39-40). Searching for evidence includes searching the deceased, their property, and the scene (ibid., p.61). The officer may also take possession of property (ibid., p.61-62).

Some interesting points about coroners and coroner's officers were made in a BBC radio programme hosted by Clive Anderson (Anderson, 2002). William Dolman, coroner for North London and one of the participants in Clive Anderson's radio programme, made the point that "Training of coroners falls far behind training of the rest of the judiciary" (Anderson, 2002); and Tom Luce, who chaired a government "review of death certification and the coroner service" (Luce, 2002b), also in Anderson (2002), added "The coroners' officers who do the case work: most of them have absolutely no training whatsoever - and they can't get any."

Recently, changes to the role of the coroner, and the coroner's court, have been proposed. These are largely due to some high profile cases such as the Shipman case. (Shipman was a G.P. (General Practitioner) who was convicted of 15 murders (Smith, 2003, p.2), but who had probably killed hundreds of his patients (Smith, 2002, p.1).) On June 12th, 2006, the Draft Coroners Bill was published. Its proposals included having a chief Coroner, and all coroner posts becoming full-time (Anon, 2006b, p.1). The Bill claimed to be based on both the Shipman Inquiry (Smith, 2002, 2003 & 2005) and the Luce Review (Luce, 2002a and 2002b).

2.1.2. The Inquest

In my quest to discover the role played by suicide notes in certain procedures, I now look at the key procedure in the coroner's court, the inquest. Inquests are held in public except when matters of national security prevail (Matthews and Foreman, 1986, p.151). However, there is no legal requirement for the public to be notified of where or when there is to be an inquest (Matthews and Foreman, 1986, p.138).

> "The functions of inquest on a dead body at the present day are really to determine certain facts about the deceased: his identity, the cause of his death, and the circumstances surrounding his death and that cause." (Matthews and Foreman, 1986, p.6, para.1.8.).

Although inquests are about "fact finding not fault finding", as one coroner, William Dolman has said (Anderson, 2002), one judge, Anthony Clark (also in Anderson, 2002), said that "In real life Coroners' Courts are debating questions of blame and truth all the time." However a coroner must adjourn an inquest when it emerges that the case involves pending, or potential, criminal proceedings (Matthews and Foreman, 1986, p.162).

Inquests do not necessarily involve juries. Apparently,

> "Of the 25,000 inquests in England and Wales in 2000 representing about one in every 20-25 deaths,... 3% were held with juries,..." (Luce, 2002b, p.35, para.72).

That very few inquests involve juries is because the coroner is only required to summon a jury in a few scenarios. To paraphrase Matthews and Foreman (1986, p.143), these include situations in which a death has occurred in prison, in police custody, and in circumstances which could continue to endanger public health or safety. Juries consist of 7 to 11 people (Matthews and Foreman, 1986, p.144). It is the coroner who ultimately decides on the exact number of jurors, and the criterion for this decision seems to be a matter of obtaining the maximum number possible within some reasonable amount of time (Paul Matthews, personal communication).

Regarding witnesses, although during an inquest "it is the coroner alone who has the power to call witnesses" (ibid., p.140), there is "one exception": if there is a jury the jury may "require a medical practitioner nominated by them to attend to give evidence before them" (ibid., p.140, fn.36). Additionally, the jury can require that this witness perform a post-mortem (Matthews and Foreman, 1986, p.84).

Witnesses commonly include those who found the body and those who identified it; they include police and independent witnesses, friends and relatives; and they may include medical and expert witnesses. A witness is first questioned by the coroner, then by anyone the coroner considers to be a "properly interested party" (ibid., p.175). The latter can include family members. Not calling a relevant witness may result in a coroner being ordered to so do by a higher court (ibid., p.140). The jury may recall witnesses and question them if the coroner allows (ibid., p.175). Coroners should stop any irrelevant or improper questioning (ibid., p.176).

> "As to the propriety of particular questions, a coroner must be careful not to allow the inquisitorial nature of his proceedings to become oppressive and an intrusion into the privacy of individuals' lives. For example, in suicide cases the coroner should not extend his enquiries into a general or detailed survey of the matrimonial or financial problems of the deceased, even though the view might be taken that this was relevant to motive." (Matthews and Foreman, 1986, p.176, para.15.20).

Copies of documents are only given to "parties to the proceedings" (ibid., p.187). It is not clear whether "given" means shown or given to keep. The definition of the "parties" would probably be at the coroner's discretion, but it would almost certainly exclude members of the public and the press, and would include lawyers, jury members, and some witnesses and relatives. Further, regarding documentary evidence,

> "The document must be read aloud at the inquest unless the coroner otherwise directs. In some cases, for example suicide notes, it may not be desirable to make the contents public" (p.187, para.15.56).

The Home Office guide (Anon, 1996, para.21) expands on this: "Suicide notes and personal letters will not be read out unless they have to be,...". However, the guide only elucidates what might amount to such a necessity as being "in the interests of justice"

(ibid., para.21).

Eventually, the coroner sums up the evidence to the jury (if there is one), and explains to them the limits of an inquest and describes the possible verdicts (Matthews and Foreman, 1986, p.188). The jury (if any) retires to consider its verdict (ibid., p.190); if there is no jury then the coroner considers a verdict (ibid., p.190). "The coroner is not obliged to accept a majority verdict, but he must accept a unanimous one, however perverse or inane it may be, if the jury after suitable explanation persist in their view" (ibid., p.191).

2.1.3. The Verdict and the Inquisition

According to Matthews and Foreman (1986, p.191, para.16.6) "The formal record of the inquest, containing the verdict, is called the inquisition". In this document the term 'verdict' includes the cause of death and various details about the deceased (ibid., p.190); however, throughout this thesis I use the 'normal', lay, meaning of the word 'verdict'. That is, I use 'verdict' to refer to the cause of death (in terms of the overall conclusion of a case (the 'Coroner's verdict'), rather than the medical cause) excluding the 'details'. Discussing verdicts, Matthews and Foreman say:

> "Note (4) to the prescribed form of inquisition gives a comprehensive list of suggested (though not compulsory) verdicts. The object of this list is to standardise verdicts over the whole country and to make the statistics based on the Annual Return more reliable by avoiding as far as possible any overlap or gaps between the different verdicts." (p.193, para.16.11).

The list of verdicts from Matthews and Foreman is quoted below in Section 2.2.1. The two of most direct relevance to the data presented in this thesis are the suicide and open verdicts. It might be worth repeating here that, as implied in Section 2.1.2 above, if the coroner suspects unlawful killing he or she is obliged to adjourn the inquest so that other legal proceedings may take place. Some of what follows is also referred to in Section 2.2.1 and in Chapter 4 of this thesis, but is repeated here for convenience.

Matthews and Foreman (1986, p.194, para.16.14) state that evidence of intent is required for a verdict of suicide. They add: "the courts will not quash a verdict of

14

suicide unless no reasonable coroner could have found suicide in such circumstances." (ibid., p.194, para.16.14). They go on to say that,

> "The suggested form of verdict in suicide cases is that the deceased 'killed himself [whilst the balance of his mind was disturbed].' The bracketed words referring to the state of mind of the deceased are of no legal significance, but some coroners may like to add them when the evidence merits it. Such coroners like to soften the impact of suicide on the relatives... However, the object of the inquest is not to determine the state of mind of the deceased when he committed suicide, but to ascertain how he came by his death. For this purpose it is sufficient to record in the verdict that the deceased killed himself, and the use of the additional words... ought to be based upon some evidence to that effect..." (ibid., p.194, para.16.14).

Atkinson (1978) interviewed coroners during his attempt to clarify the concept of evidence of suicide. His findings of what amounted to institutionalised assumptions on the part of his interviewees are mentioned in 2.2.1 below.

Regarding an "Open Verdict": this may be given only if no other verdict can be decided upon (Matthews and Foreman, 1986, p.196, para.16.19). As Tom Luce said (the full quotation is given in Section 2.2.1), suicide and unlawful killing are different from the other verdicts because they must be proven "beyond reasonable doubt" rather than on the "balance of probability" (Luce, 2002a, p.45, para.98). Thus, open verdicts can be used when there is insufficient evidence of suicide.

After the inquest the inquisition document is signed by the coroner and all those jurors who agree with it (Matthews and Foreman, 1986, p.198). Controversially, "there is no obligation on the coroner to notify the relatives of the deceased of the outcome of the inquest..." (ibid., p.251, para.20.20). However, after an inquest, except in situations that do not require registration of death, a coroner must furnish a certificate for the registrar of deaths, and he, or she, may also have to supply copies of, or facilitate access to, documents and exhibits (ibid., p.248, para.20.10).

2.1.4. Access to Documents

For the forensic linguistics researcher a key issue is what happens to documents after the legal proceedings have finished with them, because the researcher may want to see

original texts and/or acquire copies of them. As will now be seen, what happens to suicide notes after an inquest is far from clear.

> "The coroner must on application and on payment of the prescribed fee supply to any properly interested person a copy of any report of a post-mortem examination or special examination, any notes of evidence or any document put in evidence at an inquest." (Matthews and Foreman, 1986, p.254, para.20.27).

Regarding more general access to the inquisition document, Matthews and Foreman (1986, p.254, para.20.27) note that "the coroner has a power, but no duty, to supply copies of inquisitions to such persons as apply (not necessarily 'properly interested persons')." As regards other reports and written evidence, on request, at the coroner's discretion, documents may be inspected free of charge. This need not wait until an inquest is ended (Matthews and Foreman, 1986, p.254, para.20.27). According to Matthews and Foreman, a "properly interested person" would probably include a relative, beneficiary, life insurer, someone who might be responsible for the death, a representative of them, or of the deceased, or some appropriate authority, a police chief, or anyone else whom the coroner thinks fit (ibid., p.255, para.20.29 & p.150, para.13.2.). They suggest that the press and the general public do not constitute such persons.

When a coroner can envisage no legal situation in which an exhibit may be of further use, he must give it to anyone requesting it, provided he thinks they are entitled to it. In the absence of any such request, "the coroner may destroy it or dispose of it as he thinks fit" (ibid., p.256, para.20.33). Matthews and Foreman add,

> "Care should obviously be exercised in dealing with exhibits in suicide cases, such as suicide notes or the instrument with which a suicide was committed,..." (p.256, para.20.33).

They then say,

> "Any document, other than an exhibit at an inquest, which is in the possession of the coroner in connection with an inquest or post-mortem examination must be retained by the coroner for at least fifteen years. This will include... all documents put in evidence at an inquest..." (ibid., p.256, para.20.34).

Instead of waiting 15 years, however, coroners can decide to hand over documents, or copies of documents, to people they think are appropriate (Matthews and Foreman, 1986, p.256, para.20.34). From the two quotes above it is not exactly clear what the position is with regard to keeping suicide notes. A more up to date answer probably obviates the need to worry about this: according to the Birmingham Coroner's Office, the notes are "kept for six months" (personal communication, March 2003). However, clearly this is either not wholly accurate, or not strictly adhered to, as is evident from the existence of my Birmingham suicide note corpus (see Chapter 4). Perhaps part of the issue here is that an exhibit may be returned to someone who "appears" to be "entitled" to it, whereas a document which was not an exhibit may be returned to someone who "seems" appropriate for its "possession" (Matthews and Foreman, 1986, p.257, para.20.35). In the case of the deceased's property this would be their estate via their executors, and any unsent letters would be part of the estate. Matthews and Foreman conclude...

> "In general the coroner should return all personal possessions belonging to deceased persons, subject only to special considerations arising out of suicide cases." (p.257, para.20.35).

Apparently, "The records of coroners' courts... are public records for the purposes of the Public Records Act 1958..." (ibid., p.257, para.20.36). "Except for the indexed registers of deaths and papers relating to treasure trove or matters of special historical interest,.." and "records from earlier than 1875", all records are kept for 15 years and may then be destroyed (ibid., p.257, para.20.36).

> "Access to coroners' records deposited under the Public Records Act 1958 will not be given to the public until 75 years have elapsed (in the case of records relating to reported deaths) or 30 years (in the case of other records), unless the coroner of the court from which the records came gives special permission in writing for such access." (ibid., p.257-258, para.20.38).

All this shows how very confusing things are, especially because suicide notes, as distinct from any other documents, always seem to be treated as an exception to whatever rules exist.

2.2. Some Definitions

I now attempt some definitions of 'suicide' and of 'suicide note', from both the legal position and my own perspective. Before consideration of what constitutes a suicide note, I shall consider how to define 'suicide'.

2.2.1. What is a Suicide?

According to Collins Concise English Dictionary (3rd Edition, p.1346), suicide is "The act or an instance of killing oneself intentionally". Black's Law Dictionary (6th Edition, p.1434) gives "Self-destruction; the deliberate termination of one's own life." Yet another definition comes from Stincelli (2001, p.2): "Suicide is the intentional act of taking one's own life, with a conscious awareness of both the effect and finality of the act." However these definitions do not fit with the concept of getting someone else to kill one, for example, by making them act out of sympathy, as in the concept of euthanasia, or self-defence or defence of others, as in the concept of 'suicide by cop'. Indeed, the apparently simple concept of suicide is actually quite a difficult one. I shall elaborate on the 'suicide by cop' situation here because it is useful (that is, as useful as any other such as, say, euthanasia) for illustrating several of the points I wish to make.

An example of an attempted 'suicide by cop' is from 2003, in Scotland, when a 31 year-old woman with a gun engaged the police in a two hour siege (Scott, 2003). The police, who had been informed that the gun was probably an air pistol, shot her, but she did not die. According to Scott (2003, p.1), it seems that she, herself, "allegedly called the police to the house... saying she had shot her mother." This case is notable because, apart from being, apparently, only the second time that U.K. police have "deliberately shot" a woman (ibid., p.1), as Scott reports, "The police confirmed... that a suicide note had been found at the scene, saying: 'It appears she was attempting to commit suicide by being shot by a police officer.'" (ibid., p.1). Of course, without any quotations from the note, it cannot be known whether the woman did actually make explicit her intention, or whether the police made inaccurate conclusions. Although this is only a sketchy example of a failed attempt at 'suicide by cop', it does show that this method of suicide exists, and is gaining recognition, in the U.K. (see below).

Even the more 'traditional' concept of suicide proves elusive. Definitions in the academic and legal literature tend to be couched in terms of Coroner's Court verdicts. To provide some perspective on Coroner's Court verdicts, and an idea of the numbers of deaths involved in them, here are some figures published by the U.K. Government Statistical Service (Allen, 2000). According to Allen (2000, p.1), "201,000 deaths were reported to coroners in 1999..." - that is 36% of all registered deaths in England and Wales (ibid., p.3). Inquests (see Sections 2.1.1 and 2.1.2) were held on 24,400 (12%) of these. Verdicts of suicide were returned in 17% of inquests (ibid., p.5). According to Luce (2002b, p.35, para.72), in the following year, 2000, there were 25,000 coroners' inquests, and verdicts of suicide were returned in 16% of them.

As stated in Section 2.1.1 above, usually when a suicide is suspected to have occurred it is dealt with by a Coroner's Court. Coroners look at (among other things) deaths which are violent or unnatural, or sudden with unknown cause. However (as stated in Section 2.1.1 above), according to the authors of one of the major resources used by Coroners, regarding violent and unnatural deaths, "There is no legal definition of what constitutes an unnatural death" (Matthews and Foreman, 1986, p.62).

Matthews and Foreman (1986, p.193) list the usual verdicts given by Coroners' Courts: "natural causes; industrial disease; dependence on drugs/non-dependent abuse of drugs; want of attention at birth; suicide; attempted/self-induced abortion; accident/misadventure; sentence of death; lawful killing; open verdict; unlawful killing; stillbirth.". Atkinson (1978) came up against much resistance when he was trying to research into the categorisation of the different verdicts and the assumptions surrounding suicide. He interviewed several coroners but could not get a straightforward answer as to what was the legal definition of suicide.

Some of the difficulties in determining a verdict of suicide are mentioned by Etkind (1997). According to Etkind, "Many people try to disguise their suicides as accidents", (ibid., p.40); and there is "what researchers call chronic suicide", (ibid., p.67), a sort of self-neglect over an extended period of time with "continuous self-destructive behaviour

such as drug use, drinking, and reckless driving", (ibid., p.67). And, of course, as mentioned above, there are possibilities such as pointing a gun at an armed police officer, knowing she/he may well open fire pre-emptively.

This brings us back to the topic of 'suicide by cop'. According to Bresler et al. (2003, p.1), who reviewed some of the literature on the matter, instigators of 'suicide by cop' tend to leave "a suicide note or some communication of an intention to kill one's self prior to [sic] incident...". Leaving aside the dubious distinction between "a suicide note" and "some communication...", it seems that suicides who use this method might feel the need to explain their actions so that others do not mistake their cause of death for something else. If they want their suicide to be so recognised, then a note is important to convey this.

As reported in Peek (2003), some people have a problem in classifying 'suicide by cop' as any type of suicide. Their stance is that, since the word "suicide" literally means killing oneself, it cannot really be suicide when someone creates a situation in which somebody else kills them. One could imagine having two verdicts for such cases: one for the suicide, and one for the (lawful or unlawful) killing. But this might be impractical, implying that there were two events, rather than a multifaceted single event. In suicides there is a scale of involvement of others, ranging from low, when a person hangs themselves, to high, when a person engineers a 'suicide by cop'. A relatively intermediate situation is when someone lies down on a railway track. Unlike police who could theoretically decline to shoot the would-be suicide, a train driver would probably not even be aware of the person on the track until it was too late, but there seem to be few, if any, calls for incidents such as these to be classified as anything other than suicide. Because the police intent to kill, if necessary, is there in the 'suicide by cop' scenario, I believe that it is a distinct form of death, and that a separate verdict would be appropriate. However I do not object to the word "suicide" in this context.

According to Stincelli (2001, p.2), in the United States, "There have been a few deaths classified as suicide by cop, but most Medical Examiners today continue to classify these deaths as justifiable homicides...". Despite this, it tends to be thought of as a

U.S.A. phenomenon.   However, the U.K.'s first legally acknowledged instance of "suicide by cop" occurred in 2001 during a 9 to 10.5 hour siege (Peek is ambiguous about the time-frame) in which 62 year-old Michael Malsbury had attacked his wife with a rolling pin (Peek, 2003).   Apparently, "Summing up, Dr Dolman [the coroner] told the jury: 'I want you particularly to bear in mind the words "suicide by cop"'", and "After the verdict Dr Dolman told the jury: 'I think we've made legal history today.'" (Peek, 2003, p.1).   According to Peek, the verdict was "suicide by cop".

Ford & Moseley (1963) presented several cases of suicide by automobile.   They do not say whether the actual legal judgments and verdicts were "suicide" as such or some version of this (for example, "killed himself"), but they do say that suicide notes, or statements of intent, are important evidence in assessing that such cases are indeed suicides rather than, say, accidents.   Indeed intent seems to be the crux of the matter. "The fact of suicide should never be presumed, but must always be based upon some evidence that the deceased intended to take his own life." (Matthews and Foreman, 1986, p.194, para.16.14).

According to Van Wormer and Odiah (1999, p.365), in Denmark, 42%, and in the U.S.A., 5%, of murders are murder-suicides.   They say that 1.5% of suicides in the U.S.A. could, indeed, be murder-suicides.   Having looked at 22 U.S. murder cases, they determined that "execution [or hope of it] attracts murder" (ibid., p.369).   They seem to be positing that many of these murderers are killing primarily in order to kill themselves either by coercing themselves to commit murder, thus making it easier for them to kill themselves, or instead expecting a death penalty, or by instigating an armed police response.   According to Palermo (1997), mass-murderers, whose killings tend to be effected in order to illustrate various injustices, are frequently suicidal.   In terms of English and Welsh coroners' verdicts, one should perhaps include a percentage of "lawful killings" as suicides.   This is in addition to the problems with the other verdicts, such as "accidental death" and "neglect".

According to Barraclough and Harris (2002, p.577), who investigated murder-suicides in England and Wales that took place in 1988-1992, such cases represent about 0.01%

of all deaths. They actually use the term "homicide-suicide" to include "manslaughter" and "infanticide" verdicts (ibid., p.578). Barraclough and Harris found that murder-suicides accounted for 3% of male, 11% of female, and 19% of child deaths, due to homicide (ibid., p.577). They also found that "Similarly, of all suicides, 0.8% of male and 0.4% of female deaths occurred in homicide-suicide incidents" (ibid., p.577). Barraclough and Harris (2002, p.577 and 578) point out that it is the police who inform the Home Office about connected murder and suicide, and that this can be at odds with the figures derived via coroners' offices. All this emphasises the difficulty of defining suicide.

Although Atkinson's (1978) interviews (see above) with coroners and coroner's officers failed to produce a consistent account of what might amount to "evidence" of suicide, it is clearly bound up in the concept of intent. Tom Luce, in his review of the Coroners' Court system, states,

> "A good deal of effort is spent in inquests into deaths by the action of the person who has died trying to find out whether the death was an intended consequence of the action taken (in which case, if proved, the verdict is 'suicide'), or was more in the nature of a signal of acute distress and a need for help and support which accidentally went further than may have been intended, in which case the verdict is usually 'accidental death'. It may well often be impossible to know the answer..." (Luce, 2002a, p.45, para.98).

In the case of Malsbury (see above),

> "A PCA [Police Complaints Authority] spokesman, Richard Offer, said: 'Mr Malsbury did leave a suicide note and did verbally indicate what he wanted to happen. If you look at it from the point of view of the coroner and the jury who are presented with the evidence of a suicide note and verbal intention, they have to take this into account in terms of findings for the cause of death.'" (Allison, 2003, p.2).

So, as others have noted, finding a note is, apparently, an important piece of evidence when determining whether or not a case is a probable suicide - especially if it states intent. But the problem of determining intent is compounded by the degree of proof required. As William Dolman, coroner for North London, said (and as was mentioned

in Section 2.1.3 above),

> "Open verdicts may be used because the evidence doesn't come up to the standards for the most difficult verdicts, the two most difficult [verdicts] being suicide and unlawful killing which demand a criminal standard of proof." (Anderson, 2002).

Additionally, as Luce said,

> "We also question whether the verdict of 'suicide' is apt or necessary. Along with the 'unlawful killing' verdict it requires proof at the criminal standard 'beyond reasonable doubt', while for all the other inquest outcomes the civil 'balance of probability' test applies." (Luce, 2002a, p.45, para.98).

In one case, where a man had been convicted of murdering his wife whom he claimed had killed herself, the judges, in their summing up at his appeal hearing, said,

> "When the issue is homicide or suicide and... there is scope and need for a jury to consider the likelihood or otherwise of the victim having been suicidal, judges should draw juries' attention to the unpredictability of human nature... Outwardly happy and contented people may commit suicide without warning. Those who are deeply unhappy or depressed or mercurial in temperament may have more resilience than their behaviour suggests. In the end, all a jury can do is look at the signs each way and keep in mind, along with the evidence as a whole and the burden of proof, that, from such signs on their own, 'you never can tell'.", (R v Kavanagh, 2002, para.57).

This seems to typify the legal position, such as it is; and it is one reason why suicide is a difficult verdict to reach. I add that I do consider cases both of 'suicide by cop', and voluntary euthanasia, to be categories of suicide, although there were no such cases represented in my main data collection.

As a researcher, however, I have had to establish a rather more restrictive view than one that tries to encompass all the possibilities mentioned above. My position is this: a narrow view, based on Coroner's verdicts (see Chapter 4) - at least for the purposes of

the bulk of my data collection activities. Thus, in this thesis, a suicide is a death that has been given a Coroner's Court verdict of suicide. This working definition thus excludes scenarios in which a different categorisation, such as a different Coroner's Court verdict, might well be considered an actual, but legally unrecognised, suicide.

2.2.2. What is a Suicide Note?

There are two issues that are pertinent to whether or not a note is a suicide note. One is concerned with whether or not a particular piece of writing is a suicide note, and the other is concerned with whether an apparent suicide note is genuine. This sub-section addresses both issues. I begin by considering the following quotation which is from the transcription of a telephone call in Hampshire, U.K.:

"Mr K.:	'I would like to report a suicide.'
Control:	'Who is it, do you know?'
Mr K.:	'Yeah, it's me.'
Control:	'Well, I'm speaking to you so it can't be you, can it?'" (Bird, 2000, p.9).

In fact it could, and it was; and after a little more conversation with the emergency services telephone operator, Mr K shot himself. It transpired that he had left a written suicide note, but might the transcript above also be considered as a suicide note? Clearly it is not of the "note" genre of texts (because it is not written down, and therefore is not a note until a transcript is produced) - but this may not be too important compared to the alternative of omitting such communications from any consideration of suicide texts.[3]

The legal position is based on the location of the note relative to the body of the suicide, and the note's indication of intent to commit suicide. To a lesser extent any evidence of a date is considered. According to the Birmingham Coroner's Office (personal communication), the procedure regarding suicide notes is as follows. Prior to any inquest, the note's proximity to the body is important. If it was found on, or very near to, the body, then it does not matter if it is not dated. If it was found in a drawer

---

[3] This "quote" is not quite as it appeared in Bird (2000): I have not given Mr K's name here. The text shown is only the first half of the transcript reported in Bird.

however, then it is not considered to be a suicide note, unless it is appropriately dated, which means, presumably, to within a few days or so before the death. As an aside, it might be worth mentioning that some judges seem to be of the opinion that to be a suicide note, a note ought not to be have been concealed by its author: "If the Note was intended as a suicide note, why was it not left for the appellant to find?" (R v. James, 1998, p.6).

However, regarding dates, it seems reasonable to suppose that the date of a note could be more important than the note's location, and that the Coroner's Officer (see Section 2.1.1) would actually be more, or at least equally, concerned with a note dated on the day of death, wherever it was found, rather than an undated note found near the body. This could be one way of deciding on a note's authenticity. In the case of Gilfoyle (see Chapter 1 above), according to the courts, Mr Gilfoyle persuaded his wife to write a suicide note in her own handwriting, and then killed her, making it look like suicide. The prosecution claimed that Mr Gilfoyle had been in possession of the note for some time (Malcolm Coulthard, personal communication). The note was not, apparently, dated. If it had been dated, and the date had been written in the same style and handwriting as his wife's, that would have contributed to the evidence in favour of her (Mrs Gilfoyle) having written it on the day she died, and the case being a genuine suicide, because it would probably have been more difficult to plan in advance a specific day for the murder.

Of course, a date in a note could be forged, or a victim could be persuaded, or forced, to include a future date that was within a few days of a planned murder. An already dated note, prepared days, or months, earlier, or from a previous suicide attempt, could be re-used (see Section 5.2). Or the date could simply be wrong as the result of a mistake by the note-writer (see Sections 4.4 and 5.2).

Courts do not discount other genres of text, such as diaries, although they seem not to refer to them as "notes". Regarding the content of notes, the coroner seeks "intent" to further sway him in the opinion that it is indeed a suicide note (Coroner's Office, personal communication). Any indication of intent can accrue as evidence of suicide.

Thus the transcript at the beginning of this Section would be treated as if it were a note, and it would be important in determining that the case was one of suicide.

Given the imprecision of the legal definition of a suicide note, in deciding what data to collect I took a very broad view and simply amassed whatever was available in cases where a suicide verdict had been returned. My initial corpus, therefore, consisted of a very wide range of items, some of which could later be discarded if it was decided that they did not fall within my ultimate definition of suicide note. Thus, mine is an inclusive strategy, and my working definition was largely derived from what presented itself during the data collection process, albeit parameterised by the Coroner's verdicts (see Chapter 4.).

When confronted with some quite diverse data, I made the initial decision to include all genres of suicide note, i.e. diaries, letters, lists, greeting cards, etc. In Section 4.4 I consider more precisely which types of text should be counted as notes in the light of the data collected, and I quote some statements from the Birmingham Coroner's Office. In fact, I would have included the above transcript by Bird (2000), albeit in a sub-category of transcripts of spoken suicide dialogues. My position is, thus, to include any available text by a suicide which was authored shortly before the death.

This discussion of the definitions shows that definitions are difficult and extremely problematic, and not nearly as precise as the lay person might expect them to be. The researcher has to make many subjective decisions in data collection. I shall return to this topic of definitions in Chapter 4.

2.3. A Study of Case Reports

As background to this thesis, for this chapter I wanted to look into how law courts refer to suicide notes and what they make of them. Having discussed some of the problems involved in definitions, I now return to the main issue of this chapter, namely, 'how suicide notes are used in court'. In view of the lack of detail and transparency in Coroners' Courts, I felt it would be worthwhile to look at other U.K. courts. My main source of information on the use of suicide notes in courts, other than Coroners' Courts,

was the "LexisNexis" internet website database.

LexisNexis is a huge website database of newspaper reports and reports of legal cases, plus various other legal resources including journal articles, statutes and transcripts of court judgments. The coverage is worldwide. I have accessed reports of court proceedings from various courts and quoted some of the interesting comments made about suicide notes and related issues. This was achieved by initially submitting a search term of "suicide note" to the website. The point of this, as mentioned above, was to look at the way people talk about the notes in law courts.

Some cases on the database were labelled as transcripts but they did not seem to be full and verbatim transcriptions. There were, as is intimated above, no reports, nor "transcripts", from Coroners' Courts. Indeed, nearly all of the cases resulting from my search were Appeals. The Appeals were from the High Court, the House of Lords and the Court of Appeal. (The various courts are described in Anon (2007).) Appeals, of course, must stem from earlier proceedings, but the original proceedings did not seem to be on the database, and the Appeals did not repeat every detail from the original cases. Therefore suicide notes which may have been read out in original trials may not have been read out in a related appeal. I will not elaborate on the history and original trials of all the individual cases discussed; nor do I intend to give an account of all the procedures of any of the "other", i.e. non-coroners', courts. I have also avoided determining exactly whom to acknowledge as having said what because it can be a little complicated: in each Appeal case there were several judges, and they often cited and quoted not only various witnesses and lawyers, but also other judges, both from preceding, directly related, trials or Appeals, and from other cases which they were considering as legal precedents.

My search was conducted in the "Combined Courts" part of the U.K. cases database. This covered courts in England and Wales, Scotland and Northern Ireland. My primary search criteria were the phrase "suicide note", and dates between, and including, 1/1/1900 and 21/5/2003. (In Section 2.3.1, below I describe my secondary search.) Although the LexisNexis website provides details of the sources of the cases its

database contains, to a non-law student it is not clear what other sources exist that it does not contain. The different sources provided cases from various dates, the earliest being 1558. However, in the results from my search there were cases from the 1960s to the present.

My initial focus of interest was to see what the courts said about suicide notes and their contents. It was assumed, from conversations with staff in the Birmingham Coroner's Office, that, as in Coroners' Courts, other courts would seek "intent" in a suicide note - and it seemed that in some cases this was so. I was also interested in whether or not notes were disputed, or analysed by lawyers, linguists, handwriting experts or document examiners. ("Disputed", above, means disputed in whatever court case I have found on the LexisNexis database regardless of whether the notes were disputed or not in any original, or further, proceedings.) Indeed, I was interested in any remarks about suicide notes, and seeing what role the presence or absence of suicide notes have in determining verdicts. Unfortunately, there was no clear view of any of these points. However, what did emerge was a fairly clear division of the data yielded by my search term into particular classes of collocations, and these are described in sub-sections 2.3.1.1 to 2.3.1.7 below. My macro view of how suicide notes are used in court, which is reflected in the following sub-section headings, included their use as indicators of state of mind, and as partial evidence for suicide.

Ideally I would have liked to examine the analyses of linguistic expert witnesses in suicide-related cases, but no such evidence was available on the LexisNexis database. Thus, most of what follows is analysis mainly from lawyers and judges: it is their interpretation of suicide notes and the notes' contents. This can include notions of what a suicide note is, the meaning of the existence of a note, and also the meaning of the absence of a note.

It should be noted that the cases span several decades; that laws and attitudes change; and that any future proceedings concerning a case might revise, or reverse, the verdict. It seems likely that many legal judgments will continue to be disputed, particularly by the people they directly affect, but, for simplicity, I shall refer to the outcomes of the

cases according to the courts' verdicts - as I do elsewhere with the Birmingham Coroner's verdicts.

2.3.1. The Cases

The database search for cases resulted in 30 hits for the phrase "suicide note", and included three cases from Scottish courts. Curiously, while checking this, it was found that one of the Scottish cases also appeared in a search of the English and Welsh cases - which numbered 26 including the aforementioned case. That case appeared to be accessible from the wrong part of the database by mistake; I have only counted it once. However, within the "Combined Courts", three of the other cases seemed to be duplicated. In each of these pairs of cases the texts were similar but not identical, and differences included, variously, case reference codes, dates of the hearing, names of the court, and the name of parties in the case title. But it is not clear whether they are actually duplicates, inaccurately transcribed, or are the result of LexisNexis having more than one source for some of the cases on its database, or are primary and secondary appeals of the same case. Thus, the 'duplicates' of the three aforementioned cases were included, and in all I have assumed that 30 cases were found: 16 of these were from the Court of Appeal (most were from its Criminal Division, but some were from its Civil Division); many of the others were appeals from the Queen's Bench Division of the High Court (see Anon. 2007), two were from the "Probate, Divorce and Admiralty Division" (of the High Court); and a couple were from The House of Lords. The three Scottish cases were from the Outer House of the Court of Session.

To see how variations of the term 'suicide note' were used in and across a collection of legal cases via what was intended to be based on a simple search of the database, a secondary search term was applied to the already acquired data resulting from the primary search for the phrase "suicide note". The secondary search term was made case sensitive and was "suicide* note*/suicide* letter*". The asterisk is a wildcard character that allowed for plurals and for any intervening punctuation to be included, and the slash represents the logical "OR". There were also supplementary searches for the terms 'note' and 'letter' to assist in finding informative quotations that might help answer the general question of how suicide notes are used in court.

The first point of interest was that the phrase "suicide note" was itself quoted in different forms. As Table 2.1 shows, ignoring other punctuation (i.e. counting "suicide," note, for example, within the category of "suicide" note) several variations of the term "suicide note" were found. It should be borne in mind that it is not known precisely how the court reporters determined the need for quotation marks; and, more relevantly, the reports are, presumably, not post-edited by the judges.

Table 2.1. Notes and Letters on LexisNexis to 2003

Phrase	Uses	Cases
suicide note	39	24
suicide notes	8	4
SUICIDE NOTES	1	1
"... suicide note"	1	1
suicide letter	15	3
suicide letters	1	1
"suicide" letter	3	1
"suicide" note	3	1
"suicide" notes	3	2
"suicide note"	10	3
[suicide] note	1	1

N.B.: A case may contain more than one of the phrases.

The sub-sections below contain example quotations from the database: not every instance of the search term, nor every variation of it, is included. Further, a single case may contain several examples of the search term, each of which may fall into different categories. Having collected some quotes about suicide notes from the texts, to see what categories they themselves suggested that they may fall into, it became apparent that the initially seen categories tended to coincide with certain patterns of collocations. However, Section 2.3.1.6 on expert testimony, has a dual function. As well as showing collocations, it also sheds some light on how the courts consider some issues regarding suicide by including quotations that do not contain any of the search terms (although the quotations come from cases that were found via the primary search term).

In the following examples any bold, i.e. highlighted, text is mainly the result of LexisNexis' processing of the primary search term "suicide note". I also applied bold text to some phrases containing output generated by the secondary search term (such as "note" and "letter".) Any underlining or italics (to illustrate collocations) is my doing unless stated otherwise. Any paragraph numbers cited are those found in the LexisNexis database.[4]

2.3.1.1. Negatives: Implications of the Absence of a Note

In nine of the 30 cases reviewed there were no suicide notes, disputed or otherwise. Four of these concern hypothetical notes, and are discussed in the next sub-section; one (R v. Taylor) mentions a suicide note in another case, and is presented in Section 2.3.1.4; the remaining four cases are discussed here. In these four cases the phrase "suicide note" co-occurs with negative items.

As forensic linguists are well aware (Coulthard 1992, p.252), thousands of things do not happen all the time, but they do not tend to get mentioned. There are of course many legal cases in which the term 'suicide note' is not mentioned at all. Trial participants do not say "There was a burglary; this man is guilty of burglary, and there was no suicide note." - it would be ridiculous. So the fact that a suicide note is mentioned at all, and is negated, suggests that it is a possibility that there was a suicide note. As Martin and White (2005, p.118) say,

> "Negation is a resource for introducing the alternative positive position into the dialogue, and hence acknowledging it, so as to reject it."

In other words, the inclusion of a negative tends to imply the possibility of a positive. The following extracts show this, and they give a more tangible form to the reasons why the term "suicide note", or an equivalent expression, was mentioned. The negatives are underlined.

---

[4] Readers should note that the pound sterling signs have been electronically transformed into hash signs during the downloading to, and/or subsequent transferring between, the different computers used during the lifetime of the production of this thesis.

"It was not the clearest of cases, Mr Kelly submitted. There was, for example, <u>no</u> **suicide note**." (R v. HM Coroner for Portsmouth, ex parte John Keane, 1989).

In the above case it is clear that a suicide note might have helped to explain the situation.

"Whilst in police custody he had not shown any intention to commit suicide and <u>no</u> **suicide note** was subsequently found."

"The coroner has sworn an affidavit for the purposes of these proceedings in which he has given his own commentary upon the evidence. In it he says that he took into consideration that there was <u>no</u> **suicide note** and that the deceased had not said anything consistent with a suicidal intention." (R v. HM Coroner for Northamptonshire, ex parte Anne Walker, 1988).

In the above case, it seems that although the coroner had found the lack of a suicide note and evidence of intent were not a hindrance to a verdict of suicide, the appeal judges disagreed with him and ordered a new inquest.

"But sadly there was at least some evidence here derived from the circumstances of the death on which a jury could be sure that the deceased intended to take his own life. ...

Of course, there were strong pointers the other way... including <u>the absence of any</u> **suicide note**. ... But as matters stand, I cannot say that there is no evidence on which the jury could have reached that view." (R v. HM Coroner for Solihull ex parte Nutt, 1993).

In the above case, again disagreeing with the coroner, the appeal judge considered that the lack of a suicide note was, at least, partial evidence against a verdict of suicide, and quashed the inquest verdict.

"There was <u>nothing in the form of</u> a **suicide note** between himself and Danielle, and the time at which he and Danielle made the suicide pact, according to the evidence of Ian Wood, was at 10 o'clock on the Sunday morning, 21st September..." (R v. Wood, 1989).

Similarly, in the case above, the lack of a note weighed against a verdict of suicide. In instances of suicide pacts, if some members die and others do not, the survivors are at

risk of being prosecuted unless they can prove that there was a pact (Matthews and Foreman, 1986, p.205, para.17.11; p.206, para.17.14 & 17.15). Writing a note is, therefore, crucial when embarking upon joint suicides.

2.3.1.2. Hypothetical Notes: Points of Law

The cases discussed in this sub-section contain references to hypothetical suicide notes. There is a lot of debate in these extracts about the difference that the absence or presence of a note would make. However, looking at the words and phrases which collocate with the term "suicide note" (see underlined material below) does not seem to reveal much. We may also note that there are both positive and negative collocates of the phrase "suicide note".

> "But to exclude altogether the power to admit such evidence in the cases I have described would impose an extraordinary fetter upon a tribunal whose function is to ascertain the truth by an inquisitorial process. We have only to think of the simpel [sic] case of a **suicide note**, which could well be germane to the question how the deceased met his death, but which could only be admitted, if at all, as documentary evidence. It is difficult to imagine that those exercising the rule-making power ever intended to exclude such evidence, which may very well have been admitted in coroners' inquests in the past." (Devine v. Attorney General for Northern Ireland and another, Breslin v. Attorney General for Northern Ireland and another, 1992).

The above House of Lords appeal case concerning coroners' decisions in Northern Ireland shows that documentary evidence can be admitted into inquest proceedings. It involved some soldiers who had killed three alleged terrorists. The admissibility of the soldiers' evidence was challenged because it was not given in person. The challenge was dismissed on the grounds that if 'documentary evidence' were inadmissible, then suicide notes could not be considered either.

> "Where the act causing death was clearly deliberate the possibility of accident may be excluded and the circumstances may give rise to an irresistible inference of suicide even in the absence of a **suicide note** or a compelling antecedent history." (R v. HM Coroner for Newbury, ex parte John, 1991).

The above quotation shows that when notes (and relevant history) do not exist, verdicts may be still reached by a process of elimination.

The quotation below seems to imply that a suicide note might strengthen the likelihood of a suicide being genuine, but that even a genuine suicide might be enacted to obtain financial compensation for the suicide's family or beneficiaries. In this case, an appeal court judge was building his argument that the police were only partly responsible if a person in custody killed themselves. The "careless oversight" in the quote seems to refer to a prisoner being put in a cell that had a door with a flap that would not close and no glass in its spy hole. The prisoner tied his shirt through these apertures and hanged himself. The appeal was allowed, and previously awarded damages were halved.

> "Or suppose a detainee who and whose family are in serious financial difficulties and who, knowing what the Court of Appeal decided in the present case, says to himself 'the best way for me to help those I love is to commit suicide' and then carries out that purpose by taking advantage of the careless oversight. As Mr. Pannick said in argument, he <u>might even leave</u> a **suicide note** for his wife telling her this." (Reeves Respondent and Commissioner of Police of the Metropolis Appellant, 2000).

The following is part of a discussion about how Article 2 of the European law on Human Rights affects English law in cases of assisted suicide or euthanasia:

> "Indeed, if the article [Human Rights Article 2] did have that effect, it would prevent any state party retaining the offence of suicide. It would also prohibit English law from allowing would-be suicides to be rescued. <u>If a person is found with</u> a **suicide note** and an empty bottle of pills, she could not be taken to hospital...." (R (on the application of Pretty) v. Director of Public Prosecutions and another, 2001, para.43.)

It seems that when judges make points of law where suicide is an issue they take suicide notes into account, but talk in terms of hypothetical notes rather than actual notes. However, because the primary search term was "suicide note" it is not known whether any other points of law about suicide might be made without any reference to suicide notes.

2.3.1.3. Actual Notes: Partial Evidence

Along with R v. James (1998) (see Sections 2.2.2, 2.3.1.4 and 2.3.1.6) and R v. Kavanagh (2002) (discussed in Sections 2.2.1, 2.3.1.3 and 2.3.1.6), here are some other

examples of cases in which lawyers used, or attempted to use, actual suicide notes as at least partial evidence of suicide.

> "[31] In the present case, the question 'by what means' did Mr Middleton meet his death could have been investigated and determined very briefly, on the evidence of the discovery of his hanging body, the **suicide note** and the pathologist's report." (R (on the application of Middleton) v. Western Somersetshire Coroner, 2001).

> "There is no doubt that the note left by Tom Wigley for the applicant had a very profound effect upon the coroner's consideration of his verdict. He clearly came to regard it as a **suicide note**. That was not, however, how the applicant regarded that note..." (R v. HM Coroner for Ceredigion ex parte Wigley, 1993).

In Kavanagh's case, however, (as mentioned earlier in this chapter) the suicide note evidence did not help: he remains convicted of having murdered his wife, who was deemed not to have killed herself.

> "...the defence had relied on the following pointers to suicide;... evidence of neighbours of her volatility of mood; and the **suicide note** found in the bedroom." (R v. Kavanagh, 2002, para.51).

In two of the above three quoted examples the "suicide note"s are listed with other items (via 'and' conjunctions). This shows that it was not the notes alone that clinched the verdicts. However, the converse of this is that it was not any of the other items of evidence on their own that clinched the verdicts either. The existence of the suicide notes seems to have added to the weight of evidence. In the second (Ceredigion) example, by contrast, the note appears to have been the *only* piece of evidence indicating suicide, and as such is open to legal dispute.

2.3.1.4. State of Mind

Each of the cases referred to in this sub-section contains references (italicised) to the supposed state of mind of the author (or possible author) of a suicide note. In most of the examples below, the term "suicide note" is preceded by expressions (underlined) that seem to emphasise the existence of the note. R v. James (1998), (see Sections 2.2.2, 2.3.1.3 and 2.3.1.6), is one example of a note being used as evidence of a victim's state of mind.

> "Mr Treacy QC counsel for the appellant... submitted that 'the Note', as we shall call it, was <u>demonstrably</u> a **suicide note**, written by Sandra James during the ten days or so before her death, and admissible evidence of her *state of mind* at that time..." (R v. James, 1998, p.5-6).

Another example of evidence of state of mind being derived from a suicide note is the following: the first quotation shows a medical expert's opinion; the second shows the judge's conclusion.

> "Dr Cunningham Owens expressed the opinion... He regarded James Cross's illness as 'a very severe major *depression* indeed', and formed the view on the basis of what he called 'slim but compelling evidence', <u>namely</u> the **suicide note**, that latterly there was a *psychotic* element in it." (Cross and another v. Highlands and Islands Enterprise and another, 2000, para.41).

> "<u>The terms of</u> his **suicide note** are not, by themselves, in my view an inadequate basis for the conclusion that his condition was *psychotic*." (Cross and another v. Highlands and Islands Enterprise and another, 2000, para.45).

Some mentions of suicide notes in court cases, of course, merely concern precedents, as in the following reference to R v. Gilfoyle in 1996. In common with the cases in Sections 2.3.1.1 and 2.3.1.2, this case (R v. Taylor) did not itself actually feature any suicide notes. Indeed, it could, arguably, have been placed in a sub-section on its own. It does not have the same type of co-occurrences that the other examples of 'state of mind' have.

> "In this respect it may be compared with the evidence which the court regarded as admissible in R v Gilfoyle [1996]... The evidence in question was that of friends of the deceased, to whom she had made statements to the effect that the defendant had prevailed upon her to write a **suicide note**... The Court of Appeal was of the opinion that this testimony, which had been ruled inadmissible by the trial judge as hearsay, should have been admitted as evidence of the *state of mind* of the deceased...

> We accordingly consider that... the evidence given by DS Ferris concerning the statements made to him by Witness A about the reasons for his fear was admissible." (R v. Michael Anthony Taylor, 1996.)

It seems that suicide notes, and sometimes even the existence of suicide notes

themselves regardless of their content, can be considered by a court as evidence of a person's state of mind.

2.3.1.5. Classification of items as Suicide Notes

I now present some examples of courts classifying notes as suicide notes (or otherwise as in the last example of R v Torney (1997)). In most of these cases the contents of the notes are mentioned, but without the aid of any expert linguistic testimony. There is clear dispute about whether some of the notes were indeed suicide notes, and this is reflected in the phraseology (underlined) which includes items such as 'described... as', 'referred to... as' and 'so-called'. There are also phrases here (italicised) relating to the contents of the notes, sometimes indirectly. These tend to follow the phrase matched by the search term (usually "suicide note"), but sometimes precede it. One example contains only a secondary key phrase, and no primary one, because it was taken from another section of the document returned by the (primary) search.

> "Certain documents have been put in before us. One, which was <u>described</u> before the justices <u>as</u> the "**suicide note**", *being a letter* written by the wife *for the benefit of the coroner* when she apparently attempted to commit suicide unsuccessfully... From that note it is reasonable, I think, to conclude that Rowe, who had himself committed suicide, had written her a letter, which she asserts that the husband saw, justifying to her his act. We have been informed by learned counsel on instructions that Rowe's letter *contained* a sentence which could be interpreted as an admission on his part that he was the father of an expected child and as being some evidence against the wife, inasmuch as she had received this letter apparently without protest and, as I have said, acknowledged its having been *seen* by the husband *in* the "**suicide note**" *of Feb. 8, 1959*. I am not going to say any more about the validity of that evidence..." (Cooke v. Cooke, 1960).

> "...I am satisfied from the notes of the evidence, together with *the contents* of a letter which has been <u>described as</u> the "**suicide note**", that the parties did not come together again until some time after the suicide of Rowe, which took place at the beginning of February, on some date which has not been made clear. It must have been before Feb. 8, 1959, which was *the date* of the "**suicide note**"." (Cooke v. Cooke, 1960).

The above shows how the date on a note can be important. As will be discussed in later chapters, dates on notes, or their absence, can be indicators of events such as a note's being re-used following a previous suicide attempt. However, the date in the above

example does not seem to have been questioned by the court.

"In her affidavit... the mother had alleged that the father was unstable and had left notes threatening suicide. But in fact the learned Judge found that they were not **suicide notes**, but notes which he had made *reflecting his unhappy state of mind* as a result of the relationship between himself and his wife." (Stead v. Stead, 1985).

"The learned Judge deals with the allegation of instability... He says this... 'I have seen this case before, and I have certainly never seen any notes which might have had that construction' -- that is the construction of **"suicide" notes** -- 'nor have I seen nor have I any particulars of, the issue of stability being pursued before me, and that is very important in a custody application...'" (Stead v. Stead, 1985).

The above two quotations from a custody hearing, at least taken in isolation and out of context, show the appeal judge does not appear to be querying the decisions of the judge from the preceding trial concerning the latter's opinions of what is, and what form ought to be inherent in, a suicide note. There is no amplification at all on the "construction" issue itself; there is no information on what a judge thinks is the definition of a suicide note.

"A note (which was exhibit 10) was found by the body. It had apparently been written, not all at one time, by the deceased. It was referred to by a police officer as a **suicide note**. *It contained* its [sic] phrase: 'I cut myself up'. There was reference to a betting slip with #500 to be collected; there was reference to how the writer wished to be buried; and there were certainly other indications in that document that the victim contemplated that he was about to die." (R v. Dear, 1996).

In the above, the references to impending death seem to have convinced both police and the appellant's lawyer that the note was indeed a suicide note. Perhaps this was a significant part in their concept of 'intent'. As mentioned elsewhere, intent is what coroners look for in suicide notes.

"Martina came home in the early afternoon on... 12th June 1991. She was seen later in the afternoon when it was found that she had been electrocuted. She was dead. ... She had left a note. It was dated 11th June, which, in fact, was the previous day, although no one knows whether that had any significance. *She says, and I read, it verbatim*, in her own handwriting:..." (Re N (Minors), 1993).

The above is one of the rare cases in which a suicide note was read out aloud, in full, in court. There was even a mention of the date on the note (see above). The note is quite a striking one because it was written by a child. Some further comments made by the court are given below. I have included a large amount of this text because, as this brief review of Appeal judgments shows, it seems unusual to find both the full transcript of a suicide note and a discussion about it.

> "The Judge went on to say… 'Going back now to the allegations which started these tragic events, that is the allegations of sexual abuse... One has to consider what she said to Mrs Gray and to Mrs Milsom and to the police in the context of her later **suicide note** *where she says*, "Daddy had not done anything to me", and consider her motives and the meaning that ought to be given to that. ...'" (Re N (Minors), 1993).

> "Dr Glazer... was able to say that... what Martina had related to others was typical of a particular type of child abuse as was indeed, also, the <u>so-called</u> retraction statement or **suicide note**, <u>as it was called</u>." (Re N (Minors), 1993).

> "Again, when one looks at <u>what the Judge called</u> the **suicide note** or retraction letter -- <u>call it what you will</u> -- *there were a number of statements in it* which were not consistent with the evidence that was subsequently given. …

> Again, the best friend (who in the note is described as Kelly, but according to the child herself is called Keeley), denied ever having been told by Martina about any of the matters which are *stated in* the **suicide note**, and it is perhaps not insignificant that the matters which Martina *sets out in* the **suicide note** about being pressured by her father to do too much school homework are consistent with what the father told the police at his first interview, even though the interview was something like a month ahead of the <u>so-called</u> **suicide note**." (Re N (Minors), 1993).

> "The further matter relied upon by Mr Bell as detracting from the evidence upon which the learned Judge based his finding that Martina had been sexually abused by the appellant was the note which was said to have been found on the bed, after her death. It was <u>referred to as</u> a **suicide note**, but perhaps the only indication *in the note itself* that Martina, when she wrote it, contemplated any harm to herself was in the very last line, *where she had said*: 'Please Mum forgive me and tell daddy to forgive me.'" (Re N (Minors), 1993).

The suicide note seems to have been adequately addressed by the appeal court, as far as one can tell: it is used as evidence both for and against the perpetrator of the alleged abuse which apparently led to the suicide. As an aside, it is interesting to note the

mention made about the victim's not spelling her best friend's name correctly. This aspect of this note is briefly discussed again in Section 5.7.

Another rare case of a suicide note apparently being read out in full is the following. In this case the transcript of the note succinctly follows its introduction. Because so few notes are read out it cannot be ascertained whether this form of a) 'in location x', b) 'a note was found', c) 'it said y', represents any type of consistent pattern.

> "In due course a search was made of the victim's house. In the bedroom on the bedside cupboard was a **suicide note** in his own hand *which read*: 'To whom it may concern. I am leaving this note to say that I am being blackmailed by the sum of # 74,000 for something which was an accident... With this I feel I cannot carry on, so I am going to finish it tonight. Please contact my brother...'" (Attorney-General's Reference (No 40 of 2002), 2002, para.8).

In the next case, the contents of the note are clearly of some potential importance regarding the motive for the suicide. It also provides the information that suicide notes can be 'long' (although the length is not quantified), and apparently confused (although it is not made clear whether this refers to assumed contradictions in the text, such as having a "clear conscience" despite the attempted murder, or something else). However, as the mention of the "Prosecution" intimates, Campbell did not kill himself, and therefore the attributes of the note cannot safely be assumed to be those of a typical genuine suicide note.

> "He had written a '**suicide note**'. This was long, rambling and confused, but *contained* admissions to which the Prosecution attached importance, such as 'I tried to kill her. God, I don't know what came over me', and subsequently, 'I didn't want to hurt her', and 'I tried to strangle her, Mum. God, Why? Why? My lovely kids. How they would have suffered', and later 'Why I am so relieved I did not succeed to hurt Jane. At least I can go now with a clear conscience'." (R v. Campbell, 1984).

The largely self-explanatory use of capitals, in a subheading in the Appeal Court judgment below, emphasises that there was some dispute about whether the notes were indeed suicide notes.

"THE ALLEGED SUICIDE NOTES

It will be recalled that John thrust two pieces of paper into his father's hands. The Defence case was <u>that these were</u> **suicide note**s written by John but the Crown disputed this suggesting that they were in the nature of apologies, if they had ever been delivered, and were made use of by the appellant as part of his cover up... Resolutions of this issue was very much a matter for the Jury and our shared reaction is that common sense does <u>not</u> suggest <u>that they are</u> **suicide note**s." (R v. Torney, 1997).

The examples in this sub-section show that judges do indeed consider the content of suicide notes.

2.3.1.6. Expert Testimony

It seems that in none of the examples found of the U.K. cases on the LexisNexis database are there any instances of expert testimony by linguists. The experts are usually psychiatrists or handwriting experts, or psychologists. But even these experts often have their opinions discounted by judges. Expert testimony about suicide is mentioned, but often not admitted as evidence, or its arguments discounted by the court, in many cases. Examples of such evidence not being accepted include R v. Kavanagh (2002) (see Section 2.2.1), R v. Gilfoyle (2000) (see Chapter 1 and Sections 2.2.2, 3.1.7, 3.4 and 5.7) and R v. Bamber (2002) (see Section 4.4); an example where expert testimony, albeit psychiatric, was accepted is R v. James (1998) (see Sections 2.2.2, 2.3.1.3 and 2.3.1.4). But whether accepted or not in terms of its opinions, the point is that expert evidence from linguists is only rarely admitted as evidence, at least in regard to suicide notes.

Here the underlined phrases seem to show terms of analysis that were used on the notes, or opinions about them, for example, 'examined' and 'it is said that'. Interestingly, instances of the phrases matched by the search term here are invariably in quotation marks, apparently showing that the notes are disputed in some way. The examples below are somewhat random (within the parameters of my LexisNexis search), but I have included them all to give a good picture of how courts deal with experts in the matter of suicide notes.

As stated above, although this sub-section (like the previous ones) contains examples

from the cases which were found to contain the key word or phrases, some of the examples in this sub-section do not contain any instances of them. They are included because they also contribute to the view of how the courts deal with testimony about suicide and suicide notes. (Interestingly, they contain some negation.)

> "[55]… Professor Canter concluded his review by stating that, for a jury, a decision as to suicide 'would be even more difficult without the assistance of professional guidance.' We mean no disrespect to the Professor when we say that the only guidance he can have had in mind is, as George Bernard Shaw might have put it, 'You never can tell.'" (R v. Kavanagh, 2002).

> "[56] ...we do not consider that Professor Canter's proposed evidence as indicated in his review would have afforded the jury with information likely to have significantly widened their understanding of the difficulties of judging the likelihood or otherwise of suicide... the message… of Professor Canter's review is not scientific information that is likely to be outside the knowledge and understanding of a judge or jury..." (R v. Kavanagh, 2002).

In addition to the views of a handwriting expert, as shown in the two quotations above, R v. Kavanagh had a report from psychologist David Canter. As shown below, R v. Gilfoyle (2000) also featured a handwriting expert along with Professor Canter. (As mentioned above, the bold text is mainly the result of my searches, but I have also made some words bold to emphasise the many different texts with which the court was dealing.) The first paragraph below is also quoted in Chapter 1, but is repeated here to aid the reader.

> "[6] ...In April 1992, he showed her [Sandra Davies] a letter which he said his wife had written to him. This was referred to at trial as **the 'Nigel' letter**... After the deceased's death, other letters were found in notebooks in the house. One **typed letter** had been written about the end of October 1991, a day or two after the appellant had told his wife he had someone else... As a <u>result of ESDA testing</u>, another typed **'suicide' letter**, referred to as **the 'indented' letter** was <u>revealed</u> in a notebook. A handwriting expert said that there was strong evidence that it had been written before March 1992, when some domestic accounts had been written in the same book... **Another note**, of unknown date, hand-written, and addressed 'To whom it may concern', was found in a foot stool in the kitchen. It said 'I Paula Gilfoyle am ending my life. I have taken my own life and I am doing...' (R v. Gilfoyle, 2000).

> "[23] As to Professor Canter,... At trial, the prosecution wished to call him to give evidence in accordance with his report of 25 May 1993, which the judge

ruled was inadmissible. Professor Canter's conclusion at that time, having underline examined the **'suicide' note** and the other notes said to have been written by the deceased, was that the deceased's behaviour was out of keeping with that typical of women who commit suicide and that it was very unlikely that she had written the **'suicide' letter** with the intention of taking her life... In a letter to the CCRC... he changed his mind in relation to all the matters on which he had relied in his initial report..." (R v. Gilfoyle, 2000).

As shown below, the judges in Gilfoyle's appeal hearing were not swayed by Professor Canter's psychological autopsy. I quote just a few of their reasons for this. (Some of the following is also quoted in Section 3.4.)

"[25] … His reports identify no criteria by reference to which the court could test the quality of his opinions: there is no data base comparing real and questionable suicides and there is no substantial body of academic writing approving his methodology." (R v. Gilfoyle, 2000).

"[25] … We very much doubt whether assessing levels of happiness or unhappiness is a task for an expert rather than jurors and none of the points which he makes about the **'suicide' notes** is outwith the experience of a jury..." (R v. Gilfoyle, 2000).

"[25] So far as is known, there have been seventeen occasions in the United States when criminal trial judges have admitted evidence of psychological profiling: in each case the decision has been overturned on appeal." (R v. Gilfoyle, 2000).

Regarding psychiatrist Dr Weir's report the judges said the following.

"[26] His comments on the **"suicide" notes** were, as it seemed to us, in no sense scientific and contained nothing which would not have been apparent to a jury." (R v. Gilfoyle, 2000).

In the case of R v. Bamber (2002), below, as with all the above cases, the judge(s) were not swayed by any expert evidence about suicide notes, and thought the evidence was perfectly comprehensible to a jury. The case of R v. Bamber is also interesting because it featured a book that was a potential suicide note, and my data also included a book that was a potential suicide note (see Section 4.4). This again raises the dilemma of whether or not to call a book a suicide note (see Section 2.2.2).

"405. Ground 9 [of this appeal] is an allegation of non-disclosure relating to the Bible found beside Sheila Caffell's body. The precise complaint is that the prosecution failed to disclose at trial the pages at which the Bible had been opened." (R v. Bamber, 2002).

"408. The pages on which the heavy [blood] staining appears in these photographs are pages including part or all of Psalms 51-55. It is said that these pages are significant and represented "Sheila Caffell's **suicide note**". A number of passages are highlighted and it will suffice if we give one example, taken from verse 14 of Psalm 51: 'Save me from blood guiltiness O God...'" (R v. Bamber, 2002).

"419 ...The pages only became relevant if Sheila Caffell had turned them up but that begged the very question that the jury were going to have to answer, namely who killed Sheila Caffell. Was it Sheila Caffell who opened the Bible as a part of her suicide or Jeremy Bamber who did it as part of a plan to make it look as if Sheila Caffell had committed suicide?" (R v. Bamber, 2002).

In the case of R v. James (1998) (see above) the following was noted.

"Turning to the likelihood of her having committed suicide... he [Dr Nigel Eastman, psychiatrist] regarded the Note as a significant factor: it 'substantially increased', from the psychiatric point of view, the chances that she did so." (R v. James, 1998, p.6).

"The intended meaning of the words written on the Note has not been the subject of forensic debate." (R v. James, 1998, p.7).

In all the above cases it seems that the courts ultimately eschew any expert testimony, find nothing compelling about it, and believe that the evidence is clear enough for a jury to sieve through, unaided by any expert. According to Solan (1998), linguists are not generally welcome as expert witnesses in matters involving the meanings of texts (ibid., p.90). He does not mention suicide notes, but mentions texts such as statutes, contracts and patents in this respect (ibid., p.93). It seems that lawyers are seldom aware of how a linguist might be able to help them (ibid., p.88). Judges think that some texts are matters of law, for only lawyers to interpret, and that juries' native speaker knowledge is sufficient for their understanding of texts (ibid., p.91). Further, as mentioned above (in Section 2.3.1), many of the examples used in this part of the thesis are from Appeal Courts. According to Solan (1988, p.88), in an appeal case "The standard of review is very deferential to the [original] trial judge...", and if expert testimony is excluded from

a trial it is highly unlikely that it will be permitted in an appeal. However, Solan (ibid., p.89) points out that in other situations linguistic testimony is often allowed, for example, in cases involving comprehensibility of legal documents, non-native speaker issues, and trademark disputes.

2.3.1.7. Mentions of Suicide Notes

Sometimes suicide notes seem to get a mere mention by the court, as in the following cases. By "mere mention", I mean that the judgment only yields the expression matching the search term once or twice, or the expression is discussed only briefly in comparison with the other cases. As with the examples given for "State of Mind" (Section 2.3.1.4), the examples here exhibit phraseology which emphasises the existence of the note. However, here the collocates are simple; they are the indefinite article "a". But they include verbs which indicate something that happened to the note, rather than describing its content. These are underlined.

> "He was taken to hospital by ambulance, and his wife followed in her car. She did not complain to the police until 14 July, that is to say, approximately three weeks later and she did so after the offender had posted a **suicide note** through the letter box, in breach of an injunction previously obtained." (Attorney-General's Reference (No 24 of 1999), 1999).

> "On the evening in question, before he broke into her house, he armed himself with two knives and waited outside the house while he wrote a **suicide note**." (R v. Tonge, 1993).

> "But the unhappy mother had only been there for a few days before there was an episode when she inhaled butane gas and the staff had to take over the care of the baby. Not long afterwards she went out, leaving a **suicide note** behind." (Re J (a minor), 1994).

2.3.2. Summary

The purpose of Section 2.3 and its sub-sections was to give a more holistic view of how suicide notes are used by courts than was presented in Section 2.1 and its sub-sections. This study of legal judgments has revealed some interesting information about how the courts treat suicide notes. In searching for occurrences of the term "suicide note" in all the U.K. cases in the LexisNexis legal database, and then looking for references to suicide 'note(s)' and 'letter(s)', some pieces of text relating to the courts' views about

suicide notes were gleaned.

This survey allowed me to collect some information about legal and judicial views of suicide and suicide notes. In the 30 cases found by searching the legal database, nearly a third did not actually feature suicide notes, but they still talked about them. I gave example quotations from most of the cases.

The absence of a note was often held to be significant (2.3.1.1); points of law tended to be argued with reference to hypothetical suicide notes, even in cases where suicide was not directly at issue (2.3.1.2); where actual notes were counted as evidence they were supported by other kinds of evidence (2.3.1.3); actual suicide notes could be taken into account as indicating the state of mind of their writers (2.3.1.4); there was some debate as to whether particular documents counted as genuine suicide notes or not (2.3.1.5); and courts rarely defer to expert linguistic testimony in any of these deliberations (2.3.1.6): judges are very sceptical about experts, and feel that juries can cope well enough on their own.

2.4. Conclusion

In this chapter I have briefly looked at Coroners' Courts, coroners, inquests, verdicts, and the situation regarding various kinds of documents, particularly as they pertain to suicide and suicide notes. I found that the laws and/or rules can be confusing (at least to the lay person), often being qualified by the notions of usuality or the coroner's discretion.

I also considered why defining suicide and suicide notes is difficult. There does not seem to be any definition of suicide that can account for all possible scenarios. Classifying a death as suicide is usually done by a Coroner's Court, which has to find proof 'beyond reasonable doubt', including the element of intent, before reaching such a verdict. It could be argued that in many cases, Coroners' Courts have erroneously returned verdicts of suicide which should have been other verdicts; and vice-versa. For my research data, however, I have adhered to the court's verdicts. Regarding suicide notes, Coroners' Courts consider their location with respect to the dead body, and they

look for dates and intent when assessing their authenticity. For my research, however, I begin with any text written by the deceased that was found in the Coroner's Court files where the verdict was suicide (see Section 2.2.2 and Chapter 4). To do otherwise would be to prejudice my studies whose overall aim, in a sense, is to see what suicide notes are.

Finally I looked at judgments from U.K. courts (other than Coroners' Courts) to see how they used the term "suicide note" and its equivalents.

Of course, this chapter has been an overview of the legal background that surrounds the majority of suicide notes. The defining of suicide notes in general, through a detailed examination of how they are used by their authors, in terms of what is written in them is the essence of this thesis.

3. PREVIOUS STUDIES OF SUICIDE NOTES

This chapter surveys some previous work on suicide notes.

3.1. The Suicide Note Literature

There have been many studies of suicide notes, but the vast majority of these have been from psychological, psychiatric and sociological perspectives. Those which can be considered as linguistic are few in number and mostly decades old. Of these, the best known, and most frequently cited, are Shneidman and Farberow (1957e), Osgood and Walker (1959), Gottschalk and Gleser (1960), Ogilvie, Stone and Shneidman (1966), Edelman and Renshaw (1982) and Gregory (1999). Most previous studies have one other thing in common: they nearly all revolve around one set of notes which were collected by Edwin S. Shneidman (Shneidman & Farberow, 1957e) in the late 1940s - all the note-writers were male, Caucasian, Protestant, born in the U.S.A., married and aged 25-59. The knowledge base which research nourishes deserves an up-to-date study, of a larger database of notes from a wider variety of writers, and it needs to receive this from a linguistic stance, free of the psychological presuppositions about suicidal individuals.

More recent studies have been done, on larger corpora, and with other notes, and some of them are mentioned in other chapters of this thesis, but in this chapter I single out what I consider to be the core general references on suicide notes.

The most frequently cited authors appear to be Antoon Leenaars and Edwin Shneidman. Almost certainly, the most frequently cited study has to be Shneidman and Farberow (1957e), with whom I begin.

3.1.1. E. S. Shneidman & N. L. Farberow (1957)

Edwin S. Shneidman and Norman L. Farberow have, both together and separately, written many papers and books on the subject of suicide. Their 1957 edited book, Clues to Suicide contained an appendix of 66 suicide notes, half of which were genuine, and half simulated (see below). This collection of notes was the most significant element in

the book, and it had a profound influence on many of the studies that were to follow.

Shneidman had originally collected 721 notes from the "Office of the Coroner, Los Angeles County" (Shneidman & Farberow, 1957e, p.198).[5] (I have had access to some, though not all, of these.) The notes were from "the folders of all the recorded suicide cases in Los Angeles County for the ten-year period, 1945 to 1954" (ibid., p.198). Shneidman found that 12-15% of the suicides had left notes (ibid., p.198); and nearly 75% of the notes were written by males (ibid., p.199). The 33 genuine notes referred to above were a sub-set of these.

The simulated notes were

> "obtained from *nonsuicidal* [their italics] individuals contacted in labor unions, fraternal groups, and the general community... Each individual was given a personality questionnaire and was interviewed briefly. If he indicated any signs of personality disturbance or tendencies toward morbid content of thought... he was not given the suicide-note task..." (Shneidman and Farberow, 1957e, p.199).

The authors do not give the total number of simulated notes which were collected, nor do they say whether any were from women or whether a slightly less stringent matching process might have resulted in more than 33 pairs of notes. All they say on the matter is the following.

> "In order to keep the two groups homogeneous (and to emphasise whatever differences might exist in the notes), all 66 genuine and simulated notes were selected from those written by individuals who were male, Caucasian, Protestant, native-born, and between the ages of twenty-five and fifty-nine. In addition, each of the 33 simulated-note writers was matched, *man-for-man* [their italics], with a genuine-note writer who was not only of a similar chronological age (within five years) but also of the same occupational level.", (Shneidman and Farberow, 1957e, p.199).

Shneidman and Farberow (1957c) analysed the 66 notes using a method called "Discomfort-Relief Quotient (DRQ)" (ibid., p.7) which they attributed to a 1953 study

---

[5] In Shneidman (1976, p.266), however, the figure given is 906 notes.

by Mowrer. This involved counting "thought units" (ibid., p.7). These are "discrete idea[s]" (ibid., p.7), independent of the number of words used to express them, and so, in a sense, this was a discourse based analysis. The sense is that they looked at pieces of text, rather than only words, but without concern for any explicit linguistic entities within them. The thought units were categorised depending on whether they expressed discomfort (e.g. "guilt, blame, tension, aggression" (ibid., p.7)), relief (e.g. "love" (ibid., p.7)), or neither discomfort nor relief - which they termed "neutral".

Shneidman and Farberow found that there were more thought units in the genuine notes than in the simulated notes. They found no differences between the two sets of notes in the relief category. They found the discomfort in the genuine notes was more intense than in the simulated notes. They found that the genuine notes had far more thought units in the neutral category than did the simulated notes, and that these were mainly "instructions and admonitions" (ibid., p.7). Shneidman and Farberow say that the latter shows that the genuine note-writer might be "confused and contradictory" (ibid., p.8): on the one hand, being aware that he shall not exist in the future; and on the other, imagining that he shall be able to oversee events after his death.

As noted above, Shneidman's 66 notes were used in many subsequent studies by other authors. Some more of his research findings are mentioned below in Section 3.1.4.

3.1.2. Osgood & Walker (1959)

Osgood and Walker looked at suicide notes versus ordinary letters to friends and relatives as well as genuine versus simulated suicide notes. As Table 3.1 (below) shows, they had a) 100 notes from Shneidman's original set of genuine notes, b) "a sample of ordinary letters written to 100 members of a panel...", and c) Shneidman's 33 pairs of genuine and simulated notes (Osgood and Walker, 1959, p.59). (It is not clear whether there was any overlap with the two sets of Shneidman notes.) However, because of their requirement to have a minimum of about 100 words per note to suit the quantitative analyses they wanted to perform, the number of notes/letters they actually used was a) 40 from males and 29 from females, b) 13 from males and 59 from females, and c) 13 pairs, then 18 simulated and 24 genuine notes, respectively. Regarding the

source of the ordinary letters, they said, "This panel had been used for other purposes in connection with research on the communication of mental health information" (ibid., p.59), and so it is not entirely clear whether or not these letters were by potentially suicidal people. They did not include any of the full texts in their article.

Table 3.1. Osgood and Walker's Data

	Shneidman's 721 Original notes	Ordinary Letters	Shneidman's 33 paired notes
Had	100	100	33
Study 1	40 male + 29 female	13 male + 59 female	
Study 2a			13 pairs
Study 2b			24 genuine + 18 simulated

Osgood and Walker performed, basically, two studies. First, they compared suicide notes with ordinary letters using 16 measures derived from 4 hypotheses (see below). Second, they compared the genuine and simulated paired suicide notes with each other using initially 9, then 11, of the aforementioned 16 measures that seemed potentially useful in light of the results of the first study (see below). They also did some word frequency analysis on the paired notes (see below). Their methods seem to have been manual rather than computational. It is not entirely clear, but overall, they appear to have used a mixture of both individual words and phrases or sentences as their discourse units depending on the specific category they were dealing with at the time (see below).

They had four hypotheses. To paraphrase these: they expected suicide notes to show 1) greater stereotypical behaviour, 2) greater disorganisation, 3) greater frequency of criticisms, demands and requests, and 4) more conflict terms than the ordinary letters. The "measures" (as they call them) they used to test the hypotheses are listed below. (Much of the terminology used (above and below) is taken directly from Osgood and Walker, and some of it is paraphrased.)

1) Stereotypical behaviour: Osgood and Walker measured this by the average number of syllables per word; the type/token ratio (see Section 4.8); the proportion of repetitions of

two or more words in phrases; the noun-verb/adjective-verb ratio; and "allness" terms (such as "always, never, forever,... everyone, completely, perfectly..." (Osgood and Walker, 1959, p.60)). They also applied cloze procedures, which involved blanking out every fifth word in 20 suicide notes and in an equal number of ordinary letters, and then seeing whether subjects could guess correctly what the missing words were (on the basis that suicide notes should be more predictable than non-suicide notes)(ibid., p.60).

2) Disorganisation: measured in terms of structural disturbances ("grammatical, syntactical, spelling, and punctuation errors, and... awkward constructions. Points where material was obviously omitted..." (ibid., p.61)); and average length of simple sentences (derived from compound ones where they existed).

3) Directive State Measures (which, apparently, are concerned with "grammatical and lexical choices associated with the motives leading to self-destruction" (ibid., p.58-59)): distress/relief quotient (as used by Shneidman and Farberow (1957e) (see Section 3.1.1 above), but this time, apparently from a 1947 paper by Dollard and Mowrer this is "the ratio of distress-expressing phrases to the sum of these plus relief-expressing phrases..." (Osgood and Walker, 1959, p.61)); number of evaluative common meaning terms (such as "unfair, dangerous, sweet-heart, and drunkard..." (ibid., p.61)); positive evaluative assertions about people (such as 'I have always admired you' (ibid., p.61)); time orientation; and "mands", from Skinner (1957), a mand being defined by Osgood and Walker as "an utterance which (a) expresses a need of the speaker and which (b) requires some reaction from another person..." such as "Don't feel too bad about this" (Osgood and Walker, 1959, p.62).

4) Conflict terms: qualification of verb phrases (which seems to have involved counting modals and auxiliaries, and any elaborations to a simple verb or tense); ambivalence constructions (such as "but, if, would, should, because,... possibly, seems,... surely, really, except..." (ibid., p.62)); and ambivalent evaluative assertions (such as "I love you, Honey; You never trusted me;..." (ibid., p.63)).

The 9 measures initially used in the second study, with the 13 pairs of notes, were all

the ones above for "stereotypical behaviour" except for cloze procedures (i.e. 5 measures), all the ones for "conflict" (3), and mands. They then added structural disturbances and sentence length to their 9 measures, and used 24 genuine notes and 18 simulated notes from the paired set, i.e. "all notes of sufficient length, regardless of their pairing..." (ibid., p.63), and obtained similar results.

In distinguishing suicide notes from ordinary letters they found that they were correct in three of their hypotheses: suicide notes showed greater stereotypical behaviour, greater frequency of 'Directive State Measures' except for 'positive evaluative assertions' which were fewer (see above), and more conflict terms. But suicide notes did not, apparently, show greater disorganisation (ibid., p.66). Moreover, the suicide notes written by males used significantly longer sentence segments compared with "ordinary letters" (ibid., p.61).

In distinguishing genuine from simulated suicide notes, however, their results were similar but less conclusive, with only a few results being statistically significant, or nearly so. It is not clear, but it also seems that the genuine notes had shorter sentence segments than the simulated notes. The following quotation summarises all this.

> "The person faking a suicide note fails to reflect the demanding, commanding, pleading style (*mands*), the reduced qualification (*noun-verb/adjective-verb ratio*), and the evaluative ambivalence toward self and others of the genuine suicide notes." [their italics] (Osgood and Walker, 1959, p.66).

They also performed a word frequency analysis on lexical words by subject. That is, for each lexical (i.e. excluding functional) word type (rather than token) counting only the number of subjects (note-writers) whose texts contained one or more occurrences of that particular word. They noticed, amongst other things, the more frequent use of 'mother' terms, and of 'terms of endearment' in the genuine notes (ibid., p.64). And finally, they tried looking at some categories of words, such as 'job' and 'parents', to see whether there were any 'associations' between the categories, but their results were not conclusive.

It was, perhaps, their penultimate analysis that was most interesting. Counting by

subject (or by text) is an important technique, often overlooked in corpus studies. It means that results are not skewed by an individual subject (or text). It is also in this part of their paper where they say (to paraphrase them, and with their italics) that simulated notes have more mental state verbs such as *'know'* and *'think'*, and more abstract concepts such as *'life'*, whereas genuine notes have more simple action verbs such as "*tell, do, get*" (ibid., p.64), more positive concepts such as *'love'* and *'hope'*, more "ambivalence towards loved ones through high frequency of *positive evaluative terms*", and more concrete concepts (ibid., p.64). The concept of thinking versus doing (related to mental versus action verbs, and to abstract versus concrete terms) would reverberate throughout much of the suicide note literature in the years following Osgood and Walker (1959).

3.1.3. Gottschalk & Gleser (1960)

Gottschalk and Gleser were psychiatrists. They had been engaged in "measuring a wide variety of personality variables through... verbal behaviour" (Gottschalk and Gleser, 1960, p.195). These included "gender and intelligence quotient" and "various psychiatric syndromes" where they were concerned with "verbal content variables" in speech. Their 1960 study was their first on written language. They used Shneidman's 33 real and 33 fake notes. Gottschalk and Gleser performed a word-type frequency analysis in which "Each word in the notes was categorised" using both a traditional grammatical and a psychological category.

Examples of their grammatical categories were "adverbs" and "verbs" (ibid., p.196). Examples of their psychological categories were "words denoting feeling or motivation (e.g. 'happiness, anger, need')" (ibid., p.196) and "words indicating a perceptual process (e.g. 'look, itch, sensation')" (ibid., p.196). They also had psychological categories for "words that refer to different kinds of objects or agents" (ibid., p.196) such as "self ('I, my')" (ibid., p.196), and "things ('house, money')" (ibid., p.196). And they had a 'dustbin' category they called "auxiliary words" (ibid., p.196) for words, such as some "prepositions and conjunctions" (ibid., p.196), that did not fit into any of the other categories. Some of the categories were sub-divided further, giving a total of 34 categories for the words. (The exact composition of all of the categories is not entirely

clear.)

Having marked up the notes, and calculated the percentage of occurrences of words in each category, per note (ibid., p.196), they performed two procedures. Presumably, most if not all of this was done manually. They began by using a sign test on the matched pairs of notes, and then used a median test on each of the sets of notes (simulated and genuine) (ibid., p.197). A sign test is a test in which a binary decision is made. Thus, a plus or minus sign is assigned to an individual result and then all the plus and minus signs are counted. In this case, Gottschalk and Gleser seem to have given a plus sign to whichever text in a matched pair contained the greater number of words that represented a category. A median test involves putting results in numerical order and taking the middle one, or, if there are an even number of results, taking the average (the 'mean') of the two middle results. Gottschalk and Gleser seem to have taken the median of all the simulated set of notes, and the median of all the genuine set of notes, and for each set compared counts of words in a category with the median for that category for the relevant set.

Gottschalk and Gleser found that simulated suicide notes averaged fewer words than genuine suicide notes, but state that there was no statistically significant difference in the medians of the word totals between the two sets of notes (ibid., p.197). Then they found that only 5 of the 34 categories proved to be significant in distinguishing the genuine from the simulated notes regardless of the test (sign or median). (The third category in the following quotation is a combination of addressee and 'third person others'.)

> "These [five categories] were: the percentage references to substantives (nouns and pronouns); to others (third person); to others including the person addressed in the note; inanimate objects; and the percentage use of auxiliary words." (ibid., p.197).

The sign test on the paired notes found "prepositions, conjunctions, and words implying the process of cognition" (ibid., p.197) to be significant. However, as they remark, matched pairs of notes would not have much practical use in distinguishing real from fake notes in the real world (ibid., p.197). They mention only the aspect of lack of

generality in this regard; and I comment on it in Section 3.2 (below). As for the median test...

> "Two categories differed significantly by the medians test but not by the sign test: percentage references to the persons addressed in the note and percentage references to places or spatial relations." (ibid., p.197).

They then obtained an eleventh category, called 'objects', by combining all the categories concerning people and objects (ibid., p.197). Additionally, they produced a table of their most useful categories which showed that, at least for the Shneidman texts, genuine notes could be distinguished to some extent by combining the results in particular ways. For example, counting "Auxiliary words minus others (2nd and 3rd person)" in a note, and finding this to be less than 14.00% of the word count, would distinguish 27 genuine and 2 simulated notes (ibid., p.198). As outlined above, each of their final eleven categories was statistically significant in one way or another: some were so in terms of median tests and others in terms of sign tests. However, because some of these tests involved comparisons between pairs or groups of notes where half were known to be genuine and half simulated, it cannot be said that these techniques could be reliably applied to single notes of unknown authenticity.

They then went on to compare 37 suicide notes they collected from Hamilton County, Ohio, Coroner's office for 1954-1957 with Shneidman's. The Ohio notes included 10 written by women. Generally, the comparison was successful and similar results were obtained, although they found that prepositions and conjunctions did not discriminate the Ohio real notes from the Shneidman fakes, and the "auxiliaries" now did so to a lesser extent. Thus, at this point, including the combination categories of 'others' and 'objects', they ended up with nine useful categories. They also found that the women exhibited fewer "references to places or spatial relations" (ibid., p.199).

Gottschalk and Gleser's legacy seems to be that they found that genuine notes have,

> "A relatively high percentage of references to people and things, places or spatial relations, and a relatively low percentage of references to cognitive processes." (ibid., p.200).

It is difficult to say whether Gottschalk and Gleser were more successful than Osgood and Walker (1959) in distinguishing real from fake notes. Gottschalk and Gleser echoed Osgood and Walker's results on thinking versus doing in the former's 'cognitive processes' category; and, similarly, the concrete versus abstract phenomenon can be seen in the concrete 'people' and 'objects' categories.

3.1.4. Ogilvie, Stone, & Shneidman (1966)

Ogilvie, Stone, and Shneidman (1966) applied computer analysis to Shneidman's 66 notes. They used the General Inquirer software with the "Harvard II Psycho-sociological Dictionary" (ibid., p.528). (Ogilvie et al.'s study was actually a chapter in Stone et al. (1966) which describes the software and the dictionary.)

The software accessed the dictionary in order to allocate the individual words into the various categories that were identified in the dictionary. The dictionary itself, as its name suggests, was not a general, nor a linguistic, one, but contained categories pertinent to psychology and sociology. The dictionary included some complicated "disambiguation" rules to give some degree of context to the words, but Ogilvie et al. do not elaborate on this within their chapter-article.[6]

This approach is based on "Content Analysis" (the details of which are not directly relevant to this thesis); however, considering the nature of their categories (see below), it could also be described as what is nowadays called "semantic tagging" (Wilson and Thomas, 1997). It is probably arguable that Gottschalk & Gleser (1960) (above) and Edelman & Renshaw (1982) (below) could also be early examples of semantic tagging. The main achievement of Ogilvie et al. (1966) was their successful application of the computer to the problem of distinguishing real from fake suicide notes, and the successful application of automatic semantic tagging (although with respect to the other contributors to Stone et al. (1966), they were not alone in the latter achievement).

---

[6] However, the General Inquirer user guide for the comparatively more recent version of the software (Buvac and Stone, 2001) refers to Kelly and Stone's 1975 book, "Computer recognition of English word senses", which apparently explains the disambiguation rules.

The dictionary or "category system... was divided into first-order tags (discrete, independent categories) and second-order tags (nonindependent categories)." (Ogilvie et al., p.528). 'Tags' in their usage seem synonymous with 'categories'. "The first-order distinction refers to the primary explicit, denotative meanings of words." (Stone et al. 1966, p.173); whereas "second-order tags are designed to transcend the object-action basis for classification used in the first-order section of the dictionary and to 'fill out' additional levels of meaning." (ibid. p.183). The eight major categories were: the first order ones of "Roles" and "Objects" which together comprised "things", and "Emotional States" and "Actions" which together comprised processes; and the second order ones of "Institutions", "Statuses", "Qualities" and "Symbolic Referents" [my word-initial capitalisation] (Ogilvie et al., 1966, p.528). The eight categories were each composed of sub-categories. For example, 'Roles' included the sub-categories of 'self' and 'other', 'Actions' included 'think' and 'sense', 'Qualities' included 'bad', and 'Symbolic Referents' included 'understate' (ibid., p.529).

The sub-categories that they found to discriminate between the real and fake notes, giving a difference of 0.03% or more in word-counts between the sub-corpora, were as follows. For the genuine notes: self, other, male-role, female-role, artefact, place, communicate, possess, get, attack, higher-status, quantity-reference, and bad. For the simulated notes: selves, anxiety-fail, anxiety-unable, distress, think, sense, if, not, move, avoid, direct, goals, academic, death-theme, and understate (ibid., p.529). These tags are mostly self-explanatory, which is useful since, with a few exceptions, Ogilvie et al. did not elaborate on them. One such exception follows below.

Each of the sub-categories consisted of the words that the General Inquirer software counted. For example, 'think' "included words such as *think*, *recall*, *reason*, *remember*, *explain*, *consider*, *decide*, and so forth." [my italics] and know, knew and known (ibid., p.533). In addition to all this, it seems that they marked the texts for parts-of-speech, such as pronouns, and for grammatical function, "subject, verb, object, and so on" (ibid., p.530).

To paraphrase Ogilvie et al. (1966), they found, with their 0.03% difference procedure (see above), that the genuine notes had no "Emotional States" or "Institutions", but had higher proportions of 'things', i.e. "Roles", "Objects" and "Qualities", than the simulated notes; and the simulated notes had no "Objects", "Statuses" or "Qualities", but had relatively more 'processes', i.e. "Emotional States" and "Actions", than the genuine notes (ibid., p.528-529). Furthermore, when mentioning themselves and others, the genuine note writers tended to refer to themselves in the first-person and to others separately, as opposed to the simulated note writers who tended to use the word "we" (Ogilvie et al., 1966, p.530). Among the words that the software could not categorise, they found that the genuine note set contained far more "proper names, places, objects, numbers, and time" (ibid., p.530) references than did the simulated set. This seemed to confirm that the genuine notes contained "very specific, concrete references" (ibid., p.530).

They also found that in the genuine notes there were far more occurrences of "females in the *subject* [their italics] position of the sentence..." (ibid., p.531). On further examination they discovered that the genuine notes actually had fewer instructions to women than did the simulated notes, but they had a higher proportion of references about women. (An invented example of an instruction is 'Mary, get me cremated not buried', and of a reference is 'She loved dancing'.) As Ogilvie et al. acknowledge, the latter is not surprising (due to the greater number of words in the genuine note set, and presumably because the notes' authors were all male). They subsequently found over half of the instructions given to women in the genuine notes were "specific and direct" (ibid., p.532), whereas 75% of such instructions in the simulated notes were "of a vague, noninstrumental nature." (ibid., p.532).

To investigate the greater number of 'processes' tags in the simulated notes they looked into the 'think' category (see above). By separating this category into words that were indicative of either reasoning (e.g. 'think') or knowledge (e.g. 'know'), they posited that the genuine notes tend to show "that a final decision has been made" (ibid., p.533), whereas the simulated notes tended to contain evidence of on-going "problem-solving" (ibid., p.533). Some of the figures they give for this are: 19 genuine versus 8 simulated

notes employed versions of the word 'know' (ibid., p.533).

Finally, they (or rather one of them, Stone), looked at about half of the notes, in their matched pairs, and created a statistical model. Applying the other 'half' of the pairs, they were able to successfully discriminate 30 of the 33 (ibid., p.534-535). However, as they admit, working without matched pairs is another matter (ibid., p.535). Nowhere do they actually say exactly how many notes they could distinguish as being genuine or simulated when viewing them as two, non-paired, sets, and using more than one group of variables. The nearest they get to this is when they discuss the bifurcation of the 'think' category (see above). More interesting, however, was a by-product of this part of their study: in the genuine notes they found more instances of the word "love".

Ogilvie et al. (1966, p.533-534) noted that their findings concurred with those of Gottschalk and Gleser (1960), adding that they suspected even greater specificity in the "entire final communicative attempt" (Ogilvie et al., 1966, p.534) in genuine notes than had been previously noted. They summarise,

> "More often than was true of simulated notes, genuine notes contained specific information, used names of people, places, and things, made frequent mention of women, and gave instructions to others that were concrete enough to be actually carried out. By contrast, the simulated suicide notes contained a greater percentage of 'thinking' words, suggesting that the issue of suicide was being pondered, reasoned with, and probably rationalised." (Ogilvie et al., 1966, p.535).

Ogilvie et al.'s findings seem to echo those of previous studies. However, their main achievement compared with previous authors was in their methodology. Their use of computer software for semantic tagging was a major step forward for linguistics.

3.1.5. Edelman & Renshaw (1982)

Edelman and Renshaw (1982) reviewed Osgood and Walker (1959), Gottschalk and Gleser (1960), and Ogilvie et al. (1966) and tried to pull together all their categories, and to "replicate" their findings (Edelman and Renshaw, 1982, p.107). Amongst other things they wanted to use a "unified theory of discourse analysis" (ibid., p.105), and to create a profile of suicidal language (ibid., p.104). The latter is also one of the aims of

this thesis.

Edelman and Renshaw employed computer analysis on Shneidman's 66 paired notes. Their software was "Syntactic Language Computer Analysis, Version III (SLCA-III)" (ibid., p.105). They cite Cummings and Renshaw (1976 and 1979) as its authors (ibid., p.105).

SLCA-III uses "dictionary 'look up' plus positional analysis" (ibid., p.107), and "analyzes a message according to how... categories are applied to each word" (ibid., p.107). This sounds similar to Ogilvie et al.'s (1966) General Inquirer software (see above), but it is not clear what, if any, differences there are between the two programs.

> "The system utilizes traditional-grammatical categories, but concentrates on differing aspects and functions of these categories within the analysis of the message as a whole rather than specific content categories such as the TTR or DRQ.", (Edelman and Renshaw, 1982, p.106). [See 3.1.1 for DRQ.]

So, although they looked at words, their focus was on the message. However, it is not clear what constituted a message, i.e. whether it was a sentence, paragraph, non-contiguous set of phrases, whole text, or anything else. They appear to be applying a form of Discourse Analysis, but one which is not well-defined.

Edelman and Renshaw are not entirely clear about their categories, but the gist seems to be as follows. SLCA-III has 36 categories (ibid., p.108). There are three major "categories" which are grammatical (nouns, verbs, and adjectives/adverbs), and "basically eight" sub-categories which are 1) social perception, 2) sensation, 3) existence, 4) motion, 5) disposition, 6) time, 7) symmetry, and 8) definition (ibid., p.106-107). The sub-categories "apply in differing ways" to the major categories, forming the 36 lowest level categories. The following descriptions of the sub-categories are quoted or paraphrased directly from Edelman and Renshaw, and give the category name, its meaning, and its usual realisation (as 'N' nouns, 'V' verbs, 'A' adjectives and 'D' adverbs), and in some cases real-world examples, respectively.

1) Social perception: "ability to perceive and attribute characteristics"; N

2) Sensation: concrete or abstract; N, A, D; 'book' vs 'truth'

3) Existence: Exists or not; N, V, A, D

4) Motion: activity or state; V

5) Disposition: assertion or condition; V; 'I will' vs 'I could'

6) Time: past, present, future; V

7) Symmetry: intentionality; V (transitive or not)

8) Definition: modified or not; N, V. (Edelman and Renshaw 1982, p.106-107).

Edelman and Renshaw performed various statistical analyses which initially confirmed the findings of the previous studies (ibid., p.109), and ultimately suggested 9 of the 36 categories were highly relevant for their model (ibid., p.110). The categories were "Total Word Usage", "Qualifiers", "Negative Relation (Negative verbs)", "Positive Authority (Proper Nouns)", "Negative Authority (Proper Nouns)", "Defined Relational (Modified Verbs)", "Symmetric Relational (Transitive Verbs)", "Future Time", and "Conditional" (ibid., p.111). They claim an 80.3 per cent rate of accuracy in determining whether or not each of Shneidman's notes was genuine or not (ibid., p.110).

Their findings seem to be that genuine suicide notes contain: more references to concrete objects; more negative, concrete modification of "objects and actions or states of objects"; negative references to "specific people, places or things..."; a tendency for "the concept 'you'...", where present, to be negative; more positive references to generalised 'others'; fewer references to future time; and greater conditionality of objects (ibid., p.112).

Compared to the previous studies mentioned in this chapter, Edelman and Renshaw (1982) found aspects of negativity. They, like their predecessors, found more references in genuine suicide notes to people, places and objects, and they found them to be specific and concrete, but they improved upon these results by finding them also to be negative - except for references to people in the third person which tended to be positive. Another advancement was finding fewer references to the future.

It is possible that the references to 'you' along with future time might indicate the presence of instructions in the texts, but having negative 'you' and fewer, rather than more, future references seems intuitively to lessen this possibility. As mentioned above (in Section 3.1.2), Osgood and Walker (1959) had looked at time but found nothing discriminating in that regard. This contradiction with Edelman and Renshaw's finding is surely due to some particular aspect of the methodologies employed, although it is not clear what this might be as much of the paper is shrouded in statistics.

3.1.6. Antoon Leenaars (1988)

Antoon Leenaars is a Canadian psychologist who has written much on suicide. A recent newspaper article (Anon, 2000) describes him as having over 2,000 notes; indeed, he has confirmed this (Antoon Leenaars, personal communication to Malcolm Coulthard).

Leenaars (1988) derived ten protocol sentences - statements describing some potential content of suicide notes - from each of the theories of ten "suicidologists". The latter included Freud and Shneidman. Leenaars then asked psychology graduates to match the 100 protocol sentences against Shneidman's 33 pairs of genuine and simulated notes (Leenaars, 1988, p.59). (The matching process was a manual one, not susceptible to computerisation.)

This method is quite different from those of the other studies in this chapter. The source of the protocol sentences came directly from studies about suicide rather than what seem to be miscellaneous psychological studies of stress, depression and mental illness. Also, having sentences that contain precise details to match against texts might allow a greater accuracy in the matching. Here is an example of a protocol sentence about illness and rejection (see below) that he attributes to Shneidman.

> "In the suicide note, the person communicates evidence of adult trauma (e.g., poor health, rejection by the spouse, being married to a competing spouse) (Shneidman)." [sic] (Leenaars, 1988, p.177).

23 of the protocol sentences were represented in over two-thirds of the genuine notes, and Leenaars concluded that these sentences were therefore "predictive" (Leenaars,

1988, p.172).   He then found that only 18 of the sentences were significantly more frequent in the genuine notes (ibid., p.175).   Ultimately, Leenaars found that only five sentences were both "highly predictive", of suicide notes in general, and also "differentiating" between genuine and simulated notes (ibid., p.175).   The five sentences were concerned with such topics as 1) pain, 2) "forlornness", distress and grief, 3) illness and rejection, 4) contradictions, and 5) "underdeveloped personality organisation", (ibid., p.175 & 177-178); and Leenaars categorised them as "Situation; Relationship; Emotional State; Cognitive State"; and "Ego", (ibid., p.178).

Leenaars went on to test his results against some more recent notes.   He compared the Shneidman pairs from 1945-1954 with some notes from 1983-1984.   These notes had also been collected from the Los Angeles Coroner's Office by Shneidman (ibid., p.195) who, fortunately for Leenaars, was his "friend and 'teacher'" (ibid., p.7).   Leenaars took "the first 33 suicide notes (of over 100) that matched the age range (25-59) and sex (i.e. male)" with Shneidman's 'original' 33 genuine notes.   He also took "20 suicide notes written by females", which apparently represented all the notes by females aged 25-59 in the years 1983-1984, (ibid., p.231).   He found that his "clues to suicide are generally ubiquitous across time..." (ibid., p.195).   However, in accounting for both eras, by reviewing his protocol sentences his five categories (see above) became seven, now including the wish to die and "poverty of thought" (ibid., p.197-198).

Leenaars then went on to make age and sex comparisons and concluded, "We believe that suicide is generally ubiquitous across sex and age although some 'differences' exist, depending not on sex but on age." (ibid., p.206).   He had three age groups: 18-25, 26-54 and 55-77 (ibid., p.201).   The differences included 18-25 year olds being more "psychologically perturbed" (ibid., p.204).

However, in all this, other than what has been said above, it is not clear exactly how many of the genuine and simulated notes were actually differentiated by his sentences. Of his attempt to "define suicide through a study of suicide notes" (ibid., p.207) Leenaars (1988, p.207) stated,

> "It [suicide] is constituted, at least, by a painful situation, a disturbing relation, an overpowering emotional state, a constricted cognitive state and the person's own ego... It is perceived, consciously and unconsciously, as the best solution to a problem or injustices."

He concluded, "We cannot completely define suicide." (ibid., p.211).

3.1.7. Adam Gregory (1999)

One of the more recent authors on the subject of suicide notes is Adam Gregory. Gregory (1999) wanted to create a statistical model that would combine together all the variables he wanted (see below), and to improve upon the results of previous studies. He used the 66 note Shneidman corpus, although he did not reproduce this, plus another 18 genuine notes which he found in various publications, which he did not name. He seems to have tagged the notes manually rather than with a computer. He does not elaborate on his discourse units, but it seems clear that both words and phrases and/or sentences were used.

Gregory gives a useful summary of other studies of language in suicide notes, principally, Osgood and Walker (1959), Gottschalk and Gleser (1960), and Ogilvie et al. (1966), along with Edelman and Renshaw (1982). He distilled nine variables from the existing literature, and made hypotheses about each of them. His variables and the expected relative values for genuine notes when compared with simulated notes were: 1) percentage of nouns – higher; 2) percentage of verbs – higher; 3) average sentence length – lower; 4) number of words – higher; 5) percentage of cognitive process verbs – lower; 6) amount of positive affect – higher; 7) amount of instructions left – higher; 8) reasons or explanations given - none or very specific; and 9) locus of control - external to the author. He found that for the purpose of discriminating real from fake notes, the 'content' orientated variables (4 and 6-9 above) were more useful than the 'structure' orientated ones (1-3 and 5 above). (He was aware that the total word count variable is actually a structural one, but included it with the content items because it scored highly.)

Gregory was successful in building his model. Of the notes that displayed the more 'genuine' attributes, that is more occurrences of the 'content' variables (and more of the structure ones), 15 were genuine and 2 were simulated (Gregory, 1999, p.145); yet this

amounts to only 15/66 notes being correctly classified. When his other 18 notes were added to the analysis, the same procedure seems to show that 18 genuine, and 2 simulated, notes have the most 'genuine' attributes; but this is only 18/(66+18), i.e. 18 out of 84 notes classified as genuine (ibid., p.150-153). Also dubious, but from a linguistic point of view, albeit an intuitive one, is his suggestion that genuine notes tend to have a higher word count because they tend to include detailed instructions and plenty of affect, and that these "operate at the word level" (ibid., p.149), whereas expressions about control and explanations are "thematic" (ibid., p.148) - unless, of course, this has something to do with the way he tagged and counted the variables. (He may have meant that his discourse units for instructions and affect were words, whereas for control and explanations they were perhaps phrases or sentences.)

Gregory (1999) cites Casey (1993), who describes the case of Eddie Gilfoyle whose wife, Paula, was found hanged in June 1993 (see Chapter 1). The police found two 'suicide' notes, but Eddie was subsequently convicted of murder, apparently having persuaded Paula to write the notes by claiming it would be helping him with a work-project on suicide. Gregory warns of the risks in applying his (Gregory's) model to such cases (which he actually did not do) and the need to take into account the psychological state of mind of the murderer - which may have similarities with a suicidal state of mind, or at least dissimilarities with a 'normal' state of mind as attributed to Shneidman's simulated-note writers which were at the basis of Gregory's analysis.

Like many of the authors mentioned above, Gregory is a psychologist, and he is seeking underlying psychological theories about suicide (or conversely, seeking to prove or improve such theories) from linguistic evidence. However, he is included in this review because he has made some interesting comments, as shown above.

3.1.8. 'Popular' Suicide Books
As well as the academic publications, suicide attracts the attention of quasi- and non-academic authors. I class the book by Marc Etkind (1997) as one such 'popular' suicide book. Etkind (1997) includes many suicide notes, and refers to studies by academics

including Shneidman and Leenaars. Although these studies are not as adequately sourced and cited as one would expect in an academic work, Etkind does give some interesting round-ups of various research.

According to Etkind (1997), suicide notes - as opposed to other notes - often favour the use of "absolute words like 'never' and 'only'" (ibid., p.viii). Etkind characterises suicide notes as containing, typically, a) "distorted logic", b) tunnel vision, c) assignment of fault to others and seeing the note's author as a persecuted, innocent victim, and d) "a sense of weariness with life...", (ibid., p.85). In his assessment "Notes with mundane instructions are fairly typical... almost one-third of all notes contain practical requests..." (ibid., p.14). Indeed, "One of the steps on the road to suicide is obsession with seemingly small details" (ibid., p.93). Suicides' "tunnel vision" - logic which is devoid of alternative possibilities to suicide - is often illuminated by their use of "phrases like 'this is the only way' and 'I must do this'" (ibid., p.33), and there is a "detached, unemotional style that is common in suicide notes" (ibid., p.39). Etkind states that "feelings of self-loathing, disgrace, love, and hate... permeate... notes" (ibid., p.89). It is not altogether clear how he considers a "detached, unemotional style" to be capable of expressing "feelings of self-loathing, disgrace, love and hate".

Most of Etkind's findings seem to be derived from the research of others. Many are acknowledged, but some are not explicitly attributed. However, it is quite possible that some of the findings come from his own review of the suicide notes in his book (which were all gathered from previous publications). The latter findings include the "unemotional style", expressions of fault, weariness, hate and disgrace that are, according to Etkind, prevalent in suicide notes. Whoever first made these observations, they are worth repeating here because they have not been mentioned in the previous sections of this chapter.

3.1.9. Suicide Note Literature Summary

As noted by Shneidman (1979, p.151), over twenty years after his 33 pairs of notes were published, a range of studies have looked at various aspects of notes, including "logical styles" and "language characteristics" using methods including "computer count[s] of

67

key 'tag words'". He states that,

> "In general,... genuine suicide notes are characterised by dichotomous logic, greater amount of hostility and self-blame, use of very specific names and instructions to the survivor, more decisiveness, less evidence of thinking about thinking, and more use of the various meanings of the word 'love'" (ibid., p.151).

Shneidman is not clear about this, but what he seems to mean by "dichotomous logic" is contradictory and/or illogical statements. A hypothetical example (mine, not Shneidman's) might be a single note expressing both "I love you" and "I hate you" to the same person. It might also be worth mentioning that attaching blame to others, rather than oneself, as Etkind (1997) reported, is obviously different from the self-blame referred to by Shneidman.

In the years since Shneidman wrote the above, studies seem to have been increasingly concerned with statistical models, with the exception of Leenaars' innovative protocol sentences, and they appear to have moved on to looking at suicide notes within various increasingly specific contexts, such as particular age-groups of people who attempted suicide but did not succeed. (Access to information such as notes from attempted suicides implies that such studies are invariably conducted by psychiatrists and psychologists (see below)). One of the more recent studies of suicide notes comes from Giles (2007) who, in her Ph.D. thesis, studied narratives within the notes. However, being a psychologist, Giles was couching her work in terms of what she called "Ultra-Social Phenomena", and she examined her notes together with detailed case histories of their authors (ibid., p.124). Giles had 173 suicide notes from 95 note-writers (ibid., p.32), and used a "content dictionary" whose entries could be tested as either matching, or not matching, with content in a text (ibid., p.124 and p.341-342) – an approach similar to that developed by Leenaars (1988) (see Section 3.1.6).

The aim of most of the above studies was to distinguish between real and simulated suicide notes. The cumulative main finding is that this can be achieved to some extent - but not totally. From a more pessimistic view, therefore, they fail to achieve their aim. Before I consider some of the possible reasons for this, I shall very briefly reiterate what

I see as being the main accomplishments of the above studies.

Shneidman's major contribution to the enterprise was his two sets of texts. The major contributions of the other studies mentioned above are in my view the following: Osgood and Walker's (1959) was to note the greater number of endearment and mother terms, and the lack of disorganisation in genuine notes, and also their greater amount of concrete references and fewer cognitive ones; Gottschalk and Gleser's (1960) was having more than one dimension to their sets of variables, and finding more references to people, things and spatial relations, and fewer to processes, in the genuine notes; Ogilvie, Stone, and Shneidman's (1966) seems to have been their use of a dictionary and computerised semantic tagging; Edelman and Renshaw's (1982) was that they combined computer analysis with a multi-dimensional set of variables, and found negative polarity in some phrases in genuine notes; Leenaars' (1988) is his different approach to the problem; Gregory's (1999) is perhaps his attempt to assess diverse sets of variables; Etkind's (1997) lies in his succinct summarising of many previous findings, and in making some results from academia accessible to the general public.

The summaries of these studies, in the above sub-sections, are broadly compatible with each other, but there is no clear, totally consistent, pattern of evolution. This must be partly because they used different sets of categories and were looking at different features. The exact, full and explicit details of the categories are not always clear. Also, what was found to be significant (statistically or otherwise) depended on what tests were used. This is why it is difficult to know for certain whether or not the findings do actually confirm each other.

The limitations of the above studies could be said to lie partly in their reliance on Shneidman's notes. These in turn have the limitations already outlined in Section 3.1, and mentioned in Section 3.2 below, as well as being collectively a small number of words by today's standards of corpus size (see Section 4.8). There is also a corpus-related problem to be aware of: the corpus versus the single text. Saying that two corpora are different is not the same as saying that a single text can be allocated to a single group. There will almost certainly be an overlap of 'distinguishing' features in

the middle of the spectrum.

The literature reviewed shows a mixture of methodological approaches. Some authors used computational methods, others used non-computational ones. Some authors concentrated on words, others on messages, and yet others used both. They all had a corpus, whether electronic or on paper, and (with the exception of Etkind) were tagging and counting tags, and trying to find distinguishing features in one, or another, set of texts. The tags/categories were of quite different kinds, and the authors were, from my own non-psychological standpoint, extraordinarily eclectic in what they chose to count. (This is merely an observation, not a value judgement.) Some of the authors were using traditional categories, such as nouns, verbs and adjectives; others were using semantic categories, but they seem to have been quite selective. In contrast, modern semantic tagging software allocates all content words, at least, to a semantic class (Wilson and Thomas, 1997).

All the methods seem to have worked reasonably well; however, most of the work does not have a "linguistic" feel about it. This may well be because most, if not all, the authors were not linguists, but psychologists or psychiatrists (see above). This is another limitation of the above studies. Notwithstanding Chomsky's (1972, p.1 cited in Laurence, 2003, p.69) claim that linguistics is a branch of psychology, there are well-defined and long-established differences in methodology and ethos between the two disciplines. Although one cannot necessarily criticise psychologists for using psychological theories and methods, from a linguistics point of view, it is severely restrictive to be tied to having to have psychological explanations for choosing or creating a category, and for any results.

McClelland, Reicher and Booth (2000) also come from the discipline of psychology; nonetheless they want other psychologists studying suicide notes to embrace discourse analysis, do more grounded research, and move away from working with mixtures of miscellaneous traditional psychological categories. Of course, their justification is partly based on ideas that are themselves rooted in psychology, the details of which (discursive psychology) are not relevant here. What is relevant is that other researchers'

methods, they say, simply do not work, because they are missing potentially relevant material by ignoring context (such as the note writers' addressees), and by not treating the texts as functional and performative acts of communication which their authors use to justify past, present and future events. I share this conclusion.

It could be argued that the methods of counting described in this chapter are precursors to methods used by other researchers, more recently, for other purposes. For example, forensic linguist Johnson (1997) who was looking at three student essays that she suspected were plagiarised, used the method(s) of counting words that occurred only once in a text, words that were unique to a text, and words that were common in a set of texts. She found "the high use of common terms, the low use of unique lexical vocabulary... and the degree of overlap in the once-only vocabulary..." (ibid., p.223) were convincing evidence of plagiarism. The three texts in question had 26 once-only words in common with each other, whereas three control texts (randomly chosen from the same set of essays) had only one. The text she suspected of being the source for the other two texts shared two-thirds of its once-only words with them. It is the use of quantitative methods to study a forensic linguistic problem that is relevant to this thesis.

Where the existing literature does report the use of grammatical analysis, it tends to use a conventional model of grammar that is unsophisticated and devoid of a sound accompanying theory: other models could be considered. It is also possible that the various classifications of words might be improved upon. Gregory (1999) seems to think that computer analysis can only work on structure (as opposed to content) variables, and is of little use because of this (see Section 3.1.7). But this may be a tagging problem, albeit a non-trivial one. Different grammars, and tagging different contents and contexts, or finding computer-friendly alternatives to tagging, may be profitable.

It does seem encouraging that many studies seem to claim a high success rate at distinguishing real notes from fake ones, but it must be remembered that not only is this less than 100%, but most of these studies are based on the same notes (those of Schneidman)!

3.2. Difficulties in Collecting a Corpus

In order to be able to compare real and fake notes, the obvious solution would seem to be to collect a corpus of each. However, what is clear is that there is considerable difficulty in collecting a corpus of real and fake suicide notes. As Chapter 2 has shown, and as Chapter 4 will also demonstrate, creating a corpus of real suicide notes is not a straight-forward matter, and creating a corpus of fake suicide notes presents even more difficulties. Creating a corpus of paired real and fake notes, as Shneidman did, is clearly far more problematic, and is a task that seems very rarely to have been repeated by anyone. (One exception to this is Black (1993, p.699) who believed that Shneidman's simulated notes might actually have been written by psychologically stable people, rather than non-suicidal people, and that these two issues were confounded. He says that the note-writers' histories of depression and stability, if any, were ignored.) Despite these criticisms, and Gottschalk & Gleser's view (1960, p.197) that having notes in pairs is not really very useful, people (such as Pestian et al. (2008) for example) are still referring to Shneidman's data some 50 years after he created his sets of texts. There are very good reasons for this.

Aside from the practical difficulty of getting hold of any genuine notes (see Chapter 4), acquiring simulated notes would involve: either collecting notes that have been fabricated or forged, or have been written under duress - of which there are very few available; or persuading people to write them. There are practical and ethical reasons why asking people to write suicide notes cannot easily be done. These include the requirements of university Ethics Committees and the various problems that would accompany any attempt at psychological vetting of potential note writers to ensure that they were not in fact suicidal.

Another important point is that Shneidman's simulated notes were written in very specific conditions, whereas 'real' fake notes, as Gregory (1999) points out, are written under very different circumstances. So creating notes for the purposes of research into detecting actual faked notes might not be too helpful. Of course, this raises the question of what genuine suicide notes should be compared with - which I will try to answer

below (see Chapter 4).

3.3. Durkheim and Sacks

Although this chapter is primarily about literature on suicide notes, I want to mention two more authors whose works are frequently cited studies of suicide although not of suicide notes. These are Durkheim (2002) and Sacks (Jefferson, 1989), both of whom were sociologists.

Durkheim (whose work was first published in France in 1897) found four major types of suicide (Durkheim 2002, p.257 and p.239). These were Egoistic (characterised by depression (ibid. p.246), melancholy and apathy) (ibid. p.257), Altruistic (which involved concern for others) (ibid. p.257), Anomic (which was characterised by anger) (ibid. p.247), and Fatalistic (a type which rarely occurred, but which was a response to an unalterable oppressive situation) (ibid. p.239). According to Durkheim, most suicides are a "mixture" of the first three types (ibid. p.250).

Sacks' work was the foundation of Conversation Analysis (Jefferson 1989, p.7-8); and a large amount of his data consisted of telephone calls to Shneidman's "Suicide Prevention Center" (ibid. p.6). His lectures, based on this data, were published after his death (in 1975) from transcriptions of audiotape recordings (ibid. p.1-2). In one such lecture Sacks examined the spoken text of a woman who said she wanted to kill herself to see if anybody cared (ibid. p.79-88). Sacks showed that people may have very few methods of ascertaining whether anybody cares about them, the exceptions being arranging weddings, parties and funerals and seeing who attends and what is said (assuming they could oversee the latter occasion) (ibid. p.83). Sacks also introduced the concept of 'Private Calendars' (ibid. p.84) and showed that a person deprived of someone they had used as a point of reference to dates and events (for example, their mother's birthday) would feel deprived of any way forward.

3.4. Conclusion

In view of the linguistic suicide literature, my general approach to the analysis of the data for this thesis was that it should be based in the 'grounded research' concept of

seeing what the data itself 'suggested'. Within this concept I wanted to apply some corpus analysis techniques: an electronic version of the data would facilitate the generation and analysis of such things as frequencies, concordances and collocations. I also wanted to try some Content Analysis, or semantic tagging, of my own (see Section 3.1.4). Thus, my first goal was to build a corpus of suicide notes. There does not appear to be any database of suicide notes: none which is readily accessible to researchers, lawyers or judges. Perhaps this work might instigate the creation of one. In one appeal court case involving contested suicide notes (see Chapter 1 and Section 2.3.1.6 above) the judges, in declining to accept an expert's 'psychological autopsy' report, remarked,

> "His reports identify no criteria by reference to which the court could test the quality of his opinions: there is no data base comparing real and questionable suicides..." (R v Gilfoyle, 2000, para.25).

However, (as outlined in Chapter 1 above) in view of what has been said above in Section 3.2, I had to modify my original goal. One argument is that too much attention has been paid to the need to try to discriminate between real notes and fake notes - especially a particular set of fake notes - and as a result, not enough attention has been paid to the characteristics of real notes. Because of the difficulties in obtaining 'fake' notes, the study of real notes has suffered. It might, at least temporarily, be more profitable to move the focus away from distinguishing between real and 'fake' notes, and on to a larger, and wider, range of real notes. In other words, perhaps a better way of proceeding would be to consider how similar to each other the notes are within a sub-corpus, rather than how different they are between sub-corpora.

4. DATA COLLECTION & DESCRIPTION, & SOFTWARE

This chapter describes the data collection and transcription process and the inherent, and subsequent, problems encountered. The first half of the chapter revisits the problems discussed in Chapter 2, above, regarding what a suicide is, and what a suicide note is, but here these queries are made in the light of the data with which I was confronted. The chapter is mainly concerned with what I refer to as the "Birmingham notes" (see Chapter 1), but also mentions the corpora I use to for comparison. Following this there is a summary of the data I found, including some of its characteristics, such as the lengths of the texts. The chapter ends with a description of the software applications I used.

It was my intention from the outset that suicide notes were to be collected and transcribed onto a computer and the software package "WordSmith Tools" version 3 (Scott, 1998a), (referred to sometimes as just Wordsmith), was to be used in the latter stages of the data preparation process, for example, to calculate token counts. Thus, the data were prepared with specific software in mind, which will be described after the data description Sections. Another aspect that was already in mind was that I would be counting subjects (note-writers) in addition to counting suicide notes.

4.1. Data Collection

The suicide note data was collected from the Birmingham Coroner's Office. Each case has an inquest number assigned to it by the Coroner's Office. For every case there is a physical file stored in the basement, containing a collection of documents relating to that case, including originals and/or copies or any suicide notes.

In addition to the physical files, the Birmingham Coroner keeps two electronic databases. One is called "The Visitors' Database" and is open to visitors to examine; the other is completely private and is called "The Coroner's Database". I assume that the latter includes either on-going or incomplete cases, or, perhaps, cases which for some reason may not be made available to the public. Unless I specifically state otherwise, when I refer to the database I used I am referring to "The Visitors'

Database", even if for simplicity I call it "the Coroner's database"; and "The Coroner" refers to the Birmingham Coroner. The visitors' database holds information on cases from the previous five years. At the time of data collection this meant cases from 1995 onwards.

There is an 'informal' text entry for each case on the database, titled "Inquest Story". Each case record actually had two parts to its Inquest Story, which appeared to have been compiled, in turn, as the case first arose and as it neared completion. These 'stories' were a useful guide giving background information on the cases, and indicating those cases where there was no suicide note to retrieve from the physical files.

Over several months, between 1999 and 2000, 18 visits were made to the Coroner's Office to collect suicide notes. This involved inspection of the computerised Coroner's Database, searching for suicides within it and noting their reference codes (inquest years and numbers), manual retrieval of case folders (which were organised by year and inquest number), and the photocopying of 307 whole and 2 partial suicide notes, and 14 possible suicide notes (see "open verdicts" below), that were in the case folders.

The database was organised into groups according to the Coroner's verdicts (see Section 2.2.1). The verdicts I focused on were "Killed Himself"; "Killed Herself"; those "Other" verdicts which were sub-titled as suicides, for example "Killed himself whilst the balance of his mind was disturbed"; and "Open". Open verdicts are the only type which can be re-opened, (Coroner's Office, personal communication). Some of the cases in which a verdict of "Other" was returned did involve potential suicide notes: unused 'suicide' letters which were found in drawers of people who died of unknown causes. These are, apparently, fairly common (Coroner's Office, personal communication). I included such notes in my initial collection but eventually decided not to use them in my corpus.

4.2. What Counts as a Suicide?
Although it has been argued by some (see Section 2.2.1) that actual suicides might well include cases given verdicts other than suicide, such as "Self Neglect", I limited my

counts of suicides to those cases which the Coroner had deemed "Killed Himself", "Killed Herself" and "Other... suicide". It is nonetheless worth reiterating that there are possibly far more suicides than have been classified as such under the coroner system.

It was not clear what the difference was, if any, between some of the Coroner's verdicts such as "Other: suicide" and "Killed Himself", but when introduced to the "Others", I was told that the Coroner can devise his own labels for verdict categories (Coroner's Office, personal communication). The Prison Service also devise their own categories, and their label of "suicide" can overlap several coroners' verdicts. Towl (1999, p.19) notes that

> "We have seen that official Prison Service figures of recorded suicide are significantly higher than would be apparent if we were to rely solely on Coroner's court verdicts of suicide."

That is, the Prison Service definition of suicide includes coroners' verdicts of "suicide", "accidental death", "misadventure", "open", "killed self", "suicide aggravated by lack of care", "misadventure aggravated by lack of care", and "lack of care", (Towl, 1999, p.18).

According to Blatchley (2006) of the Defence Analytical Services Agency (DASA), although he is not totally clear or explicit about it, it seems that a similar state of affairs to that in the Prison Service used to happen in the armed forces, and that recently this changed due to the incidents at Deepcut Army barracks where there were a series of disputed suicides during 1995-2002 (Anon., 2006a, p.1). The changes are that violent and accidental deaths have to be reported to coroners (or the equivalent in Scotland and Northern Ireland) (Blatchley, 2006, p.8). Indeed, writing this thesis on the shifting sands of changing official policies, it is probably safe to assume that the Prison Service must now also refer all deaths in custody to the coroner.

All this shows that during the years in which my suicide notes were written coroners could 'inherit' certain cases from other services who had different sets of views and categories from those of coroners, and these surely confound the situation regarding

what may count as suicide - unless, of course, one just accepts the coroner's decisions.

According to Blatchley (2006) the office for National Statistics uses coroner's court verdicts, and includes some "Open" verdict cases in its classification of "suicide verdicts". This practice now seems to be the norm for some researchers too. According to Linsley, Schapira and Kelly (2001) it is common practice to include coroners' "Open" verdicts in studies of suicide, only excluding from this sub-set the obvious non-suicides such as cases involving babies. However, because it conflicts with the concept of an Open verdict (see Section 2.2.1) I have decided against combining both types of verdict.

4.3. Problems and Limitations

There were various problems with the Coroner's database such as cases being classified under the wrong verdicts (see Table 4.1 below), and there were a few other issues which are worthy of mention for the sake of completeness, although they are not really pertinent to the data collection itself. The categories used in the database showed evidence of bias on the basis of gender[7] and culture[8].

I was unable to collect data on subjects' race or religion - as Shneidman (Shneidman and Farberow, 1957e) had done - because it was not present in the files. It was also impossible in practice to determine the socio-economic status of the subjects. However, I did collect data on the following attributes: Inquest number, name, sex, age, date of birth, date of death, mode of death, reason for death, marital status and occupation. The latter two, on reflection, were deemed irrelevant because they did not reveal information

---

[7] One section, although it seemed to be used regardless of the person's sex, had data entry fields labelled "if deceased is female... (only)... Husband/Widows [sic] Name... occupation...".

[8] The word "surname" was used where "family name" might have ensured more accurate information. In some cultures the family name is not a surname, i.e. not a last name, but is the first name mentioned in a person's set of names; in some the family name is in between the person's other names; and in some it is the last in a set of names. However, some cultures may not use "family" names at all, but variously names referring to ancestry, generations, parents, sex, place of birth, etc., and they could be offended by the term "family name". Despite all this, recording names is a bureaucratic necessity, and it would be difficult to find a system to suit everyone.

which could be useful.[9] For example, knowing that a person was single, married, divorced, separated or widowed, does not tell us whether that person had a good relationship with a steady partner or was living alone and was lonely; and a couple of words (which was typical) for "occupation" could not adequately describe a person's working environment, financial status, or the responsibilities their job had entailed. Indeed, examination of some files revealed that the "occupation" field on the database did not always accurately indicate whether or not a subject was even still employed.

Because of time limitations, with very few exceptions, I only examined the physical (non-electronic) case files of individuals where the computer database mentioned that there was, or might be, a note. The exceptions to this were when I happened to be within arm's reach of a file whose inquest code on its label caught my attention because the case was on the database as a suicide, and I looked at it. These exceptions yielded only a couple of notes. However, in my counting of suicide notes I was very aware that there could have been notes in files I did not examine.

Some notes may have already been destroyed. According to Etkind (1997, p.62),

> "It is common for suicides to destroy their notes, since many who feel they are unworthy to live also feel their final thoughts aren't worth sharing".

This statement seems a little dubious, and Etkind does not give any figures for just how common the destruction of notes is. And, of course, social, religious and cultural taboos surrounding suicide, and fear of not benefiting from a life insurance policy, can lead relatives to destroy notes prior to their coming to the attention of anyone else. Another point noticed in the database was that at least one of the suicides who did not leave a note could not write. It is open to speculation as to how many non-note-writers might have written notes had they been able to do so.

Where possible, I made copies of the original notes. However, some were severely damaged, or soiled, and the coroner's photocopies had to be photocopied. In some

---

[9] In fact, it transpired that most of these collected attributes were not utilised in this thesis.

cases the note was missing from the file (either lost, or reclaimed by relatives or friends) and only a photocopy was available. In other cases, it seemed that there had been a note, but both it and any photocopy were missing from the file. Had these notes existed, and been found, they might have incremented my corpora by 7 suicide notes and 5 suicide-note-leaving subjects, plus 1 note and subject in the "open verdicts". I decided to exclude them from my figures (see below) because it was not always certain that these notes had existed.

There were a couple of cases where it seemed clear that part of the note was missing. Both of these notes "ended" with "P.T.O." at the bottom of a page, yet had no further pages in the coroner's files. There was one foreign language note, and there were three audio transcripts. All these would need special consideration. I decided to accept the audio transcripts as verbatim despite the fact that there were also a few transcripts of handwritten notes, presumably made by someone in the Coroner's Office, where the original existed and was also in the file, and several transcription errors were evident, indicating a lack of attention to detail. Additionally, there were 5 male subjects who had, presumably, typed all or some of their notes. These notes were treated in the same way as the handwritten ones, as was one note from another male subject which was made from printed words, possibly from a newspaper or magazine, which had been cut out and pasted on a sheet of paper.

Table 4.1 shows the total number of suicides in Birmingham for the years 1995 to 1999. (It excludes any cases which may have been opened during these years but were either not closed or not updated in the "Visitors' Database" by the end of 1999.) It also shows that it is risky to rely on other people's data: included under the "killed himself" verdict were six "killed herself"s and one case of "natural causes"; while under "killed herself" were one "killed himself" and a "manslaughter" case. In the remainder of this thesis, unless specifically stated otherwise, all references and figures pertaining to the "killed himself" and "killed herself" verdicts have been amended to account for these mis-classifications. (The Coroner's Office, incidentally, was not concerned by females and males being mis-classified when I brought this to its attention, but it was concerned about the other cases.) Those verdicts of "other" which were sub-titled by the Coroner

as some type of suicide have been split to show males and females.

```
    Table 4.1. Total Suicides in Birmingham 1995 to 1999
             as represented on the Coroner's Database

        Killed Himself
Total under this verdict    236
includes 6 killed herselfs   -6
includes 1 natural causes    -1
Adjustment from herself      +1
                            ---
                            230
        Killed Herself
Total under this verdict     73
includes 1 killed himself    -1
includes 1 manslaughter      -1
Adjustment from himself      +6
                            ---
                             77
        Other
Total under this verdict    n/a*
Other Suicides, Male         22
Other Suicides, Female        7
                            ---
                             29

Grand total suicides  =     336

  *Not applicable
```

Some statistics relating to notes are presented in Tables 4.2-4.4. The number of suicides leaving notes in Table 4.2 are all those whose notes are accounted for in Table 4.4, including the audio messages, foreign note and partial notes, plus those whose notes were "illegible". The years given in Table 4.2 are the years in which the Coroner's inquest was opened. Some people who died (or were found) late in December of one calendar year had their files dated as the next year. The "Other"s are explained in Section 4.1 above. Unless I state otherwise, when I use the term "Others" I am referring to suicides found within that set of verdicts.

From Table 4.2 it can be seen that the percentage of Birmingham suicides who left notes is considerably higher than one might expect from some of the previous research: Shneidman and Farberow's data (see Section 3.1.1) gave them a figure of "about 15 per

cent" (Shneidman and Farberow, 1957c, p.8). However, when Salib, Cawley and Healy (2002, p.189) reviewed several other studies they found that figures ranging from 15% to 42% were cited. A more recent paper by Toshiki et al. (2005), apparently found that 23.4% to 36.2% of suicides in Kobe, Japan, between 1981 and 2001, left notes.

Table 4.2. Suicides leaving Notes: Birmingham 1995-99
as represented on the Coroner's Database

	1995	1996	1997	1998	1999	Total
Killed Himself	17	20	19	20	19*	95/230 = 41.30%
Killed Herself	7	6	5*	4	2	24/77 = 31.17%
Other, Male	1	-	1	-	1	3/22 = 13.64%
Other, Female	1	1	-	-	1	3/7 = 42.86%
Grand total	26	27	25	24	23	125/336 = 37.20%

(To 2 dec.pl.)
* includes 1 in each case who died late in previous calendar year.

Table 4.3. Open Verdicts in Birmingham 1995 to 1999
as represented on the Coroner's Database

	Open
Total Open	232
Note Leavers (all male)	8
Notes Found - legible	8
Notes Found - partly illegible	1
Notes Found - wholly illegible	5

Table 4.4. Suicide Notes found in the Coroner's Office, for Birmingham, 1995 to 1999

Coroner's Verdict	Written, Whole, Readable, English	Taped	Partial	Foreign	Totals Legible	--Illegible-- Partly	Wholly
Killed Himself	212	3	1	-	216	-	5
Killed Herself	74	-	1	-	75	1	-
Other, Male	8	-	-	-	8	-	-
Other, Female	3	-	-	1	4	-	-
	---	---	---	---	---	---	---
Grand Totals:	297	3	2	1	303	1	5

Of the "illegible" notes (see Table 4.4 above), the "killed himself"s were subjects for whom I already had other notes, but the "killed herself" subject had left only the one (illegible) note. Of the "partial" notes, I already had other notes from the "killed himself" subject, but I had no other notes from the "killed herself" subject. Of the three "Taped", that is, audio transcripts, in Table 4.4 above, one was a dialogue between the person about to kill himself and the emergency services telephone operator. All three "taped" notes were from different subjects for whom I had no other notes. As Table 4.6 shows (see Section 4.9 below), separating the audio transcripts from the rest of the "killed himself"s, and excluding the illegible, foreign and partial notes from all the sets of notes, left me with 92 "killed himself", 22 "killed herself", 3 "other, male" and 2 "other, female" subjects.

The differences between the figures for usable notes and subjects is due to some, in fact nearly half, of the subjects each having written more than one note. The number of multiple-note writers in the "killed herself"s is 9 and in the "killed himself"s is 42. These subjects account for 61/74 and 162/212 notes, respectively. The subjects for whom I have only one note each number 13 and 50 in the "killed herself"s and "killed himself"s, respectively.

4.4. What Counts as a Suicide Note?

What counts as a suicide note may seem obvious, but it is far from clear. As a starting

point, as mentioned in Chapter 2, I have taken as 'suicide notes' notes which the Coroner's Office have labelled 'suicide notes'. Thus I have to start from the assumption that that is what they are. But generally in this thesis as a whole I have widened this 'definition' in an attempt to be more inclusive, and not to prejudge what is in a note, and therefore my own study of them.

As has also already been discussed (see Section 2.2.2), the police and coroner's officers look for a note left on or near the body which states the suicidal intent. The further the note is from the body, the further back in time - relative to the death - when the note was written, and the less it says about the intent, the less likely it is to be called a 'suicide note'. What people, including the Coroner's Office, call a 'suicide note' also seems to be influenced by the number of other notes left (by the same person), their size in words, and the sub-genre.

Example sub-genres in my data were notes, letters, notebooks, diaries, poems, greetings cards, and transcribed audio tapes. If a "note" existed in a diary it was initially counted as another note. As Leenaars (1988, p.37) says "A suicide diary is a suicide note with a history." However, diaries can cover many days, months or years. The Coroner's Office seemed to have compromised by acquiring only the last dozen, or so, pages of diaries. My intention was to use the last day's entry of a diary as a suicide note. In practice, however, due to human error, I only followed this procedure with one of the two diaries I found in my "killed him/herself" data collection. The count, after this oversight, was that each of the two subjects (note-writers) concerned was deemed to have left two suicide notes.

If anything other than a name and/or address or telephone number appeared on an envelope, I counted it as a separate note. Letters and poems were likewise counted as notes. Single sheets of paper containing, say, two separate addressees and two separate signatures were counted as two notes. One exception to this occurred through human error: one of the "killed himself" notes should have been counted as three. On the other hand, two sides of a sheet of paper, or an envelope, unless distinct in some way, such as their addressees, were counted as being part of the same note. In two cases, one of

"Post It" notes being stuck on a wall, the other of writing actually being on several walls, I made the perhaps more intuitive decision to count the former as individual notes, and the latter as pages of one note.

There were three transcripts of audio events. These were all included in my concept of suicide note, but because one was a "999 call", a dialogue, it could not be treated in the same way as the other two. Further, to avoid compromising the speech versus writing situation, it had to be borne in mind that the two monologue transcripts were indeed speech when, and if, any results or analyses that included them were considered.

Here are some extracts from the Coroner's database which illustrate the "official" notions of what is, and is not, a suicide note. These were found in the 'informal' text entry part of the database, titled "Inquest Story" (see Section 4.1). In fact, only one note was actually found from the following cases but there are numerous references to items which could have been considered as notes but were discounted by the Coroner's Office. (The references are from my log of inquest codes.)

"No suicide note has been left" but 2 notepads were found, (M99/272).

It was not made clear why the notepads (above) were discounted. They may, or may not, have contained text.

"No suicide note found at this stage.. but quantity of letters found", (M99/139).

It was in the above case that I found a note: it consisted of one word, namely "SORRY ?". The "quantity of letters" were not in the file. It was not clear why none of them was considered to be a suicide notes/letter. As with the notepads in the first of these examples, it could have been their location, their genre, their dates, or a judgement on their contents.

"There were no notes at the scene but she had a Filofax and also kept a written journal...", (M99/295).

Again, in the above example, the Coroner's Office gave no reason for not counting

these items as suicide notes, but it seems reasonable to assume they were influenced by location and genre.

> "No suicide note was found..." but he left "a number of labelled items... outlining what was to be done with them", (Other, M95/433).

The labels referred to were not in the file. In this case they each might have been counted as a separate suicide note, but because they were not there they could not appear in my corpus.

> "Property taken includes... 'love letter' [sic] written within last day by deceased. No suicide note found", (97/136).

> Letters were sent to a medical insurance company, but were "not dated", (96/351).

In the above two examples genre again seems to have played a part in the decision to call none of the letters 'suicide notes'; however the potential importance of dates seems to be clear and thus recorded.

> "A number of notes were illegible although from the general text there is a possibility this may be part of a suicide note", (Open, M99/367).

The informal nature of the data entry might account for the singular "note" in this last example above. It would have been interesting if the Coroner's Office had described the texts they found because this example entry clearly shows an influence by content.

Although I looked for suicide notes in only suicide (and 'other… suicide' and open) verdict case files, it was not always clear what was a suicide note and what was a handwriting sample in the files. Each file included a document summary, but this was not always complete, and I wanted to be sure I included every suicide note I could find. However, this gave me more problems such as whether to count copies of notes made by their authors. One subject had been an ardent believer in euthanasia. Her file included various descriptions and affirmations of this belief, a living will, and power of attorney documents. She had written many of these long before her suicide, and had

repeatedly signed the affirmations over many years. She had also deliberately made many copies of her documents. These were in the file, presumably, both to act as handwriting samples and for the Coroner to prove intent to commit suicide. Where the copies seemed sufficiently different graphologically and lexically, and were not dated too far from the date of death, I counted them as separate texts, but this was something of a judgment call; there were several handwritten 'copies' of some texts which were very similar but not identical, and had I counted all of these as separate suicide notes it might well have skewed any analysis of my corpus, so I opted not to count the copies.

I also had to consider whether a will, or a euthanasia declaration, could count as a suicide note. If an informal will was the only document found in a suicide case, and it was written on the same day as the suicide, then arguably it indicates an intention to take one's life. Conversely, a suicide note can also be "counted" as a will. This was the judgment in Ryan v Kazacos (2001). Interestingly, in this Australian legal case the judge remarked, "The fact that he displayed it [the note] on his desk shows that it was to be operative" (Ryan v Kazacos, 2001). In these sorts of cases I generally counted all of these items as suicide notes, but with the provisos outlined above.

Yet another issue was whether a book, found near the body, with a name written on the cover, should count as a suicide note. What if that book had a page with its corner turned over, or included a passage which had been marked by the subject? In my data there was only one subject (a "killed himself") who left a book. It had his name, address and telephone number in it, and I counted these items as one note. As I recall, there were no other markings on or in the book, and this relieves me of yet another potential dilemma, namely what else should be considered as part of a note. Ought one to count a book's title, or any, or all, of its contents? As mentioned in Chapter 2, much deliberation went into the situation and meaning of a book, in fact a bible, in R v Bamber (2002). The court judged that Bamber had killed his family, but Bamber maintained that his step-sister had killed them and then killed herself. A bible was found near her body. The court considered whether the bible might have been a kind of suicide note, or was indeed made to appear as one by the killer. That this book was a bible emphasises one of the problems with counting books as suicide notes. Even a

short marked passage might well exactly match an individual's feelings, but it is very unlikely to reflect their own language style. For the linguist, none of the text of a book itself, as opposed to any text written on a book by a suicide, should count as a suicide note.

Notes from recent, but previous, suicide attempts, such as one found in the "Open" verdicts that was written, apparently, 20 days prior to the victim's death, were counted as at least potential suicide notes. Of course, in the case of an Open verdict, it cannot be certain whether the death was actually a suicide.

Regarding length of notes, it is easy to see how some may discount a one word note when an author has also written many other long letters. But if the one word was all that there was it certainly ought to be counted as a suicide note, and it seems reasonable, therefore, to count such a note as a suicide note even when it is one of many much longer notes. I had one such note in my data. It consisted of the word "GARAGE" [sic](M99433b).

I have taken the view that in principle nearly all the above instances, with the provisos given, ought to be counted as suicide notes regardless of the sub-genre, materials used, medium, number of notes left by the individual, their size in words, or whether the note mentions suicide or not. However, a text that consists of only the one word, "Bye", found 100 miles away from the dead body, and written 12 months prior to the act, would have to be judged in a deeper context before it was construed as being a suicide note. Fortunately, there did not seem to be any notes in my collection presenting problems of this nature.

Having decided what I would accept, in principle, as a suicide note, and from the notes I actually had in my possession on what I would like to have in a corpus of suicide notes, I now had to decide what I reasonably should have in the corpus. The cells of Table 4.4 above (excluding those in the "Illegible" and "Foreign" columns) along with the legible notes from the open verdicts were all initially considered to be potential sub-corpora. However, the "other" (suicide) males and females, "open", and non-dialogue and

dialogue "tapes" had to be abandoned because their use as separate sub-corpora was not practical, largely due to the very small number of them. I have already stated why I did not want to absorb the "open" verdicts into the "killed himself" and "killed herself" sub-corpora (see Section 4.2, above), and similar reasons apply for my not wanting to absorb the "other"s and the audio tape transcriptions, namely, they are of different sub-genres. Thus, the main work was to be done on the "killed himself" and "killed herself" verdicts which were written (as opposed to spoken), complete - as far as one could tell, mostly readable, and in English. The "killed himself"s and "killed herself"s will usually be referred to, henceforth, as the "Him"s and the "Her"s, respectively.

Despite the decision not to use some of the texts in any corpus, they are an important part of the story of what I found in the Coroner's files, and so I show some of their characteristics in Table 4.6 (Section 4.9, below) along with those of the sets of texts I did use. For the purpose of Table 4.6 only here are some abbreviations. The "other" (suicide) males and females, "open", and non-dialogue and dialogue "tapes" are referred to as the "Otm"s and "Otf"s, "Opn"s, "Tap"s, and "Dia"s.

4.5. Data Preparation & Anonymisation

The subjects' non-linguistic attribute values (age, date of death, etc.) were tabulated and sorted. The suicide notes were transcribed onto a computer as text files, one note per file. This would accommodate most software packages and therefore whatever software would be used to interrogate the resulting corpus. The references which I will use for the notes are the codes described in Figure 1.1 (Chapter 1).

Note-writers' crossings-out were respected: I did not attempt to reproduce electronically within the corpus any portion of text which had been crossed out. (Although such text is worthy of research, my decision to exclude it meant I could concentrate on what I presumed the note-writers wanted to be read.) However, any pieces of text clearly marked for inclusion in some other places in a note, for example, by arrows, were put in the indicated places. Multiple pages in a text were signified by my adding a line of dashes ("-"s) between the pages. One female subject had included a line of Pitman's

shorthand in a note. The longhand equivalent was included in the transcription[10].

It was noticed that there was an inconsistency between the software used for editing the texts and the software used for counting the tokens that meant that the pound sterling signs in the British notes were not being counted properly. (The text editor uses an old IBM extended character set.) The Birmingham notes were appropriately corrected.

Finally, to be able to meet the requirement of including the texts with the thesis, full anonymisation was needed. I describe this process here before returning to the other transcription-related issues in the preparation of the data (see Section 4.6, below). At the expense of losing some original layout, I decided to anonymise by changing all the addresses to 6 characters, starting from the beginning of the alphabet. For example, "Cherry Lane, Edgbaston" becomes "Aaaaaa Lane, Bbbbbb". When "Llllll" was reached, combinations beginning with "Aabbaa", "Bbccbb", etc. were used. Unlike Shneidman (Shneidman and Farberow, 1957e) and Leenaars (1988) (see Chapter 3), I do not like using pseudonyms, and I would rather make it explicit that something has been anonymised. Thus, I also changed people's names (including surnames) to 4 characters, adding a 6-character variant if one person appeared to be being addressed in two ways, starting from the 'middle' of the alphabet ('N') (for example, 'Di' and 'Dianna' might be Fnnn and Fnnnnn respectively). Where one person has two forms of reference, the length of the anonymised form does not necessarily relate to that of the original form.

Except in the case of surnames or family names, in people's and animals' names the first letter indicates the individual's sex where this seems obvious. For example, the hypothetical "Algernon Bloggs" and "Mary Bloggs" became "Mnnn Pppp" and "Fnnn Pppp", but "Mr Philip" became "Mr Qqqq". This made the reasonable assumption that none of the authors were referring to people in the 'title forename' format. Names of organisations that are also geographical places were treated as names, rather than addresses, unless it seemed intuitive that the subject was actually referring to an address. Gender was not applied to street names, so "Mary Road" became, for example,

---

[10] This was translated by Rivelyn Shapero.

"Gggggg Road". This meant that ambiguity between names and addresses could be more easily avoided, and that the 4-letter name scheme could be extended when all appropriate alphabetical letters (from n to z) were used up: "Mary Freda Bloggs" could become "Fzzz Fbbb Cccc".

When the context seemed likely to create confusion, particularly between names and/or addresses, then another sub-scheme was used. So, for names: after the 4 letter words using the letters n-z (sub-scheme 1) were exhausted, 4 letter words using letters a-l were used (sub-scheme 2), then 4 letter words based on a 'scrolling window' through the alphabet, i.e. abcd, bcde, to vwyz (sub-scheme 3) were used. Finally, when all these combinations were exhausted, another sub-scheme using words beginning with "aa", "ma" or "fa", i.e. aabc aacd, or maab, or facd, was used.

Other attempts to avoid potential confusion included not using the letters M and F where they might be conflated with a the individual's sex, thus avoiding combinations such as Mdef and Fdef; avoiding Uuuu and Vvvv in the same subject's notes due to their visual similarity; not using the letter A for initials (potential confusion with indefinite article); and not using the letters I (confusion with the pronoun), O (confusion with zero) and X (confusion with symbol used for 'kiss'). Initials produced some loss of gender information. For example, if "Mary Bloggs" was "Fnnn Pppp", then "M. Bloggs" was "N. Pppp". When an initial's source (the name it represented) could not be determined, "B" or "Z" was usually used. The use of S was also avoided where it could be confused with a plural or possessive.

Attempts were made to preserve, or at least signal, capitalisation. For example, "maRy" would become, say, "fqQq". Inconsistent, or erroneous spellings were usually signalled by suffixing a word with an extra letter, often n, a or z. For example, "Edgbustin" and "Edgbaston" used by the same subject would be turned into, say, "Ccccccn" and "Cccccc". Attempts were also made to imitate oddities such as spurious hyphenation within names, names that were split over two lines, and names with prefixes (such as "McDonald", for example). To this end, forms such as "mcPppp" were used. Nicknames were usually anonymised. Names almost certainly referring to public

figures, such as Princess Diana, were left alone. For addresses, some locations, in some contexts were also left alone where they were unlikely to identify any of the individuals concerned. In particular, "Birmingham" was not anonymised. All phone numbers, and almost all address-related numbers, were replaced with "9"s. No dates or monetary amounts were altered.

The above anonymisation scheme was applied consistently within each subject's note, or notes, generally starting with the first name encountered being changed to "Mnnn", "Fnnn" or "Nnnn", but it was not applied at once over any entire sub-corpus. In other words, "Mary" might be "Fnnn" in one note, but "Fqqq" in another, although "Mary" would be represented consistently in all the notes from the same subject. For the sake of readability, examples, or extracts, used within the text of the thesis, however, may have their anonymised names changed to equally anonymous but more realistic names, such as Shneidman's ubiquitous Mary and Bill. Here is an example of a note with a) its names changed, b) the anonymisation effected, and c) names that are suitable for citing in the thesis. (That the first name in 'b' is Fsss rather than Fnnn is because this is the fourth of four notes in my database that were written by the same author.) (I have used a non-proportional font because it better represents the original layout of the notes.)

a) Note (with changed names)
```
"Dear Camilla & Charles,
              I can't get
over the lies that has been
told about me . It is real
malicious gossip . Friends or so
called friends are not even
speaking to me because of
the lies . Everything you need
is in a Bag with your
name on it . Please get
something nice for Philip and
Elizabeth for Christmas.
        Thanks for everything
           Gordon" (M99510d)
```

b) Note anonymised
```
"Dear Fsss & Mttt,
              I can't get
over the lies that has been
```

```
        told about me . It is real
        malicious gossip . Friends or so
        called friends are not even
        speaking to me because of
        the lies . Everything you need
        is in a Bag with your
        name on it . Please get
        something nice for Mwww and
        Fyyy for Christmas.
                Thanks for everything
                    Mnnn" (M99510d)
```

c) Note as might be cited in thesis

```
"Dear Mary & Tom,
                    I can't get
over the lies that has been
told about me . It is real
malicious gossip . Friends or so
called friends are not even
speaking to me because of
the lies . Everything you need
is in a Bag with your
name on it . Please get
something nice for Fred and
Betty for Christmas.
        Thanks for everything
            Bill" (M99510d)
```

Lost from the original texts, in addition to the entities already explicitly mentioned above, were all the words, particularly in surnames and addresses, that could contribute misleadingly to frequency counts of colours, modals, etc. Hypothetical examples are "Green Lane", "Joe Black" and "May Bloggs". In fact, there was only one instance of a name that could be miscounted as a modal.

The problem is that this makes a nonsense of the type counts and the Type Token Ratio (TTR) figures (see Sections 4.8 and 4.9). Suppose one had "Joe", "Jack" and "Rupert", that is three types. If one changed all of them to "Mmmm", that is one type. Therefore, it must be stated that the type counts are subject to anonymisation distortion. However, the analyses of other authors with which I might wish to compare my own are subject to similar distortions. After all, Shneidman's "Mary"s might originally have been "Jo"s or "Antoinette"s, so type counts of his sub-corpora have characteristics and flaws similar to mine.

93

Indeed, concerning the work done in this thesis, there are no analytical issues affected by the anonymisation process. For example, it is assumed that the details of the addresses themselves are probably not very important or crucial to any of the analyses in this thesis: subjects do not have quite the same choices of words when they use names and addresses as they do when they use other items.

4.6. Transcription Problems

More problems were encountered as the notes were transcribed. As with most data there are the problems of apparent gaps in it and incomplete parts of it. In my case this applies both to the existence of notes themselves and to their content. Regarding the latter, again some judgment calls and assumptions had to be made.

It was often difficult to ascertain the order of multi-paged notes: pages were seldom numbered, and the pages of notepads were often adrift from their pads. The police photographs, which were also in the Coroner's files, sometimes showed the order of a couple of pages, but some files were in great disarray giving no clues about page order. Thus, the page sequence of some multi-paged notes could only be guessed at. A similar problem illustrating the files' disarray was that the Coroner's files often had multiple-note writers' notes in the wrong envelopes. This was apparent when, for example, two notes addressed to, say, A and B were found in envelopes addressed to B and A, respectively.

Some texts, or parts of texts, were, to various degrees, illegible. Sometimes a few characters went off the page. Where it seemed reasonable to guess what the missing or illegible characters were, based on the context of the rest of the note, this was done. The same applied to punctuation marks. In the case of spelling and punctuation, authors were generally given 'the benefit of the doubt'.

There were many authors whose spellings were unconventional, and some of these authors also tended to be inconsistent in their spellings. It was obvious that this would add to the problems that could be encountered when using any software to automatically

search the texts. As an example, some subjects spelt 'know' as 'no'. I decided not to correct misspellings of this kind and simply to bear in mind that there would be a 'margin of error' when observing TTRs and the like.

Some authors used non-standard capitalisation. This usage seemed to favour the capitalisation of nouns. In many cases it seemed that this simply represented the writer's preferred, or indeed only known, way of writing a particular word or letter, rather than an intention to give emphasis to any word. In some texts, such authors appear to have used larger lower case characters to represent capitals. With the exception of the latter, I have tried to replicate these features in my transcriptions.[11] Some of the authors emboldened or underlined words for emphasis but I did not reproduce these features. The reason for this is that I intended to subject my files to automatic analysis by various software packages which would only accept plain (ASCII) text which does not permit different font types. The only way of indicating bold type or underlined characters in plain text format would be by manually adding mark-up tags; alternatively I could have created more than one version of each sub-corpus. Either of these would have been time-consuming and unnecessary for the kinds of analysis I would be conducting, and thus beyond the limits of this thesis. I did, however, make some exceptions and employed square brackets as a rudimentary form of tagging, as discussed below. There were also cases of words being split over lines without the use of hyphens, which I did replicate.

I have tried to preserve the original punctuation in the texts - in many cases there was little or none. Where it did exist, however, there was much evidence of punctuation preceded by spaces, such as in the following where there is a space before the full stop.

"a few words . Then more words".

Square brackets were chosen to demarcate unreadable parts of the texts. There was only one instance of a subject using square brackets, in the "Hims". Totally illegible words, and the illegible parts of otherwise legible words (assuming, that is, that all these entities

---

[11] I used only one font, and only one font-size for the transcriptions.

are indeed words) were represented by "[?]"s. Examples are "find [?] the...", "conta[?]" and "m[?]a". More easily guessed, but still illegible words and parts of words, were enclosed in square brackets. Examples are "[for]" and "ha[d]".

The problem with this is that Wordsmith (Scott, 1998a) (see Sections 4.8, 4.9 and 4.11 below) treats square brackets as white spaces and ignores punctuation marks, and this renders the count of tokens inaccurate. So, "conta[?]" and "[for]" are counted, acceptably, as one token each; but "find [?] the..." will be counted as two tokens when it is fairly clear that there are really three; while "ha[d]" and "m[?]a". will be treated by Wordsmith as two tokens each when it is reasonable to assume that in each case it is really one.

Consequently it appears that about 7 extra tokens were wrongly added to the "Him"s while some 38 were wrongly omitted. Similarly, approximately 2 extra tokens were wrongly added to the "Her"s while about 4 were wrongly omitted, while 1 token may have been wrongly subtracted from the "Open" scores. However, it proved difficult to assess this accurately, and I felt that the two phenomena were best treated as more or less cancelling each other out. All this shows that transcribing texts is not necessarily a simple process.

4.7. Other Suicide Notes & Comparison Corpora

Because of the ubiquity of Shneidman's 66 paired suicide notes (Shneidman and Farberow, 1957e) in the literature, and the fact that they were published and are available, I have added them to my data, although I have kept them separate (see Chapters 1 and 3). The "Simulated" and "Genuine" sub-corpora of notes are referred to by the abbreviations "Sim" and "Gen" respectively. Although they were printed in book form (ibid.) and thus fully legible, a little preparation was still required to convert them into text files. This mainly involved a mixture of electronic scanning and manual copying.

I wanted to be able to compare the suicide notes, not only with other suicide notes, but also with unrelated texts. The non-suicide related corpora used for comparison included

the British National Corpus (BNC) "written" sub-corpus obtained via Mike Scott's website (Scott, 1998c), and also available (albeit from another version of the BNC) for use with the WMatrix software (Rayson, 2007) (see Section 4.11), the Bank of English (BoE), and the Bank of English's "brbooks" (British books) and "usbooks" (United States books) sub-corpora.

Both Wordsmith Tools and Wmatrix (see Section 4.11) are able to compare a user's corpus with a reference corpus namely the BNC. To be precise, the version of the BNC used by Wordsmith is known as the "BNC World Edition" and was created in 2001, and the version of the BNC used by Wmatrix is known as the "BNC sampler" (which is a subset of BNC World). Therefore both applications are capable of comparing a text, or set of texts, against some norm (which is what a reference corpus is usually used for). Moreover, their creators both considered not only that such comparisons form a useful part of a user's methodology, but also that the BNC is an appropriate default, general purpose, reference corpus.

Wordsmith Tools version 3 (Scott, 1998a) can compare a single text with a word list of frequencies provided the latter is in its proprietary file format. Word lists in this format were not available for the Bank of English. For the BNC, on the other hand, the word lists (including the "written" one mentioned above) were available - although the texts from which they were made were not. Table 4.6, in Section 4.9 below, summarises all of the corpora and sub-corpora used (or acquired) in this thesis.

4.8. Comparison Problems

Every corpus is a compromise; no corpus is big enough or representative enough. The corpora used in this thesis are extreme examples of this. Ideally, I would have texts that are clearly written and of suitable, similar lengths. However, all corpora are a compromise between what one wants and what is available.

I now look at the issues affected by the compromise, namely the issues of type and size, and of normalisation. In the following chapters I compare some of my sub-corpora with each other, and sometimes with some other much larger sub-corpora, or corpora. My

sub-corpora are of various sizes, some being relatively small; they contain variously different numbers of texts (notes); and the texts themselves vary greatly in their sizes. According to Kilgarriff (1997, p.4),

> "It is not clear what, if anything, a measure of the similarity of a thousand-word corpus and a million-word corpus, or a one-text corpus and a thousand-text corpus, would mean.".

He goes on to discuss the same topic in relation to comparisons of corpora containing texts of unequal sizes, or containing complete versus sampled texts (ibid., p.4). Kilgarriff also asserts that "Corpus similarity can only be interpreted in the light of corpus homogeneity." (ibid., p.1).   He states, "It is not clear what a measure of similarity would mean if a homogeneous corpus was being compared with a heterogeneous one." (ibid., p.1).   However, this would be part of establishing the linguistic features which defined its homogeneity.

According to Sinclair (2001, p.2), "million-word corpora must be regarded as small by today's standards."   But without a million words, should one give up on the hope of having a corpus? There needs to be some way, however primitive, to deal with the sub-corpora so that comparisons may be permitted. The normalization I have applied is the division of each raw frequency result by the token count of the relevant sub-corpus. That there remains some discrepancy seems obvious; nevertheless, this was my approach. I have taken a similar approach regarding the numbers of notes and subjects contributing to any result, i.e. extracting separate figures for them and dividing by their totals (see Section 4.9). Many studies of corpora may acknowledge the bias given to their results by small numbers of constituent texts, but they seldom seem to isolate these phenomena.

One potential means of comparing corpora might be the TTR or type/token ratio. One common use of a TTR is to give the average frequency of use of a type, but Wordsmith can also calculate 'average type/token ratios' which may allow for some comparison of texts or corpora of different sizes (Scott, 1998b, p.92-93); although it is not clear how to interpret this information - particularly given the problems described in Sections 4.5 and

4.6. Although I did not utilise TTRs in this thesis, for the sake of completeness Table 4.6, in Section 4.9, presents the TTRs for all the sub-corpora.

Albeit perhaps out of context, and most probably not as he intended it to be used, Sinclair's comment offers a little hope.

> "Comparison uncovers differences almost regardless of size." (Sinclair, 2001, p.3.).

4.9. Corpora Summary

It should be noted that for the Birmingham sub-corpora, as stated in Section 4.4 above, only the "Hers" and "Hims" are used in the remaining chapters of this thesis. Table 4.5 below shows the abbreviations for the suicide note sub-corpora I actually analyse in this thesis (see Chapters 5, 6 and 7).

Table 4.5. Corpora Abbreviations

Abbreviation	Explanation
Gen	Shneidman's genuine suicide notes
Sim	Shneidman's simulated suicide notes
Her	Birmingham 'Killed Herself' verdict notes
Him	Birmingham 'Killed Himself' verdict notes

Some abbreviations introduced earlier are summarised in Table 4.6 which shows the word tokens per corpus, and the number of notes and subjects, where applicable, of most of the data I collected. The numbers of tokens have been used as divisors to create the percentage occurrence figures that are used in some of the result tables of this thesis (as mentioned in Section 4.8, above). The Bank of English gives its frequency values as averages per million words. Any percentages for the BoE shown in the remaining chapters in this thesis are their figures divided by 10,000.[12] In any columns labelled "Notes", a score of n means the word (or phrase) was found at least once in n notes. In any "Subjects" columns, a score of n means the word was mentioned at least once by n

---

[12] All percentages given in this thesis have been rounded to two decimal places.

subjects. The counts of tokens and types used for all my sub-corpora were made using Wordsmith Tools version 3 (Scott, 1998a) (see Section 4.11).

Table 4.6. Tokens, Types, Notes & Subjects per Corpus

Corpus	Tokens	Types	TTR	Notes	Subjects
gen	3,524	760	21.57	33	33
sim	2,152	546	25.37	33	33
her	7,372	1,455	19.74	74	22
him	30,046	3,338	11.11	212	92
otm	1,247	461	36.97	8	3
otf	95	61	64.21	3	2
opn	613	287	46.82	8	6
dia	141	75	53.19	1	1
tap	551	251	45.55	2	2
BoE	450,000,000	n/a	n/a	n/a	n/a
brbooks	43,333,620	n/a	n/a	n/a	n/a
usbooks	32,471,311	n/a	n/a	n/a	n/a
BNC written	90,748,880	377,384	0.42	n/a	n/a

4.10. Suicide Note Size

One common assumption among the general public is that suicide notes are invariably short. Because many of the results in the following chapters are based on word counts it seems appropriate to look at this now. As Figures 4.1 and 4.2 show for the Birmingham sub-corpora, it seems the assumption is generally true. However, the largest and smallest of these notes, in tokens, are 469 and 8 in the "Hers" and 1,838 and 1 in the "Hims", respectively. This suggests that, generally, there is a wide variety in the length of suicide notes, although most of those in my Birmingham corpus were under 100 words long.

Nearly half of the notes written by females (51%), and nearly half of the notes written by males (49%), are less than 75 words long. Indeed the numbers of notes that are under 25 words, 25-49 words and 50-74 words long are very similar for the "Hers" and the "Hims". Beyond the 74 word length, however, there are clear differences between the sub-corpora. Over three times as many 75-99 word notes were written by females than were written by males. Yet the notes that are over 99 words long were mainly written by males rather than females, the figures being 44% and 27% of the notes,

respectively. Further, the "Hims" that are 200 or more words long (23%) are more than double the number of "Hers" in that category (11%).
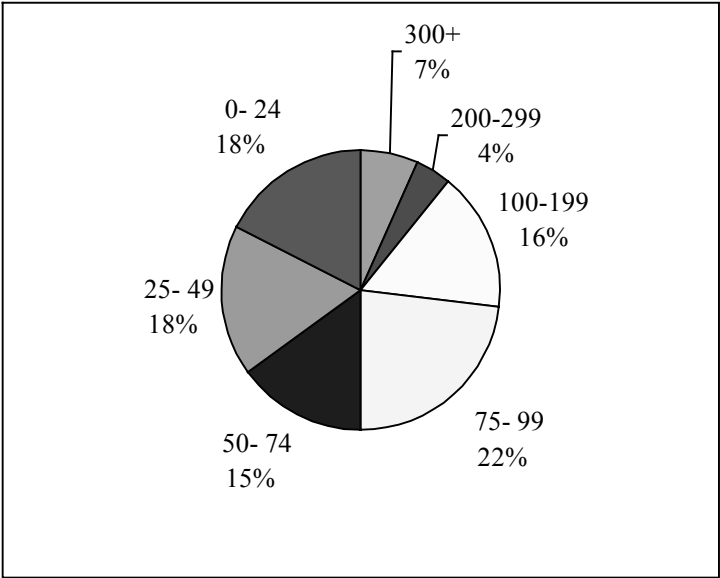
Figure 4.1. Note-Size in Words: Hers
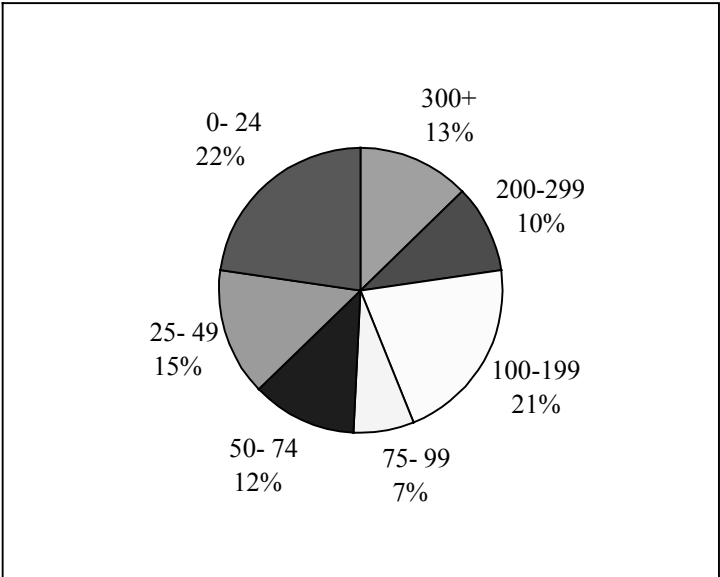
Figure 4.2. Note-Size in Words: Hims

The mean note-size in the "Hers" is 99.62, and in the "Hims" is 141.73. There are a

couple of outliers in the "Hims" with word counts of 1,060 and 1,838 (and the third highest word count is 726). The median note-size of the "Hers", 76.50, is very similar to that of the "Hims", 77.00. The mode for note-size is 83 in the "Hers" (due to four notes from three subjects), and 9 in the "Hims" (due to five notes from four subjects). The mode for the "Hims" is not representative of the whole sub-corpus. The numbers for the mean, mode and median are not enormously different from each other in the "Hers". However, for the "Hims" they are very different from each other.

4.11. Software

To complete this chapter on my methodology I now discuss the software applications used in the thesis. I used both Wordsmith Tools version 3 (Scott, 1998a) and Wordsmith Tools version 4 (Scott, 2007a). Wordsmith 3 was used mainly for token counts of the sub-corpora (see below). When Wordsmith 4 became available it was also used because, unlike version 3, it could produce word lists that gave the number of texts involved in a word's frequency. Wordsmith 4 was used for basic word and 'key word' lists (see Chapter 6), although Wordsmith 3 was used for the concordance lines (in Section 6.3) and to supplement Wordsmith 4 in the work on clusters (see Section 6.4). For semantic tagging (see Chapter 7) I used Wmatrix2 (Rayson, 2007) (see below).

I also used "megrep" (Microsoft Corporation, 1987)[13] which is derived from the Unix[14] standard string-handling command 'grep' (Bourne 1983), along with an application I wrote (called 'Readgrep') to count the lines it produced. Here follows a simplified description. 'Grep' is an acronym for 'global regular expression print'. It is a utility that allows one to search for regular expressions such as sequences of characters (which may be words or various patterns of letters such as 'abca') across several files at once and outputs the lines in which the sequences are found. It is essentially a concordance program minus any word alignment or sorting facility. The output lines are prefixed with the filenames that contain the strings. Since each of my suicide notes is an individual text file whose name includes the subject (author) code, this would prove

---

[13] This Microsoft Editor version of grep for the MSDOS operating system is a rather limited version of that which is included with versions of the Unix computer operating system.

[14] UNIX is a registered trademark of Bell Laboratories.

useful for the further investigation of any results from the other software that might be biased in favour of a few subjects or texts. The 'Readgrep' software utilises sub-corpus token counts acquired from Wordsmith 3, and any differences in the figures for overall occurrences (or the numbers of texts) between those produced by this software and those produced by Wordsmith 4 are due to the different token counts calculated by the different versions of Wordsmith. Figure 4.3 shows some example output from 'grep' (used with the command "megrep kill *.txt" in the "Hims" file directory).

```
M95214B.TXT 65 8 :and it kills you I'll come with you anyway
M95214I.TXT 15 5 : OF KILLING MYSELF FROM CUTTING WRISTS
M96242A.TXT 22 18 :       AND I WILL KILL MYSELF TUESDAY 4th1996
M96242C.TXT 14 43 :LIKE IT MUCH NOW.  DON'T DRINK & DRIVE IT KILLS
M96318C.TXT 62 16 :DISPAIR. MOMMY KILLED DADDY.
M96382A.TXT 6 19 :FNNN knew I would kill
M96480B.TXT 31 8 :either kill my neighbours or myself within a couple
M96480B.TXT 37 47 :and it failed there is nothing left to do but kill myself
M96488A.TXT 5 9 :her its killing me inside . I know I should just get
M96488B.TXT 2 13 :go. But its killing me inside wish you could of give
M96488B.TXT 12 45 :bad I wanted to help you through life youve killed
M98059.TXT 1 24 :THE PAIN IN MY BACK IS KILLING ME
M98059.TXT 3 18 :SIDE EFFECTS ARE KILLING ME
M98173B.TXT 11 1 :KILLS GeRms
M98362.TXT 19 11 :WOULD NOT KILL MY SELF , BUT BY NOW
```

Figure 4.3. Example Grep Output

The default Wordsmith Tools settings were used for the work in this thesis. The settings relate to what will be treated as a word and how the words will be counted, and are as follows. Words were processed in a case insensitive manner (i.e. 'A' was treated in the same way as 'a'). The minimum frequency of a word's occurrence for it to be included in a word-list was one. Hyphenated words were treated as separate words, so "note-writers", for example, counted as two words. Numbers that were written as digits (e.g. '5' as opposed to 'five') were not displayed in the output word list, but their individual instances were counted and summed and they appeared collectively as a hash 'token' (#), one hash per word list (see below). Digits were, however, included in the total of running words (see below). Another default setting was to ignore apostrophes at the ends of words. The following invented text is an example of how Wordsmith Tools 4 counts items under these settings.

Individual types: 1 2 3   4   2 5
'The 92 note-writers' 6 texts.'
Individual tokens: 1 2 3   4   5 6

The two sets of digits are counted as the same type: '92' is counted as one token, and '6' is counted as another. The word 'note-writers' is counted as two tokens, and 'The' and 'texts' are each counted as one token. This gives six tokens altogether. Table 4.7.a shows the relevant 'statistics' (as Wordsmith calls them) for the above. However, there is a problem with Wordsmith 4: it is not clear where the 8 'running words' come from. Removing the apostrophe, making the '92' a single digit number '9', and even removing the full stop, all produced the same results. So the basic token and type figures seem reliable, but the figures for 'running words' do not.[15]

Table 4.7.a. Wordsmith Tools 4 'Stats'

tokens (running words) in text	8
tokens used for word list	6
types (distinct words)	5

Table 4.7.b shows the 'frequency' view of the same data. In Table 4.7.b, in the column headings, 'N' is the line number of the result row, 'Freq' means raw frequency of occurrence, and 'Texts' refers to the number of texts in the sub-corpus. The '%' columns give the percentage versions of the raw figures to their immediate left. For example, in Table 4.7.b, the fourth row shows that 'the' occurred once; that it comprised 12.5% of all the words in the text; that the word occurred only in one text; and that that one text was 100% of the whole corpus.

---

[15] Wordsmith version 5.0 (Scott 2008) (not used in this thesis) has corrected this anomaly and gives 6 running words.

Table 4.7.b. Wordsmith Tools 4 Frequencies

N	Word	Freq.	%	Texts	%
1	#	2	25.00	1	100.00
2	NOTE	1	12.50	1	100.00
3	TEXTS	1	12.50	1	100.00
4	THE	1	12.50	1	100.00
5	WRITERS	1	12.50	1	100.00

As can be seen in the fourth column of Table 4.7.b, there is another problem with Wordsmith version 4. Here, all five lines together represent the entire corpus, but the percentages representing the frequencies of occurrence of all the items in the text do not add up to 100%. This is because the frequency percentages are based on the total number of running words which, as illustrated above, is not reliable. However, although the absolute percentages are not accurate, the relative percentages are, so the words can be compared to each other in terms of their comparative frequencies. So despite the problems with Wordsmith 4, it was felt that using it consistently would give worthwhile results that would be consistent across corpora.

Additional default Wordsmith Tools settings used for keyword lists (see Chapter 6) were as follows. The minimum frequency of occurrence for a word to be included in a keyword list was three. The statistic used to calculate 'keyness' was log likelihood, and the maximum p value (the significance) was 0.0000001.

Wmatrix2 (Rayson, 2007), the most recent version of Wmatrix, is a web-based semantic tagger.[16] It is described by its creator as a front-end to both a semantic tagger and a part-of-speech tagger. As mentioned above (Section 4.7), it provides access to both the BNC written and BNC spoken sub-corpora frequency lists which can be used as reference corpora for comparisons with user-supplied corpora. Since it relies heavily on the BNC, and is U.K. based, Wmatrix employs British English spellings. The tagger operates on multi-word units as well as single words, so it can deal with phrasal

---

[16] The work in this thesis was in fact done with Wmatrix2, but for convenience I refer to it as "Wmatrix".

concepts (such as 'more and more'). The results of comparisons are produced in descending order of Log Likelihood. The Wmatrix category system is described in Chapter 7.

Some extra data preparation was required before using Wmatrix. All texts within a corpus (or sub-corpus) had to be put into a single computer file; and any angle brackets had to be removed from the text. Six 'greater than' brackets ('>') were found in the "Hims", and these were replaced with spaces. It is worth mentioning that since Wmatrix is a web-based application, i.e. installed on a server computer (in Lancaster, U.K.), one has to upload one's corpora in order to use it. This has the potential for incorporating glitches into the data which may well be the main reason for any discrepancies between Wmatrix results and those from other software.

According to Rayson, Archer, Piao and McEnery (2004, p.7), the "UCREL semantic analysis system" used by Wmatrix is 91% accurate in its allocation of semantic tags.[17] The problems with Wmatrix are not in what it does, but in how it does it, principally in the user interface. Getting results which include the category names in the display is a far from intuitive process; and it would be a great deal easier to use if all its columns of word lists, concordances and semantic tags were in the same font size and in line with each other. It would also be extremely helpful for those organising their corpus as I have if Wmatrix could process a set of individual files and give results which referenced each of them. Finally, being internet-based, and subject to continuous improvements by its authors, the information displayed can sometimes change between uses.

4.12. Conclusion

In this chapter I have described the method used for collecting the data that would become my corpus, and the many problems involved in its collection and subsequent transcription.

I looked at what should be counted as a verdict of suicide, and what should be counted as a suicide, in the light of all the data I initially collected. These decisions are very

---

[17] However the Wmatrix website gives an accuracy figure of 92%.

much more problematic than many people might think. Having discussed how I collected the corpus, I described how I labelled different parts of it. I then introduced some corpora that I could use for comparison. I also looked at some potential concerns regarding the comparisons of non-homogeneous corpora. I then presented some statistics relating to the data. Next, I looked at just how short, and how long, some suicide notes are, and the proportions of notes of various lengths that were in the corpus. The chapter ended with a description of the software I use in this thesis.

My final collection of data comprised a viable corpus of 286 suicide notes from the Birmingham Coroner's Office, and a corpus of 66 suicide notes (half of which were not genuine but fabricated) from Shneidman and Farberow (1957e). As well as a considerable variation in the length of the texts, for the Birmingham corpus there was a discrepancy between the number of notes and the number of subjects, some subjects having written more than one note. With the corpus in place, it was now time to perform some analyses on it, and the following chapters, between them, would satisfy this aim.

5. SUICIDE NOTES: ANATOMY AND ODDNESS

5.1. Introduction

Suicide notes can be difficult to analyse, not least because they tend to be short (see Chapter 4). This can aggravate the problem of trying to tell whether a suicide note is genuine or falsified. As shown in Chapter 3, until now most methods proposed to aid the distinction between real and fabricated suicide notes have come from psychologists, and usually required the use of statistical techniques. Winter (1996, p.141) points out that such methods are not readily applicable to short texts. This chapter presents a linguistic approach to the problem, which might be effective even for texts too short to admit statistical analysis. I begin by deriving an outline of the anatomy of a typical suicide text from a set of genuine suicide notes by means of analysis at the discourse, rather than lexical, level. It may then be possible, by assessing how well this fits a selection of simulated notes, to identify a group of salient characteristics. Conversely, by analysing a set of simulated suicide notes, then applying the results of this to some genuine notes (a method evolved from one developed by the Birmingham University Forensic Linguistics Group), it may be possible to find a different group of characteristics which can then be used as diagnostic criteria which can contribute some insight into equivocal suicide texts.

My examination of real suicide notes focuses specifically on how they begin and end, some topics within them that appear salient, and some other categories. I shall also look at how each of these features in turn is mimicked by Shneidman's fabricated suicide notes (Shneidman and Farberow, 1957e) (see Chapter 3). The "topics" are based on what other researchers have found, and what I have noticed in suicide notes. Most of the authors discussed in Chapter 3 are referenced again, at various points, in this chapter, and any repetition of their findings here is to aid the reader. This notion of "topic" is more qualitative than the largely quantitative notion employed in Chapters 6 and 7. At the end of this it may be possible to characterise the prototypical real suicide note.

Next, I shall look at the topic of "oddness", and I shall take a different approach from

that employed above and look first at Shneidman's fabricated notes (Shneidman and Farberow, 1957e) seeking any "odd" features that tend to prevail among them. I shall then consider whether any such features are present in real suicide notes. The approach of starting with fake rather than real notes does not seem to have been used before. At the end of this exercise it may be possible to identify the prototypical fake suicide note.

5.2. Beginnings and Endings

Aspects of how the texts start and finish can be categorised using the following category labels: 'Date of Note', 'PostScript', 'Addressee', 'Salutation', 'Valediction', and 'Signature'. Most of these categories have been considered before by various authors, with the exception of postscripts. Addressees, on the other hand, seem to have been included in more studies than any of the others. The existing literature has tended to treat each characteristic in isolation rather than taking a holistic view as I intend to. The following are some of the studies that have used, or mentioned, the above categories.

Dates on notes, or other indications about when they were written, are mentioned in some previous studies. Salib and Maximous (2002, p.159) dealt with notes of suicides who were over 59 years old and noted that very few were dated. Salib, Cawley and Healy (2002), using the same data as Salib and Maximous, found 17% of their notes were dated. Black (1993, p.701), who acquired his own simulated (as well as genuine) suicide notes, found that the genuine notes were dated more often than were the simulated ones.

Addressees are also mentioned in the literature. Gottschalk and Gleser (1960, p.196-197) found that addressees were referenced more in genuine notes than in simulated notes, but they conflated this by including pronouns in the procedure. Darbonne's (1969, p.47-48) categories included one of addressee, but he did not find any significant differences in addressee between his four age groups of genuine suicide note-writers. Arbeit and Blatt (1973, p.290) found that whereas genuine notes tended to include the names of people, simulated notes rarely addressed them directly (see Section 5.7). McClelland, Reicher and Booth (2000, p.229-230 & 235) examined the concept of blame in terms of the relationship between note-writer and addressee. Salib and

Maximous (2002, p.159) said that most of their notes were addressed to people. Salib, Cawley and Healy (2002) stated that 80% of their notes contained explicit addressees. Tadros and McGrath (2004)[18] looked at addressees but found nothing significant in this regard between the age-groups they were investigating. Tuckman, Kleiner and Lavell (1959, p.60) mention addressees and salutations but do not give any figures for them.

Gregory (1999, p.137-138 & p.149) mentions **salutations** and **valedictions** as being typically "Dear…" and "Love…", respectively, in ordinary (non-suicide) notes. It is not clear whether he counted them separately or along with other 'affectionate' lexis throughout the texts, but he certainly seems to have included such items within his category of 'positive affect'. **Signatures** are mentioned in three sources of which this author is aware: Tuckman, Kleiner and Lavell (1959, p.60) say that most of their notes were signed, as do Salib and Maximous (2002, p.159); and Salib, Cawley and Healy (2002) who give a figure of 87% for signed notes.

From the existing literature, then, it seems that most genuine suicide notes contain explicit addressees and are signed, but are not dated. Also, it seems that simulated suicide notes rarely contain addressees, and have even fewer dates than do genuine notes. It will be interesting to see how my data compare to these findings.

The finding that genuine notes are not often dated is unfortunate because, clearly, a note's being signed and dated is of primary importance for any legal disputes regarding it. A dated note can be an important piece of evidence with respect to when the text was written, or when its author wanted other people (such as police or coroners) to think the text was written. In this study 'dates' need not include all the trio of day, month and year. They can be just for example, "Tuesday". Although, intuitively, it seems more likely that a date, if present, will occur at the beginning or end of a text, a date may actually occur anywhere in a text. In this study any date that seemed to date the note was counted regardless of where it occurred in the text.

---

[18] It should be mentioned that Tadros and McGrath (2004) use the same data that was collected for this thesis, although the focus of their work was entirely different, comparing notes on the basis of age from a non-linguistic perspective.

Postscripts in a text may indicate that the text has been re-used at a later date. Postscripts need not include an explicit "P.S." denoting them. If the note contains some farewell lexis followed by a signature, for example, and then continues with more text, the latter can be counted as a postscript. A postscript can also occur in the 'middle' of a note. For example, if the note contained an explicit "p.s.", but the text then continued and had a signature nearer the end of it, the "p.s." would count as a 'Postscript'. There are various permutations of all this, and although what to count is ultimately a matter of the researcher's judgement, it is considered that these categories are generally uncontentious, and obvious and relatively easy to determine. The following corpus extracts are examples of postscripts.

(5.1. Postscript – without 'p.s.')
"Sorry for Inconvience" [Sic] (F96431)

(5.2. Postscript – with 'p.s.')
"P.S. look after your Mom for me." (M98036e)

The 'Addressee' category is for the name or role of the person, or persons, (or other entity or entities) to whom the note was written. This may be, for example, the note-writer's family, a partner, the police, or the author's cat. To count as an instance of 'addressee' in this study an address term had to be explicitly stated and had to occur at, or near, the beginning of a note. The addressee was taken to be the first addressee encountered in a note: some notes directed different parts of their text to different people. Here are some example addressees.

(5.3. Addressee – title & name)
"Mrs Smith" (F97158c)

(5.4. Addressee – role)
"FLORIST" (M95214f)

'Salutation' in the context of this section is concerned with how the addressee is addressed. Regarding the salutation-addressee as a pair, I have counted each in terms of function rather than form, that is, what they are rather than where they appear within the

pair. For example, if occurring at the start of a note, "Dear" (not followed by any name), is counted as a salutation, but "Mary" (not preceded by any actual salutation such as "Dear") is not counted as a salutation (but is counted as an addressee). Beginnings such as "Mary Dear" are counted as addressee and salutation, respectively. (It could be argued that the actual semantics of the latter example should be interpreted as salutation ("Mary") followed by addressee ("Dear"), partly owing to the usual order of a salutation preceding an addressee, and partly because it cannot be known for certain exactly what the note-writer meant. However, I have counted these terms as described above.) I have dealt with instances of "Dear Wife" by counting "Wife" as the addressee, and with instances of "Dear Wife Mary" by counting only "Mary" as addressee (and "Dear" and "Dear Wife" as the salutations, respectively). Thus, "Dear Wife Mary" counts as only one salutation, which is "Dear Wife", and only one addressee, which is "Mary". The extracts below show an example salutation and an example salutation-addressee pair.

(5.5. Salutation)
"My Darling" (F96302)

(5.6. Salutation-Addressee)
"Dear Social Worker" (F96049b)

'Valediction' included, for example, "Yours Faithfully". The dilemma of whether to count instances of salutation in terms of form or function also applied to valedictions. That is, the question of whether or not a valediction could occur at the beginning of a note had to be considered. Ultimately, unlike the postscripts (see above), I prioritised form (position) over function. So valedictions were only counted when they were at the end of a text (just as salutations were only counted when they were at the beginning). Here is an example of a valediction.

(5.7. Valediction)
    "all the best for now
and Merry Christmas" (M95468d)

'Signature' included any partial or full names, or pseudonyms, of the notes' authors. "Fred" was counted in this category, as was "Dad". The following examples show a signature and a valediction-signature pair.

(5.8. Signature)
"Bill Smith" (M96242b)

(5.9. Valediction-Signature)
 "Bye
Mary" (F96049)

The Birmingham notes were split into sub-corpora of males and females (see Chapter 4) partly to see any gender differences that might occur, but primarily to enable some comparison with the Shneidman notes because the Shneidman notes were all written by males. The categories were counted 'by note'. That is, a maximum count of 1 per note was made regardless of how many times a category occurred within an individual text. The figures (numbers) below are the totals of counts per category, per sub-corpus, calculated as percentages of the total number of notes in the sub-corpora. When comparing the Shneidman notes with each other, or with the Birmingham notes, I have taken a somewhat arbitrary cut off point of 10% as being important for differences between the sub-corpora. This represents approximately four notes in the Sims and in the Gens. Intuitively, 10% seems a reasonable amount of difference for a single category, but fewer than three texts would seem to be an extremely dubious difference. (For both the "Sims" and the "Gens" 3.03% represents one note (100/33), therefore three notes are 9.09% of each sub-corpus. A cut off point of 10% will equate to more than three notes, and should be a somewhat simpler calculation than the 12.12% which actually represents four notes.)

In Figures 5.1 to 5.4, below, the categories 'Date of Note', 'PostScript', 'Addressee', 'Salutation', 'Valediction', and 'Signature' are shown as columns 'DoN', 'PS', 'Add', 'Sal', 'Val' and 'Sig', respectively. The vertical axes in the Figures represent the percentages of the total number of texts.
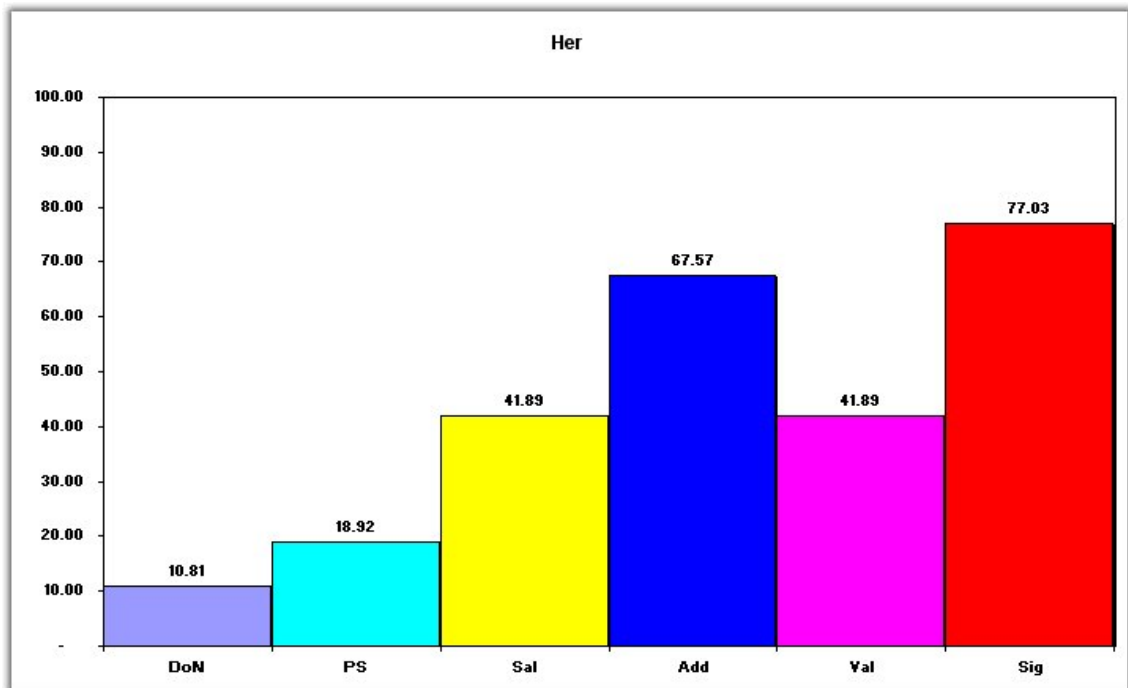
Figure 5.1. Beginnings & Endings: Birmingham females
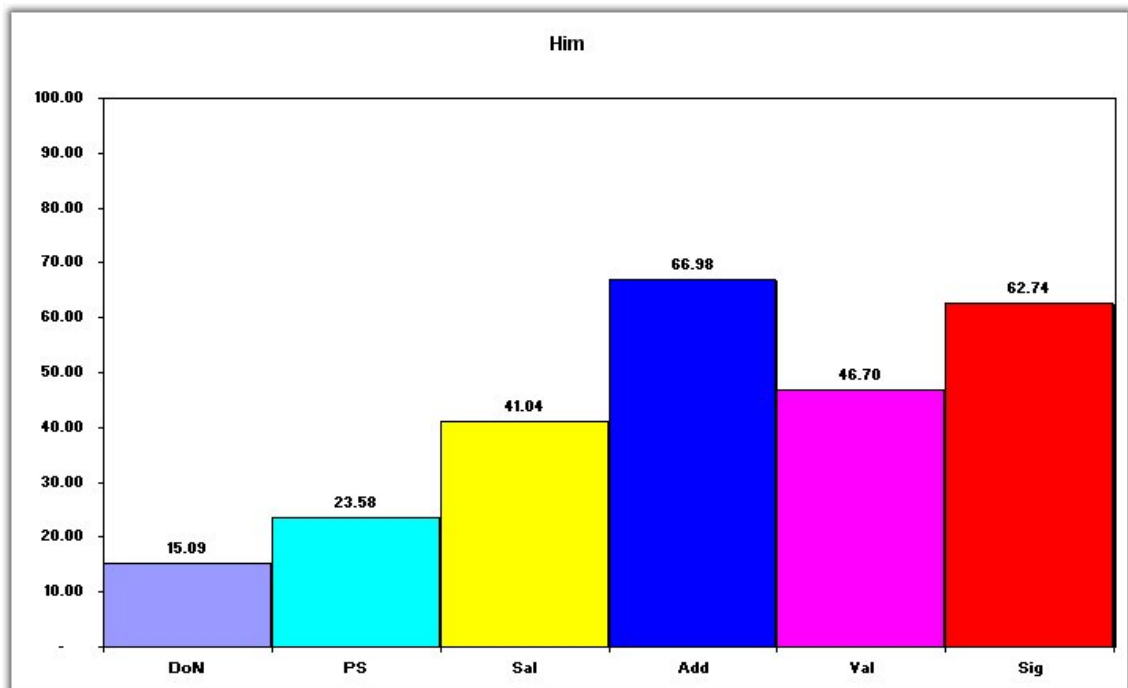
Figure 5.2. Beginnings & Endings: Birmingham males

Figure 5.1 shows that not too many of the "Her" notes are dated (just over 10%) or postscripted (fewer than 20%). It also shows that over 67% of the "Her" notes are addressed and over 77% are signed. And equal amounts, over 41%, of the "Her" notes have salutations and valedictions.

As can be seen in Figures 5.1 and 5.2, these results for the Birmingham males are similar to those for the Birmingham females in terms of the overall relative distribution of frequency of the categories (shown by the height of the columns relative to each other). Further, there are similar percentages of notes with salutations in both sub-corpora, and similar percentages of notes with addressees in both sub-corpora. However, the males used over 14% fewer signatures.
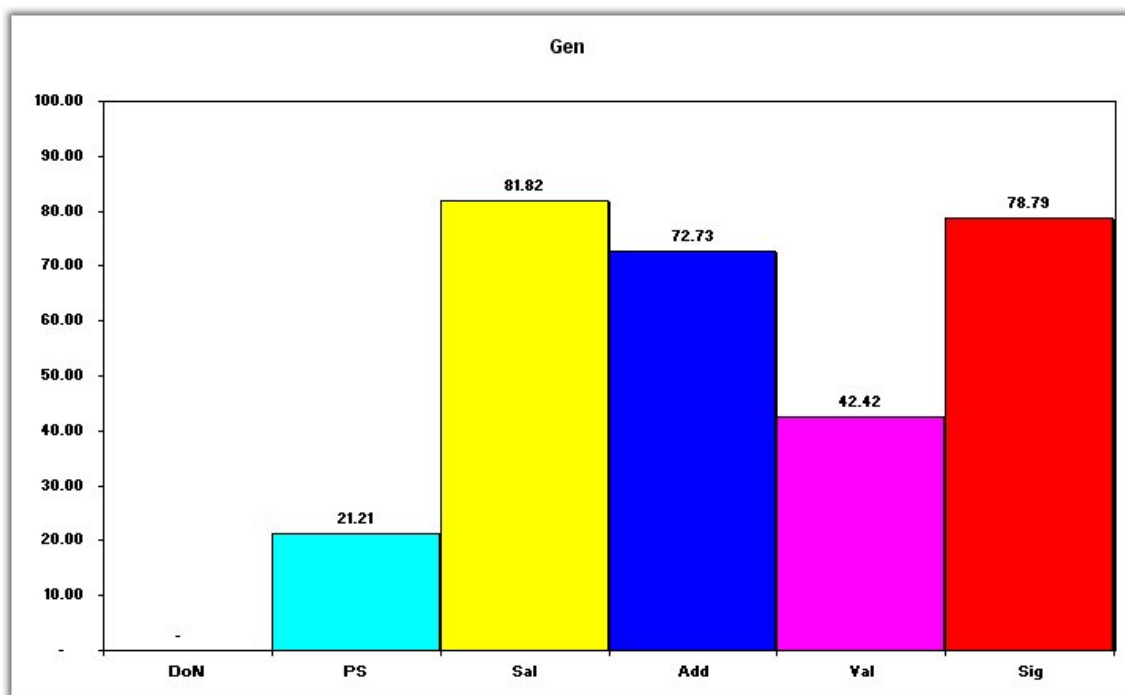
Figure 5.3. Beginnings & Endings: Shneidman Genuines

As Figure 5.3 shows, Shneidman's Genuine notes are somewhat different from my Birmingham corpus in terms of the relative frequency distribution of the six categories. In the "Gens", there are no dates. Also the "Gens" have a proportion of salutations which is roughly double that in the "Hims" (81.82% as opposed to 41.04%), and

considerably more signatures, than the "Hims". Of course, any dates could have been lost in Shneidman's transcription, or their absence could be a 1940's versus 1990's phenomenon or a difference between U.S. and U.K. note-writers. The latter two points might also account for the differences in the numbers of notes with salutations and signatures. There is a further point that might account for the differences between the "Gens" and the Birmingham corpus. The figures suggest that there is little uniformity between the notes and that any corpus, particularly a very small one, may be different in almost accidental ways from any other corpus. Therefore, the larger the corpus the more reliable it should be.
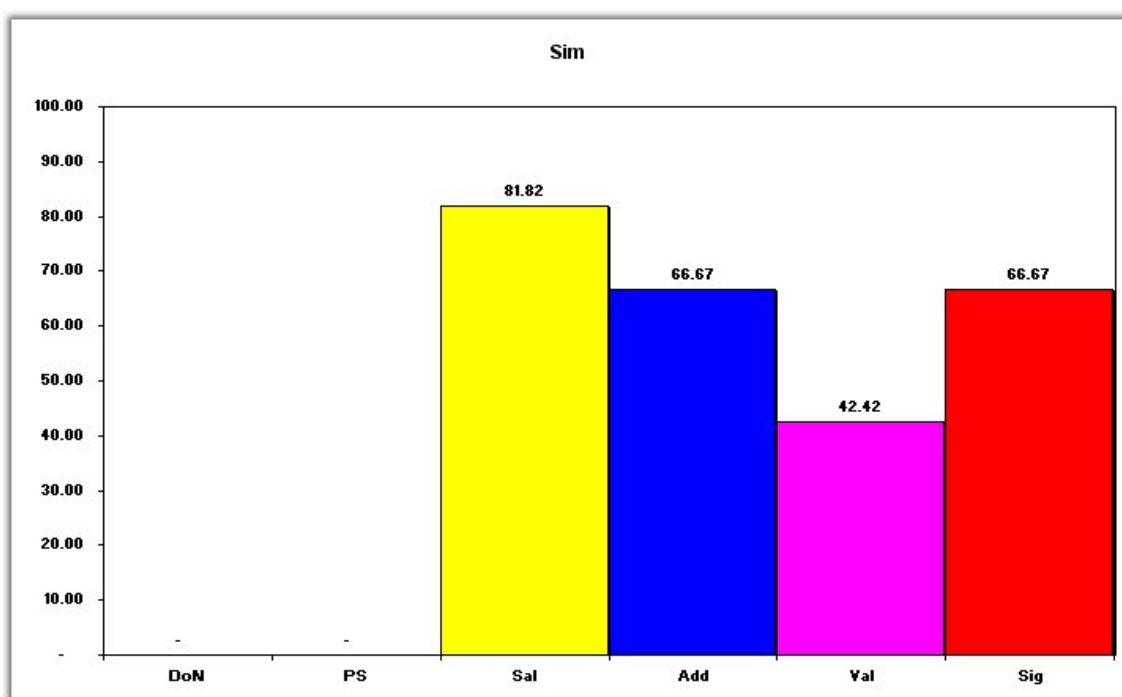
Figure 5.4. Beginnings & Endings: Shneidman Simulateds

From Figure 5.4 it can be seen that Shneidman's Simulated notes contain no dates or postscripts, whereas over a fifth of the Genuine notes, as shown in Figure 5.3, contained postscripts. This could be significant, but, as mentioned above, they could have been lost in transcription. The "Sims" also have over considerably fewer signatures than the "Gens" (66.67% as opposed to 78.79%).

If one sets a threshold of differences between sub-corpora of 10 or more percentage points, it can be concluded that the "Hims" are more like the "Hers" than the "Gens"; and the "Gens" are more like the "Sims" than the "Hims". However, caution is needed in taking any small group of texts, such as the Shneidman notes, as being representative of suicide notes in general. They are, largely, inadequate for distinguishing between genuine and fake notes. That they do not really stand up as two distinguishable corpora is itself an important finding.

Of course, this does not reveal how many notes have various combinations of the categories. Tables 5.1 and 5.2 show the combinations as percentages of total notes in each sub-corpus. In Table 5.1 'Both' means that both salutation and addressee are present in a note, and 'Neither' means that a note had no salutation nor any addressee in it. In Table 5.2 'Both' means that both valediction and signature are present in a note, and 'Neither' means that a note had no valediction nor any signature in it.

Table 5.1. How Suicide Notes Begin

	Her	Him	Gen	Sim
Salutation only	1.35	1.89	12.12	18.18
Addressee only	27.03	27.83	3.03	3.03
Both	40.54	39.15	69.70	63.64
Neither	31.08	31.13	15.15	15.15

Table 5.1 shows the "Hims" and "Hers" to be similar to each other, and the "Gens" and "Sims" to be similar to each other in terms of the relative amounts of the features, but the Birmingham notes have a different profile from the American notes. Whereas very few American notes (one "Sims" and one "Gen" in fact) contain only an addressee and no salutation, the feature associated with the smallest figures for Birmingham notes in Table 5.1 is that of salutation and no addressee.

117

Table 5.2. How Suicide Notes End

	Her	Him	Gen	Sim
Valediction only	4.05	4.72	-	6.06
Signature only	39.19	20.75	36.36	30.30
Both	37.84	41.98	42.42	36.36
Neither	18.92	32.55	21.21	27.27

Table 5.2 shows that very few notes in any sub-corpus contained only valedictions. Most notes were signed either with or without valedictions. However, far more "Hims" than "Hers", or "Gens", had no valediction and no signature. And far fewer "Hims" than the other sub-corpora contained only signatures. Looking for differences of 10 percentage points or more, it can be seen that the "Hims" are no more like the "Hers" than they are like the "Gens"; but the "Sims" are similar to the "Hers", "Hims" and "Gens". Indeed, there is no clear pattern in Table 5.2 (unlike in Table 5.1). This is important because it shows that one cannot say that the "Sims" are different from all the three real-note corpora in respect of 'how suicide notes end'.

A couple of other features of notes (including notes other than suicide notes) that are typically found near their beginnings and ends are dates and postscripts. How prevalent are dates and postscipts in suicide notes? Can the dates on notes suggest advance planning of the suicide act? Can the presence, or absence, of a date and/or a postscript indicate that a note has been fabricated? It is possible that a date might be used to cast suspicion on a note's being genuine or not – but is the date data reliable? I now want to consider these questions.

There has been some suggestion, as mentioned above, that postscripts might indicate a more current, contemporaneous, kind of writing, and thus, perhaps, signify a genuine rather than fake suicide note; alternatively, if a note was left over from a previous suicide attempt, a postscript could have been added to "update" the note. There might even be clues to either of these scenarios, and to the general question of note "reuse", by examining those notes which were both postscripted and dated, and also by comparing the dates of dated notes (whether postscripted or not) with the dates of death. Although dates of death were not available for the Shneidman notes I have looked at dates in the

"Hers" and the "Hims", and I have looked at postscripts in all the four sub-corpora. Although relatively few notes were postscripted and even fewer were dated in the sub-corpora considered above, this was considered worth investigation because it might have revealed a small part of the picture about suicide notes. In Table 5.3, 'p.s.' means postscript, and all the figures are percentages.

Table 5.3. Percentages of notes with Postscripts and Dates

	Her	Him	Gen	Sim
p.s. & no date	14.86	17.45	21.21	-
date & p.s.	4.05	6.13	-	-
date & no p.s.	6.76	8.96	-	-
no p.s. no date	74.32	67.45	78.79	100.00

From Table 5.3 it is clear that most notes, especially the "Sims", have neither postscripts nor dates; very few contain both postscripts and dates; and postscripts are more prevalent than dates. Clearly, there are not enough postscripts nor dates in the data to enable any definitive analysis to be made. However, regarding fabrication or otherwise, about a fifth of the "Gens", a quarter of the "Hers" and a third of the "Hims" do contain a postscript or a date, or both, and this itself is an interesting finding.

Inspection of all dated notes versus dates of death (in the "Hers" and the "Hims") showed that many notes were dated one day before the recorded date of death; other notes were dated two, three and four days before the date of death; one was 23, another was 29 days "pre-dated"; one was dated three months before death. Before surmising about the significance of any of this, it should be mentioned that some anomalies were present in the data. There was one note which was dated two months after the death. Another six notes by this note's author were all dated with the day before his official date of death (four included postscripts). I did not collect sufficient data to be able to determine how frequently bodies tended to be found on the actual day of death as opposed to being found a number of days after death, but it seems reasonable to assume that many officially recorded dates of death could be inaccurate estimates, at least in terms of the few hours which would establish the date as being early on one particular

day or late on the previous day. It seems that some suicides are planned in advance, and the notes can be written several weeks before the act itself takes place.

Some notes might have been re-used, but it is difficult to know which they were. There is at least one clear example of note re-use: one "Her" subject re-used her various declarations including her belief in euthanasia (see Chapter 4, Section 4.4). It is of course possible that some of the notes, particularly those which were dated more than 20 days prior to death, had only been retained by the Coroner's Office as samples of the deceased's handwriting and were not really suicide notes at all. This possibility became clear when a note (neither dated nor postscripted) was found in which the subject said he was 36 - although he died aged 38. But it seems at least equally possible that most of theses oddities were simple errors made by the notes' authors. As an aside, one note, which amounted to a last will, was clearly dated, but the heir's name was illegible. This might imply that even if a suicide is being careful by dating their note, rather than merely being habitual in dating their writing, their concentration may lapse at times during the writing process. One must assume that dates on suicide notes cannot always be relied upon to be accurate, and caution should be used when including them as evidence. On the other hand, one should always query whether an 'inaccuracy' is in itself an indication of non-authenticity.

What is remarkable in all this is that so few notes are dated, less than three-quarters have an explicit addressee, and over a fifth are not even signed. This holds true for all the sub-corpora. Since many people seem to see a suicide note as a means of making known their final wills and wishes (see Section 5.4 below), it is a significant finding that that these documents may, sadly, not have the legal validity that their writers imagined they would have.

5.3. Topic Categories

In this section I look at several categories that appear at first sight to be ubiquitous in the Shneidman notes. The categories consist of topics that are typically expressed by the phrases indicated in Table 5.4. (Please see the further discussion of topics in Chapters 6 and 7.) Most if not all of these have been mentioned by other authors, and many have

been used as categories, or subcategories, in their own right, but the labels used for them here are my own. I will give a few examples of how other authors have used them before presenting the figures I obtained.

Shneidman & Farberow (1957c, p.7) found more blame in genuine suicide notes. Osgood and Walker (1959, p.65) mentioned references in suicide texts to goodbyes and farewells, and concepts of forgiveness and of the 'way out'. They discuss the associations of these concepts with each other, and also mention that writers of simulated notes, too, included references to a 'way out'. Tuckman, Kleiner and Lavell (1959, p.60) looked at hostility and affect, and these categories seem to overlap some of the categories used in this section. Ogilvie, Stone and Shneidman (1966, p.534) found more references to the word "love" in genuine notes. Jacobs (1967, p.67-69) found forgiveness was a characteristic tendency of two of his six categories of suicide notes, and references to the 'only way out' were characteristic of many notes. Three of his other categories included blame and being sorry or not. Darbonne (1969, p.47) had sub-categories of 'goodbyes and farewells', 'can't go on', 'tired of life' and 'fault'. Arbeit and Blatt (1973, p.286-288) investigated the judgement of various groups of people regarding whether a set of sentiments was indicative of a genuinely suicidal person. Their sentiment categories included self-blame, positive affect (including love), expressions of suicide being the "only way out", and the apparently general "for the benefit of others". Lester and Leenaars (1988, p.106) successfully fitted Shneidman's 66 notes into Jacobs' (1967) note categories, but found that forgiveness occurred more in simulated notes than in genuine ones, while blame occurred more often in genuine notes. Gregory (1999, p.139), pooling other people's categories, noted aspects of positive affect, blame, 'the world is too cruel' and 'the only way out'. McClelland, Reicher and Booth (2000, p.230 & p.232) looked at fault, love, apologies, forgiveness, and expressions of inability to continue (with life), being a burden, a failure, and suicide being the best solution. Salib and Maximous (2002) mention that some of their notes contained references to guilty feelings. Salib, Cawley and Healy (2002) grouped guilt, regret and blame together within their categorisation scheme.

The examples of each category in Table 5.4 are all my inventions, (although some

probably do exist in the suicide texts in the actual form (lexis and syntax) given below).

Table 5.4. Topic Categories

Category	Example realisations
BeenGood	it's been a good life
BestWay	this is the only way; this is best for you
Burden	I don't want to be a burden; I wouldn't wish to drag you down
Can'tFind	I can't find my place in life / this world
Can'tTake	I can't take any more; I can't stand this; I can't live without you
Fault	it's my fault; it's your fault
FindBetter	find someone / something better than me; I hope you find someone else
Forgive	forgive me; don't think ill / badly of me
Goodbye	Goodbye; Bye; Farewell
GrieveNot	don't grieve for me; don't feel bad about this
IFailed	it's all my own doing
Ilove	I love you; I have much love for our son
NoMore	There's no more to say; I can't say any more; there's nothing else to tell you
Sorry	I'm sorry for what I did; I apologise for this suicide; you'll be sorry; I'm not sorry
TakeCare	take care of Betty; be good to Betty
Tell	tell Mary that I love her; don't say how I died
Tired	I'm tired of living; I'm tired of trying
WillHelp	Betty will help you through this; my savings should make things easier for you
YouBest	you've been the best wife anyone could have had

Note that 'Fault' includes (and doesn't distinguish between) "it's my fault" and "it's your fault". 'IFailed', on the other hand, includes only instances where the note-writer had expressed the view that they themselves have failed in some way. 'Goodbye' is reserved for explicit expressions of saying "Goodbye", but includes words such as "Bye". 'Sorry' refers to the suicide itself, or events related to it; so it excludes instances such as "I'm sorry I didn't see you before I died". 'TakeCare' is an instruction referring to people (or animals) other than the note-writer, rather than objects or things, and excludes instances such as "take care of my bills", "Fred will take care of my body", "you'll be taken care of" and "my money will take care of you". 'Tell' includes both negatives and positives, for example, "tell Mary I love her" and "don't say how I died".

The former example illustrates that these categories are not mutually exclusive[19]: there would be one count for 'Tell' and also one for 'Ilove'. 'WillHelp' means 'someone/something will help you', and so does not include the more indirect instances of someone (or something) helping the addressee, such as "Fred will take care of my body", but does include instances such as "my money will take care of you". (Note that this is different from the 'TakeCare' category outlined above.)

Again, as described above in Section 5.2, the figures in Tables 5.5, 5.6, 5.7 and 5.8 are 'by note', and comparisons have a cut off point for differences of 10 percentage points.

Table 5.5. Most Frequent Topics: percentages of Shneidman notes

Gen			Sim	
Ilove	42.42	BestWay		30.30
BestWay	30.30	Ilove		30.30
Sorry	30.30	Sorry		27.27
Fault	27.27	Forgive		24.24
Forgive	27.27	Can'tTake		18.18
Goodbye	21.21	Burden		18.18
Can'tTake	18.18	WillHelp		15.15
Tell	18.18	Tell		12.12
TakeCare	15.15	BeenGood		9.09
NoMore	12.12	FindBetter		9.09

Table 5.6. Least Frequent Topics: percentages of Shneidman notes

Gen			Sim	
FindBetter	6.06	Goodbye		9.09
Tired	6.06	IFailed		9.09
WillHelp	6.06	Tired		9.09
YouBest	6.06	Fault		6.06
BeenGood	3.03	NoMore		3.03
Can'tFind	3.03	TakeCare		3.03
GrieveNot	3.03	YouBest		3.03
IFailed	3.03	Can'tFind		-
Burden	-	GrieveNot		-

---

[19] Indeed, none of the 19 categories was counted as being mutually exclusive with respect to any of the other 18.

As Table 5.5 above shows, the first three items ('Ilove', 'BestWay' and 'Sorry') along with 'Forgive', 'Can'tTake' and 'Tell' were used in at least 12% of the Simulated notes as well as the Genuine notes. However, the "Sim" note-writers seem to have used 'Fault', 'Goodbye' and 'TakeCare' far less than the "Gen" note-writers. As Tables 5.5 and 5.6 above also show, although none of the genuine note-writers wrote about being a burden, 18.18% of fakers did. Of course it should be borne in mind that 3.03% represents only one note.

In this section, because of the categories being derived from looking purely at the Shneidman notes, I have examined the "Gens" and "Sims" first, before considering the Birmingham notes. I now present the latter in Tables 5.7 and 5.8.

Table 5.7. Most Frequent Topics: percentages of Birmingham notes

Her			Him	
Fault	27.03		Ilove	43.40
Sorry	25.68		Sorry	42.92
Ilove	24.32		Fault	31.60
BestWay	18.92		Can'tTake	21.23
GrieveNot	16.22		Forgive	14.62
Can'tTake	14.86		Tell	13.21
Forgive	8.11		TakeCare	13.21
Goodbye	8.11		Goodbye	12.74
Tell	6.76		BestWay	11.32
TakeCare	6.76		IFailed	7.55

Table 5.8. Least Frequent Topics: percentages of Birmingham notes

Her			Him	
Burden	5.41		GrieveNot	6.60
BeenGood	5.41		WillHelp	4.25
Tired	2.70		YouBest	3.77
IFailed	-		NoMore	2.36
WillHelp	-		Burden	1.89
YouBest	-		FindBetter	1.89
NoMore	-		Can'tFind	0.94
FindBetter	-		BeenGood	0.47
Can'tFind	-		Tired	0.47

What is the significance of these findings? It transpires that many of these apparently salient phrases, when actually counted, are not indeed present to any great extent in either the "Sims" or the "Gens", nor in the "Hers" or the "Hims". Of the 19 categories in this section, only 5 differentiated the two sets of Shneidman notes in terms of a difference of 10% or more. More "Gens" than "Sims" contained instances of 'Ilove', 'Goodbye' and 'TakeCare', and far more "Gens" than "Sims" contained instances of 'Fault'. However, more "Sims" than "Gens" contained instances of 'Burden'. It could be speculated that the "Sim" note-writers were familiar with the most frequent topics of suicide notes, but that they overestimated the extent to which suicidal people feel they are becoming a burden to others, and they underestimated the extent to which suicides want to assign or assuage fault, explicitly express love and say goodbye.

The most distinctive topic categories in both the "Hers" and the "Hims" were 'Fault', 'Sorry', 'Ilove', 'Can'tTake' and 'BestWay'. In the "Hers" another distinctive category was 'GrieveNot'. In the "Hims" 'Forgive', 'Tell', 'TakeCare' and 'Goodbye' were also distinctive categories. In general, the most and least distinctive categories were the same in the Birmingham notes as they were in Shneidman's "Gens". One exception to this was the category 'GrieveNot' which was used by very few of the "Gens" (and indeed none of the "Sims"), but occurred in 16.22% of the "Her" notes. Interestingly, occurring in only 6.60% of the "Hims" makes 'GrieveNot' the main topic category to distinguish between the "Hers" and the "Hims". This research suggests that topic, as defined here, is only partially successful in distinguishing between real and fake suicide notes.

5.4. Meta Categories

Next, I looked at categories (mostly used my others) which I labelled 'Who', 'Where', 'Why', 'How', 'Events', 'Memories', 'Humour', 'Instructions' and 'Trivia'. These are mostly self-explanatory, but some explanations are given below.

These categories (with the exceptions of 'Who' and 'Where') seem to occur at a relatively higher level of discourse than the other categories in this chapter (which is why I have called them "meta categories"). Whereas the topic categories used in

Section 5.3 (above) were typically realised by the example lexis and phrases given in Table 5.4 (and were thus often visually salient in the text − and consequently relatively easier to count), the categories here did not generally seem to be realised by any stereotypical set of words or phrases, but by a larger vocabulary and over several sentences or even paragraphs. (Also, they are not all truly concerned with 'topics' in the linguistic sense, as such.) The exceptions mentioned above, 'Who' and 'Where', intuitively seem to be expressible with a relatively short span of words. For example, saying 'I am Fred Bloggs' or 'I am in the garage' does not require a great deal more text to convey the essential message. Of course, the other categories here can take similarly short forms. One could find in 'Humour', 'Here's a joke...', or in 'How', 'I'm going to hang myself with this rope', but it seems less likely to be the case.

Before the explanations here are some references to other studies that have used these categories. Indeed, with exception of 'Who', 'Memories', and possibly 'How', they have all been used, or at least mentioned, by other authors to various degrees.

'**Where**'. Capstick (1960, p.1181) found that 14 of his 136 Welsh suicide notes contained warnings to those who might find their author's body (and that the purpose of these was to "reduced stress" in others and/or to advise them of danger, such as a room full of gas).

'**Why**'. Capstick (1960, p.1180) found that over 50% of his note-writers tended to "explain" why they were killing themselves. Darbonne (1969, p.47-48) had a category for "reasons stated" (and found the over 59 year-olds gave more of them). Gregory (1999, p.138 & p.148) also had a category for reasons, and found the inclusion of reasons was more typical of simulated, as opposed to genuine, suicide notes, but that the reasons given in the genuine notes were very detailed. McClelland, Reicher and Booth (2000, p.230), with one primary category of 'blame', and a secondary category (under blame) of 'reasons given', found that 104 of their 172 notes contained reasons.

'**How**'. Although researchers have looked at details of the methods employed by suicides, there does not seem to be anyone who has considered the tendency of note-

writers to discuss this in their texts.

'**Events**'. Osgood and Walker's (1959, p.62.) 'time' category (which yielded no significant differences for them) seems to encompass some of my notion of 'Events'. McClelland, Reicher and Booth (2000, p.232) found that 111 of their 172 texts contained mitigation of self-blame, and that 10% of the 111 also contained 'Events'. There does not seem to be any substantial study of '**Memories**' in suicide notes.

'**Humour**'. Humour is, apparently rare in suicide notes. Jacobs (1967, p.72) found that only four of his 112 notes contained any humour. O'Donnell, Farmer and Catalan (1993, p.47) found only one incidence of humour in their 37 suicide texts from suicides on the London Underground.

'**Instructions**'. Shneidman and Farberow (1957a, p.255) found "definite instructions" in suicide notes. Shneidman and Farberow (1957c, p.7) found that many of their genuine notes contained instructions. Capstick (1960, p.1180-1181) found 30 of his 136 notes contained instructions. Ogilvie et al. (1966, p.531-532) looked at instructions to females where a female occupied the subject position in the sentence, and counted the number of them and whether they were specific or not (see Chapter 3, Section 3.1.4). (They found more instructions in the simulated notes, but twice as many specific instructions in the genuine notes.) Edelman and Renshaw (1982, p.104) mention Ogilvie et al.'s (1966) finding that specific instructions were more prevalent in genuine notes. Jacobs (1967) divided his suicide notes into six different categories, two of which were 'wills and testaments' and 'instructions'. In this section I count both of these as instructions. Darbonne (1969, p.48) found instructions tended to be given most by suicide note-writers who were 50-59 years old, and least by those who were 20-39 years old. Arbeit and Blatt (1973, p.291-292) found that people who are better at distinguishing simulated from genuine suicide notes considered that instructions were an important indicator. Lester and Leenaars (1988, p.106) found instructions, including wills, occurred more often in genuine suicide notes than in simulated ones. O'Donnell, Farmer and Catalan (1993, p.47) found their notes tended to contain instructions. Black (1993, p.701), using his own collection of simulated suicide notes for comparison,

found that genuine notes contained more instructions. Etkind (1997, p.14) stated that about a third of all suicide notes contain "mundane instructions". Gregory (1999, p.138 & p.149) confirmed that instructions were more likely to occur in a genuine note. McClelland, Reicher and Booth (2000, p.230) found that in notes that did not contain blame 13% contained instructions; and, apparently, in 50 of their 172 notes instructions were also represented, but as a sub-category of 'blame'. Salib and Maximous (2002, p.159) and Salib, Cawley and Healy (2002) (who used the same data as Salib and Maximous (2002)) found that 17% of their notes were mostly concerned with instructions.

'**Trivia**'. Shneidman and Farberow (1957d, p.32-33 & p.35) looked at logic in suicide notes and stated that trivia and detail can be a symptom of the type of logical error that they call "semantic fallacy" in which the note-writer conflates how others see him with his own view of himself. According to Shneidman and Farberow (1957d) this particular type of error is common in genuine suicides. Frederick's (1969, p.26) review of the suicide note literature includes the statement (closely paraphrased from Shneidman and Farberow, 1957d, p.33), "There is concern with trivial and practical details such as repairing automobiles, distributing goods, and cancelling appointments.". Etkind (1997, p.8) states that suicide notes are "jammed with petty details". Shneidman (1999, p.303) says that he is not surprised that sometimes suicide notes are full of trivia because their writers are so ambivalent. While a rather confusing range of causes have been adduced to account for such findings, it does seem to be accepted that trivia tend to form part of the typical suicide note.

Here now are the promised explanations of the categories, along with a breakdown of what was, or was not, counted as falling within each of them. It should be noted that the categories are not mutually exclusive (so, for example, an instruction might be trivial, or an event might also be a memory). Where it seemed that there was an instance of any category, but it was not too clear or obvious, I opted to give it the benefit of the doubt and included it in the count for that category.[20] An example of this is "I never seem to

---

[20] This is the opposite of the approach taken with the "Topic Categories" (see Section 5.3) where the benefit of doubt was not given.

feel happy deep inside…" (F95456a) which I have counted as a 'memory' although the note from which this quotation is taken is set, almost entirely, in the present tense.

'**Who**' is any occurrence of the note-writer explicitly stating who they are. This excludes signatures unless they are of the form: 'I am, yours faithfully, Joe Bloggs'. However, I have included in this some exceptions to the explicitness principle where the note-writer seems to be uniquely, and easily, identifiable due to other aspects of their text (such as their having included a relative's name, role and address).

'**Where**' is any occurrence of the note-writer stating where they are. This need not be explicit, for example, "Bedroom" as a lone paragraph without any prior or following context concerning, say, where certain assets are located, can reasonably be taken to mean that that is where the note-writer's body can be found.

'**Why**' refers to text given about the note-writer's reasons for their suicide. A large amount of text that attempts to explain the reasons is counted as 'Why'. A very small amount of text that succinctly gives some explanation also counts as being in this category. However, in an attempt to capture all variations of explanations for suicide, any implicit reason is also counted, as is any instance such as "I don't know why I'm doing this". '**How**' follows the same concepts as 'Why', but is with respect to the note-writer's method of suicide rather than their reasons for it.

'**Events**' are references to happenings past, present or future, but are usually concerned with the past. '**Memories**' include passages of the type "do you remember when…" as well as those non-explicitly stated as remembrances. 'Memories' and 'Events' can sometimes be difficult to distinguish from each other, and the resolution into any one particular category necessitated a judgement call. I have tried to keep these two concepts, and therefore categories, mutually exclusive, except where both do seem to be represented in the text. '**Humour**' refers to any attempt at humour, including jokes, or anything which I interpreted as an attempt at humour.

'**Instructions**' are potentially: imperatives; inventories; statements such as "bins need emptying"; clauses containing requests such as "will you please...", "please will you...", "look after...", and "take care of x". I have included in this category, sentiments such as, "take care of x" (where x is, for example, "my son" and the addressee, has not previously had this responsibility). Indeed I have excluded instructions about feelings, but included instructions concerned with physical actions. Thus I excluded sentiments such as "don't hate me", "don't blame me", "be happy", "don't grieve", "don't feel guilty", "forgive me", and "take care of yourself". This was partly because they appear to be ubiquitous and would dwarf the other categories, but mainly because I wanted to find the essence of instructions, rather than any form-related, and possibly formulaic, phrases as most of these examples appear to be.

'**Trivia**' is a difficult concept. In principle, one might consider "Put bins out on Thursdays"[21] and "Library book will need returning" to be typical examples of trivia. However, the context and details are crucial. The first example above is in fact from a landlord to his heir regarding the latter's new responsibilities to his tenants, which arguably renders it non-trivial: it might even refer to some legal contractual arrangement. However, the text from which this example is taken also includes considerable detail concerning various other responsibilities which might well be viewed as trivia. As for the second example above, returning library books is a matter of common courtesy and integrity. If these attributes were of high importance to the note-writer, then returning property to its owner could have been important, although to an outsider it might appear relatively trivial in the context of suicide. The relative length of the section of text in question may also be a relevant criterion. For example, if 90% of a text were concerned with dusting a particular table this would almost certainly count as trivia, but "Wish I had dusted the table" in a 600 word note would probably not. Nonetheless, dusting, cleanliness, and the table itself might have great significance for the author. The remarks here illustrate the difficulties inherent in using non-linguistic categories in a linguistic study. However, trivia has been used as a key point of difference between genuine and fake notes (see above), and it is therefore worth the

---

[21] The "bins" examples given for 'Instructions' and 'Trivia' are essentially the same. The categories are not mutually exclusive. So a single utterance can be placed in more than one category.

effort of finding examples of it, as I have done. In addition, trivia is a subjective concept. What is important to note-writers is not trivial to them. Essentially, it seems there is a conflict between what outside observers may think is trivial for people who are about to kill themselves and what they consider to be important. (One could argue, as Etkind (1997, p.93) does, that the reason people find themselves in suicide situations is that they have lost the ability to distinguish between what is trivial and what is not.)[22] The more the researcher attempts to adopt the note-writer's perspective, the fewer items are likely to be left in the "trivia" basket. However, I have counted all the above examples as 'Trivia'.

Table 5.9 shows the various meta categories and some invented examples for each of them.

Table 5.9. Meta Categories

Category	Examples
Who	I am Fred Smith
Where	I am in the garage
Why	I am in debt
How	I have taken cyanide
Event	You laughed at me when I fell over
Memory	Do you remember when you hit me on the head?
Humour	Gotcha. Ha, Ha!
Instruction	Have me cremated not buried
Trivia	Pay my library fine of £5.76

In Figures 5.5 to 5.8, below, the above categories are shown as columns. In all cases, when I was not totally sure whether to include something within a category I gave the benefit of doubt to my initial intuition and counted it. The vertical axes in the Figures represent the percentage of notes containing at least one instance of the category in question.[23]

---

[22] Etkind (1997, p.63-64) notes that an obsession with trivia can be a symptom of depression.

[23] For each category a maximum of one per note was counted regardless of the number of times that category was represented within any note.
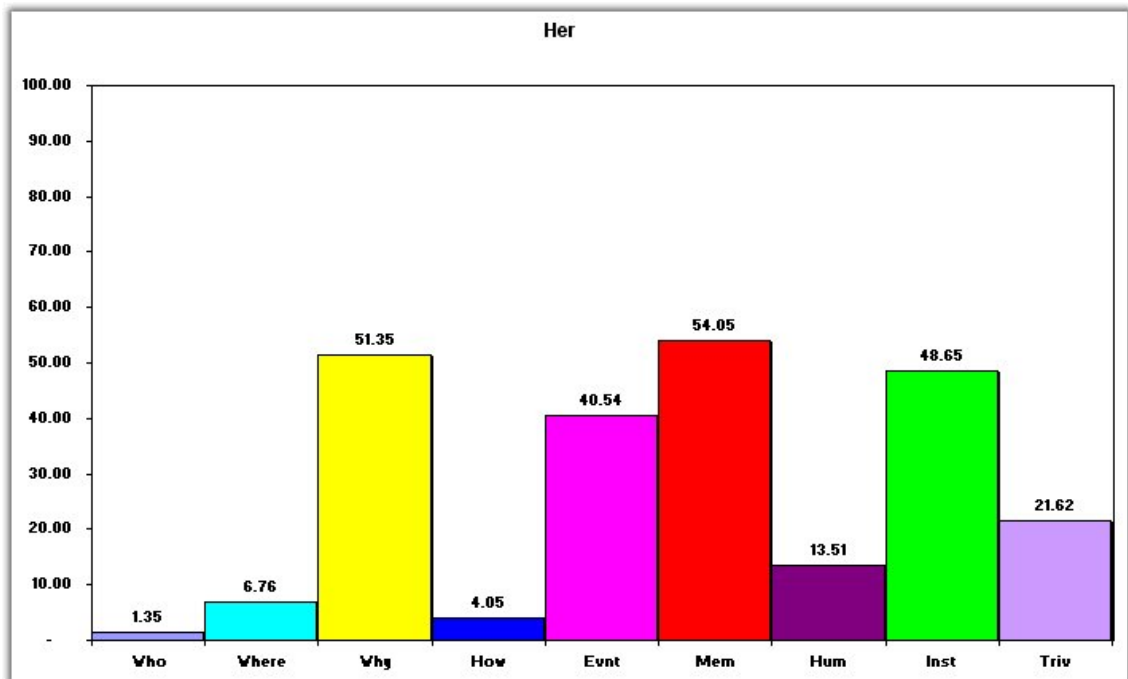
Figure 5.5. Meta Categories: Birmingham females
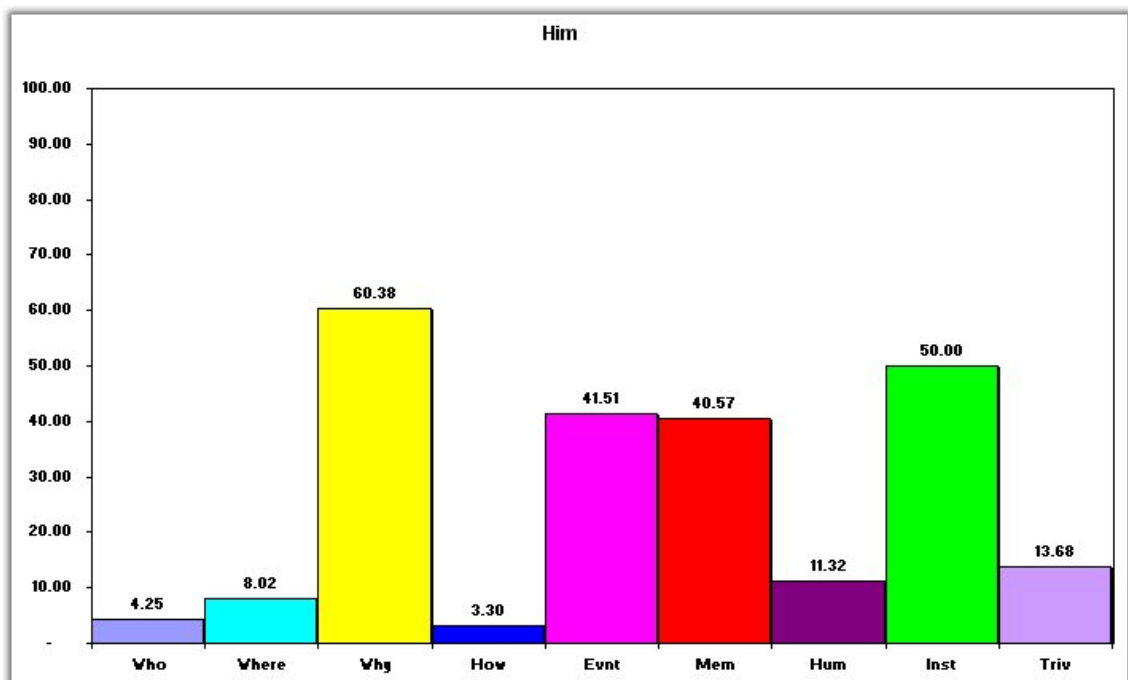
Figure 5.6. Meta Categories: Birmingham males

132

As can be seen from Figures 5.5 and 5.6 the categories that are most heavily represented are 'Why', 'Events', 'Memories' and 'Instructions'. 'Humour' and 'Trivia' are the next best represented, while least frequent are notes containing 'Who', 'Where' and 'How'. As Figures 5.5 and 5.6 show, considerably more of the notes written by females have 'Memories' and 'Trivia' than notes written by males. Males are more likely to mention reasons ('Why') than females. Following the decision to consider differences of 10 percentage points or more as important (see Sections 5.2 and 5.3 above), the most interesting observation here is that the greater number of notes referring to memories was written by females. With the exception of memories, the relative percentages of notes containing each category (as shown by the column heights relative to each other) is similar in the "Hers" and the "Hims". Very few of these Birmingham note-writers stated who they were, where their body could be found, or how they killed themselves. What cannot be seen from the Figures, regarding the notes with 'Instructions', is that, interestingly, over 36% of the "Hers", and over 40% of the "Hims", are by multiple-note writers.
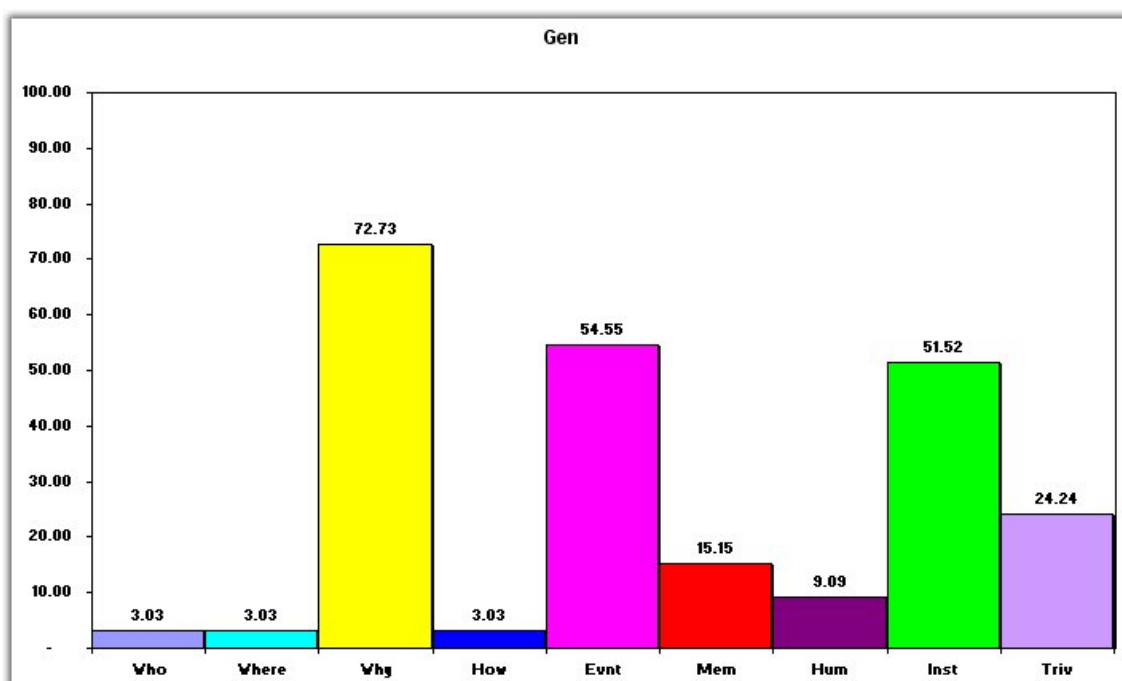
Figure 5.7. Meta Categories: Shneidman Genuines

133

Figure 5.7 shows that, compared to the "Hims", more notes in the "Gens" contain reasons ('Why'), 'Events', and 'Trivia'. However, only 15% of the "Gens" have 'Memories' as compared with over 40% of the "Hims". In terms of the overall pattern of frequency distribution (the column heights relative to each other) the Birmingham males are arguably more similar to the Birmingham females than either category is to the Shneidman Genuine notes.
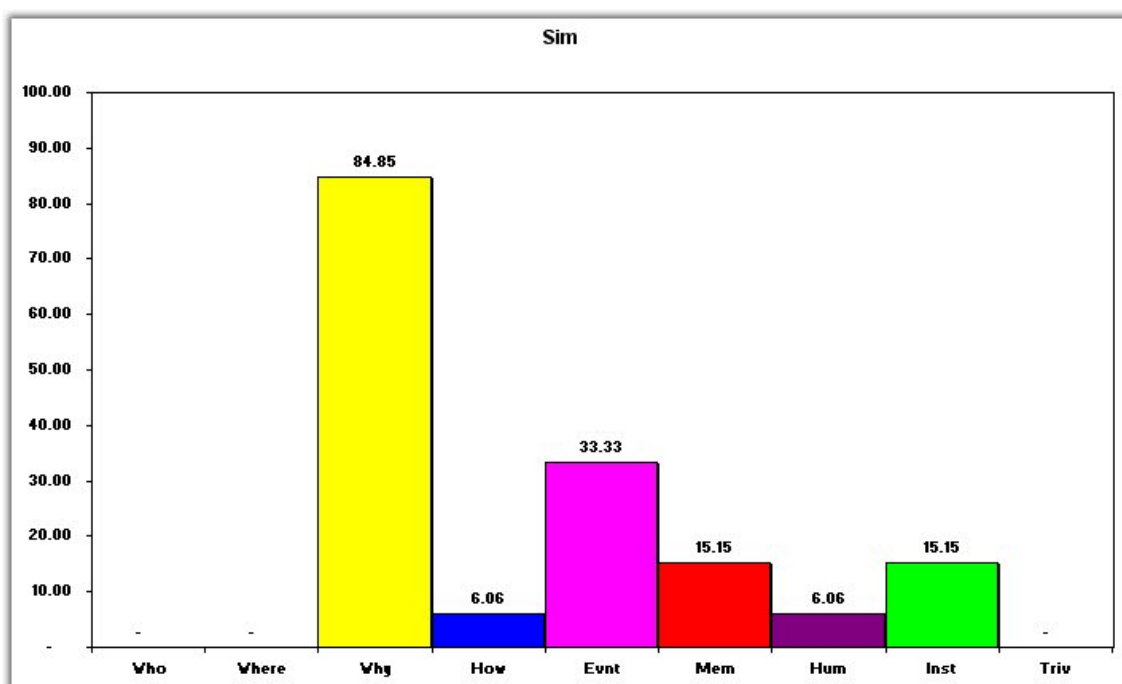
Figure 5.8. Meta Categories: Shneidman Simulateds

From Figure 5.8 it can be seen that the faked notes contain no instances of 'Who', 'Where', or 'Trivia'. Over 24% of the "Gens" contain 'Trivia'. More of the fakes state reasons ('Why') than the "Gens", but substantially fewer "Sims" than "Gens" cite 'Events'. Only 15% of the "Sims" have 'Instructions' compared with over 51% of the "Gens".

As was the case with the formal categories of Section 5.2, in terms of the number of differences between the sub-corpora of over 10 percentage points, it can be concluded that the "Hims" are more like the "Hers" than the "Gens"; but the "Gens" are no more

like the "Hims" than they are like the "Sims". However, there is a key difference. Taking all three genuine note corpora together (the "Hers", "Hims" and "Gens") there are notable omissions in the "Sims", particularly that of 'Trivia'. This does suggest that this particular aspect of topic is a good distinguisher, but it cannot be taken as a litmus test because 80% of the genuine notes do not contain any trivia. So the fact that a note does not contain any trivia does not mark it as fake; but the presence of trivia does strongly suggest it is authentic.

5.5. Practical Points

Given that certain percentages of notes have particular topics in them it might be interesting to see how many correct and incorrect identifications ('real' or 'fake') of notes would obtain by applying a test, such as taking some of these categories and interrogating the results. For instance, if one expects to find a 'Why' category in genuine notes, then 'Why' is being used as an indicator that a note is genuine; and under these circumstances the questions are: "In what percentage of cases would one correctly identify a genuine note?", and "In what percentage of cases would one correctly identify a fake note?"

Here are a couple of examples of how some of these findings could be used. Tables 5.10 and 5.11 show the results, in percentages (rounded to integers), for the categories 'Why' and 'Trivia' for the "Gens" and the "Sims".

Table 5.10. Simulated vs Genuine: Why

	Gen %	Sim %
presence	73	85
absence	27	15

Table 5.11. Simulated vs Genuine: Trivia

	Gen %	Sim %
presence	24	0
absence	76	100

If, say, the researcher were to use 'Why' as an indication of a genuine note, then in 73% of cases they would correctly identify a genuine note as genuine, but in 27% of cases they would incorrectly identify a genuine note as fake because it did not contain any 'Why' category text. Further, in 85% of cases they would identify a fake note as genuine, and in only 15% of cases would identify a fake note as fake. Looking at the table cells for the correct identifications we see that there is a reasonably high number of correct classifications of genuine notes, but an even higher number of incorrect classifications of fake notes.

Here is the second example. If the presence of 'Trivia' is deemed to indicate that a note is genuine, then 24% of genuine cases would be marked as such, but 76% would be marked as fake because they did not contain any trivia. Furthermore, none of the fake notes has trivia in it. So as a marker of something not being fake, 'Trivia' is quite good. If a note has trivia in it, one can be reasonably sure that it is genuine, but if it does not contain any trivia this is not a reliable indication that it is a fake.

'Trivia' gives many false negatives because it identifies three-quarters of genuine notes as fakes. The 'Why' category, on the other hand, gives far fewer false negatives, but it identifies too many fake notes as genuine (false positives). When trying to establish whether a note is authentic or not, these categories can at least be used as an indication of likelihood, but none of them is conclusive on its own. However, it may prove possible to attain much more reliable results by combining categories with each other.

5.6. A Fabricated Suicide Note

I will give here an example scenario as a staring point in order to determine whether it might be possible to identify the features of a fake note, and whether it is possible to determine authenticity by examining fake notes first, before any genuine notes.

An example legal case:

> "The suicide note was a doozey. It was handwritten, 15 pages long, quoted Mark Twain, lamented 'My battle is over now I must expire,' and repeated '[I'm] dead as a doornail' [**3] 28 times.

But surprise surprise. Unlike Marley's ghost, Feichtinger was not 'dead as a doornail' despite his repetitious assertions that he was. Feichtinger, it turned out, was alive and kicking in Arizona, where he was nabbed 10 months later in August of 1995." (Evans, Circuit Judge, U.S.A. v. Frank Feichtinger, 1997).

The above quotation is from the law report of U.S.A. v. Frank Feichtinger, 1997, United States Court of Appeals for the Seventh Circuit, No. 96-2375, 105 F.3d 1188 (accessed via LexisNexis, see Chapter 2). Unfortunately the full text was not in the appeal transcript. Given the judge's brief description of the text, in the first paragraph above, is it possible to tell whether the note was real or fake? And how could this be determined? Was it obviously fake only because a) its author was known to the police? or b) the appeal judges knew he was alive?

For the purpose of background information, here is the story of Feichtinger, taken from the same Appeal judgment.

"Frank Feichtinger has a common complaint: he thinks his sentence is too long...

Feichtinger and another man concocted a plan to break into and rob two banks... To carry out their plan they stole military equipment from an Illinois Army Reserve Center. Among the items stolen were night vision goggles, ... and - astonishingly - a two-and-one-half ton cargo truck. One vehicle apparently not being enough, they obtained another Army truck from the Jefferson Barracks in Missouri... In all, the military equipment they snatched was worth almost a half million dollars.

To understate, Feichtinger failed to appear for trial. What he did instead was to flee in dramatic fashion. He faked his own suicide by leaving his abandoned car and a suicide note on a Mississippi River bridge...

Upon his arrival back in Illinois, a second superseding indictment was returned, with a sixth count for failure to appear for trial added to the old charges.

This appeal is dismissed." (Evans, Circuit Judge, U.S.A. v. Frank Feichtinger, 1997).

Returning to the main point, how could one tell whether or not his note was faked? Most of the Birmingham corpus is handwritten, and it does contain some long texts,

although it does not contain too much in the way of quotations and repetition. So if handwriting and length are not clear determiners might the use of quotation and repetition indicate a fabricated text? This does not appear to be so because Shneidman's simulated notes seem to contain neither of these features. What strikes me as being odd about Feichtinger's text is use of the relatively infrequent lexis, "expire", the use of metaphor, in "dead as a doornail", together with this phrase being repeated many times, and the very light-hearted way of referring to death and not being alive any more.

In trying to approach this question, my inspiration came from The Birmingham University Forensic Linguistics Group, convened by Professor Malcolm Coulthard. Its seminars involved up to 30 participants (staff and post-graduate students), all examining the same text. Typically, after a while one of the participants said, "That's odd!" and there would be resounding agreement, subsequently qualified by a contribution along the lines of: "But where I come from people often say that". However, the first speaker would often maintain that the feature they had noticed was significant, until in the end the "oddness" was generally accepted by almost everyone. The studied texts were usually legally contested texts (often sub judice) of various types, and the common thread between most of them, in terms of what was considered to be 'suspicious' (meaning that a text was not what it purported to be) was usually that some aspect of it was "odd".

The method I actually used is a version of the above, derived from it. My method is also evidence-driven, and avoids having a priori categories. It begins with intuition and ends with corroboration, but instead of having other people to corroborate the findings, it uses corpora of other texts, namely my own suicide note corpus and the Bank of English. Having found a perceived "oddness" through manual observation of a text, I next seek a linguistic rationale for it by considering why it is odd, and then testing this potential reason (where appropriate) against a corpus. I then attempt to classify the oddness.

5.7. Oddness in Suicide Notes
As outlined above, the aim of this section is to see what strikes me as odd in the "Sims",

and to do this initially without reference to any other sub-corpus. I then look for oddness, in the same way, in other sub-corpora. Following this, I look at how these oddities can be classified in the light of categories derived from some other studies (and legal cases) that have found some types of oddness in equivocal texts. First, however, I shall present some examples of oddness from other forensic linguistic studies.

Gibbons (2003, p.303) mentions some of Coulthard's work in which Coulthard gives examples of statements fabricated by police. In these cases the interference of "police-speak" (Fox 1993) can be seen to mark the confessions as false. "Police-speak" refers to the register in which police tend to speak and write (often, according to Gibbons, to exercise their power over others or to exhibit what they see as formality), and it can manifest itself as over-specification, or as Gibbons (2003, p.85 & p.91) calls it "over-elaboration". Gibbons (2003, p.305) also refers to Coulthard's 1994 work on the Birmingham Six appeal which also featured over-specification in a statement which later turned out to have been "enhanced" by malicious police officers: the phrase "white plastic bag" was repeated in full several times. Indeed the forged statement also included apparent use by the suspect of his friends' surnames although his normal way of referencing them was by their first, or sometimes full, names (Gibbons, 2003, p.305). Regarding over-specificity, Coulthard (1994, p.417-419) mentions that specific details used inconsistently in a text, and details not relevant to the narrative, are suspicious.

Coulthard (1994) himself describes the Birmingham Six case, specifically the alleged confession of William Power, as concerning disagreement about how the confession statement was created and whether it was augmented by the police, who claimed that the entire statement was dictated by Power. Coulthard's evidence, therefore, included a difference between speech and writing, namely, the use of compound noun groups (not frequently found in speech) and their repetition. Although, as stated above, there was very little repetition in general in the Birmingham corpus, it seems reasonable to consider these findings of Coulthard's in relation to suicide notes because a fake note might have been dictated by a murderer (as the court, refusing to admit any linguistic evidence, decided was what had happened in the case of R v. Gilfoyle (2000)).

In addition to over-specificity, and the way people's names were referenced, Coulthard's (1994) evidence included mention of identical parts of texts which (particularly when occurring in more than one document) is highly unlikely to have occurred naturally and is thus highly suspicious (ibid., p.420-422). Coulthard (1994, p.417) also looked at suspicious documents in terms of coherence which he says is normally disrupted in personal texts, such as suspects' statements to police. He characterises coherent documents as being detailed, well-formed, well-organised, free from doubt, specific, and not violating the sequential time-line (to paraphrase Coulthard, 1994, p.417).

As well as overspecificity, underspecificity can be an indication of something being amiss with a text. Covering both, Coulthard and Johnson (2007, p.128) refer to pragmatics and the Gricean Maxims (Grice 1975), particularly the maxim of quantity. This refers to normal, non-suspicious, texts containing sufficient and only sufficient information, and that that information is based on what the speaker, or writer, thinks their addressee already knows (to paraphrase Coulthard and Johnson, 2007, p.128).

Spelling, specifically incorrect or inconsistent spelling of a family member's or friends' name, can also be an indicator that something is amiss with a text. Here is an example case from the LexisNexis legal database (mentioned in Chapter 2) in which the victim did not spell her best friend's name correctly.

> (5.10. Inconsistent Spelling - Odd)
> "Again, the best friend (who in the note is described as *Kelly*, but according to the child herself is called *Keeley*), denied ever having been told by Martina about any of the matters which are stated in the suicide note..." [My italics] (Re N (Minors), 1993, Court of Appeal (Civil Division) 27 October 1993 (Transcript: John Larking)).

Another writer who uses the concept of oddness, or inconsistency, in forensic work is Olsson (2004). He gives a useful set of indicators of dubiousness in texts (Olsson 2004, p.127-134). These are concerned with inconsistencies in the matters of time, place, narrative sequence and focus, superfluity, tense, character names, register, and pronoun omission.

Some of the features mentioned above are not relevant here, and others are not appropriate. There is no way to know whether names were mis-spelled, for instance, because there is no supporting information about the Shneidman note authors, nor are there any other example texts from the Shneidman note authors. There is likewise very little such information for the Birmingham texts. Ill-formedness and ill-organisation seem too nebulous to be easily assessed. However, combining the above with the findings of some of the previous studies of suicide notes (see Chapter 3) the candidate categories that seem most likely to be of use for distinguishing fake suicide notes (from genuine ones), in general, are complex noun groups, absence of detail, inconsistent detail, irrelevant detail, doubt, absence of specificity, inconsistency in tense, repetition, and character names.

To see whether these categories are indeed valid, I now embark upon the manual examination of the "Sims". However, rather than look directly for text that conforms to the categories, I want to look at the "Sims" with no particular categories in mind, and retrospectively see whether the categories mentioned above fit my findings. Of course, it is probable that awareness of the various potential categories influenced the types of oddness detected, but intuitively the exercise still seems worth doing. With respect to the following examples it is important to point out that oddities do not occur only in fake notes. I shall look at examples in real notes in due course.

So what exactly is the oddness that I am looking for? Towards a definition, at least a working definition, linguistic oddness is some feature of a text that seems to a native language speaker to be out of place, or unnatural, within its context. (The converse, that another such speaker would not find something odd, may or may not be enough to render something 'not odd'.) The context can probably be any linguistic aspect. Within this remit the following observations were made. (The examples given of text that is not odd, to contrast with the various examples of oddness, were obtained retrospectively.) The entire set of oddnesses in the "Sims", including those presented here, is given in Appendix A.

5.7.1. Naming References

Here is an example genuine text. Clearly, it cannot be certain that the "Joe" referred to in the text is the son of the note-writer and the addressee as opposed to, say, their friend, but it seems reasonable to assume that he is.

(5.11. Name - Not odd)
"Mary dear. Im sorry that I have been making you unhappy-I'm all twisted up inside. You and *Joe* will be better off this way-start over.
    Love, Bill" [Sic] [My italics] (Gen27b, Shneidman & Farberow, 1957e).

And here is an example simulated text.

(5.12. No name - Odd)
 "Dear, please forgive me for leaving you with all the responsibilities that this is certain to bring on you. If there is anything of me that can be used in any medical or scientific way please dont refuse to let them as my last request. I am very proud of *our son*, and his high potential in *his chosen field* for which he has real talent. Bye for the last time, and never forget that you were the best thing that ever happened to me. Have *my brother* help you, I know he will want to very much." [My italics] (Sim02b, Shneidman & Farberow, 1957e).

In the above simulated note, "our son" and "his chosen field" are odd. ("My brother" is also odd, see below.) The author's partner would probably know the name of her son, and his work. It seems awkward to reference them in this way. As mentioned in Section 5.2, this phenomenon (as exemplified by "our son" – for "his chosen field" see 'logic' below) was noticed by Arbeit and Blatt (1973, p.290) who found that, whereas the genuine notes tended to use names of people (and more of them), the simulated notes rarely addressed people or did so by their role. Arbeit and Blatt (1973, p.290) give examples of "Dear Wife" and "the children" which they found to be typical of the simulated notes, versus "Dear Mary" and "Tom and Betty" which they found to be typical of the genuine notes. There were nine notes in the "Sims" that contained such references to children. One possible caveat regarding these findings is that when actual names are used by the note-writers, as shown in the example genuine note above, one cannot always be sure of the role of the person being referred to. It might be that genuine note-writers tend not to write about their children, or tend not to have any. Intuitively, however, this seems unlikely.

There are issues of "Audience Design" in extract 5.12 above (Bell, 1984, p.197).[24] It seems that the simulated note writer got confused. When people write, or speak, they are constructing an imaginary audience, and the problem here is that the audience design is wrong. It is set up for a third party rather than the addressee. (The matter of audience design also concerns whether the note-writers envisage single or multiple addressees (in a single note). Although in Section 5.2 (above) I looked at how and whether addressees are addressed, time restrictions did not permit me to examine the linguistic cues involved in single versus multiple addressees. However, this is worthy of future research, and it might have important implications regarding the issue of distinguishing fake from real notes.)

Indeed, this type of oddness (in referring to people) is not restricted to simulated notes – and therefore is not diagnostic of them. In some cases it seems odd to refer to children by their roles even when their names are included. Here (extract 5.13), below, is an extract from a genuine suicide note that seems odd. Surely the addressee would know her daughter's name and that Betty was her daughter.

    (5.13. Role - Odd)
   "Please be good to little *Betty, our daughter*, I love her so." [My italics] (Gen07b).

Of the nine notes, containing a total of ten odd references to children, seven of the notes, possibly eight, were written to wives, and one was written to the note-writer's children. Thus the oddness seems to be largely dependent upon the role of the addressee. However, there are such instances of 'children' that do not seem odd, as in the following. It is not clear to whom the first example below was addressed, the second example below was, apparently, addressed to an ex-wife/ex-partner, and the third example below was addressed to the note-writer's current wife/partner.

---

[24] Bell (1984) seems to have coined the term "Audience Design" (ibid., p.147), and to have derived the idea of it, partly, from Goffman's 1981 book entitled "Forms of Talk" (ibid., p.159).

(5.14. Nickname – Not odd)
"Good by *Kid*. You couldn't help it. … Bill" [My italics] (Gen33b).

(5.15. Role – Not odd)
"I insist that she… attend my funeral, she has been very good to me, so please ask *the children* to be nice to her…" [My italics] (M97145b).

(5.16. Role – Not odd)
"YOU LACK THE BASIC UNDERSTANDING OF FEELINGS… LOVE, ACCEPTING THE PERSON FOR WHO HE IS… WITHOUT DAMMING HIM TO THIS HELL... ALL I PRAY FOR… IS THAT YOU LEARN FROM MY ACTION, SO *OUR CHILDREN* DO NOT SUFFER." [My italics] (M99438b).

The author of the second example, above, wrote many instructions, as far as one can tell, both to his current partner and to his ex-partner. Although his notes do not seem to exhibit any anger, they do seem to be absorbed with trying to persuade his partner and his ex-partner to co-operate in order to fulfil his wishes. It could be speculated that this absorption was similar in effect to the anger expressed by three other subjects who also wrote about their own children as "children" (without naming them). The third example above is an edited extract from one of the angry notes. Whatever the true reasons might have been, within the aforesaid contexts, these notes do not seem odd. One possible explanation for this is that address terms (such as "Kid" in the example above) are always different and permit terms that might otherwise seem peculiar. Another possible explanation, concerning the latter two examples above, is that the terms are plural, and plurals seem to normalise what might otherwise seem odd.

It is interesting to note that the non-naming oddness does not seem to extend to parents or siblings: "sisters", for example, does not necessarily strike one as peculiar, even when addressed to the family, as in the following example, 5.17. This is possibly because "sisters" is plural and is minimising the effort. Someone with a dozen sisters would certainly find it easier not to name each of them.

(5.17. Role – Not odd)
"All *my sisters* I love you all" [My italics] (M96488c).

144

I now return to things that are odd. Regarding Arbeit and Blatt's (1973, p.290) reference to the use of "Dear Wife" (see above), I decided to count this phenomenon as odd too. I included all instances of "Wife" occurring anywhere in the salutation/addressee pair. It transpired that there were seven instances of (that is, notes containing) this, but in each case "Wife" was not the only oddness in the note. In similar vein I also counted any instances of "Husband" occurring anywhere in the valediction/signature pair. There were two occurrences of this in the "Sims", and again, they were not the sole oddnesses in the notes in which they occurred. Of course, the Shneidman notes were all written by men, and mostly to women, so there were no "Husband"s in salutation/addressee pairs nor "Wife"s in valediction/signature pairs. However, to avoid gender bias (particularly when looking at the other sub-corpora) I counted all such references. Extract 5.18 below is an example of one of the wife/husband references.

(5.18. Role – Odd)
"Your ever loving *husband*" [My italics] (Sim28b).

As far as can be seen from the evidence addressing one's spouse or child by referring to their relationship role is odd, but addressing people by nicknames, or by their roles in the plural, is not odd.

5.7.2. Detail of Phraseology

Another category seemed to be odd uses of phraseology, specifically uses of nouns preceded by 'the…'. There were two cases of this (see extracts 5.19 and 5.20, below). In the first of these (5.19) the nominalised "**suffering**" seems odd.

(5.19. Phraseology – Odd)
"I cannot stand *the suffering* any longer". [My italics] (Sim08b).

There are no occurrences of "suffering" in the "Gens", but there are 3 in the "Hers" and 4 in the "Hims", and there is a second one in the "Sims", yet none of them is preceded by "the". (They are preceded by 'and', 'less' and 'or' (in the "Hers"), by 'be', 'is', 'I'm' and 'are' (in the "Hims"), and by 'my' (in the "Sims").) However, it appears that it is not the simple presence of the definite article before "suffering" which is the cause

of the oddness. In the Bank of English "the" is the most frequent collocate in the N-1 position (i.e. 'first word to the left' of (the node word "suffering".[25] This was evident in 11% of the concordance lines. Further, the use of "cannot" in itself does not seem to be problematic. While "I can't" occurs 35,070 times in the BoE, and "I can not" only 89 times, the phrase "I cannot" appears 6,532 times, which is far from trivial. In any case, being infrequent, or even unique, does not necessarily mean that something is odd; being odd is diverging from a pattern. There is no evidence of such divergence here.

If one searches the BoE for occurrences of "bear", "stand" or "take" followed by one other word, followed by the word "suffering"[26], in all the available sub-corpora (that is 450 million words), one finds only 8 lines, all unique, containing the pattern. None of them corresponds to the above extract. Only one line seems to refer to the suffering of its author. The line contains the text, "I have to take my suffering as it comes", and that seems to be a quotation from another source. The remaining lines refer to the suffering of others. Of the latter, two lines show that the author is also undergoing some suffering by proxy: "I can't bear animals suffering" and "I can't bear anyone suffering"; the first of these lines again apparently a quotation. It is worth noting that in both these cases, "suffering" is a non-finite verb, in effect creating a subordinate clause, rather than a noun.

In other words, there are no examples in the BoE of anyone who can, or cannot, bear, stand or take "the suffering" themselves directly. The closest match is someone who seems to be saying that they can endure their ("my" rather than "the") suffering. (An examination of 100 concordance lines of "suffering" used as a noun in the Bank of English[27] shows that the verbs that usually precede "suffering", either one or two places to the left of it, are typically words concerned with relieving it or revealing it, although there is no really clear pattern.)

---

[25] This was ascertained by taking 100 random (as calculated by "LookUp", the software used for interrogation of the Bank of English database) lines from the available 24,548 concordance lines, across all sub-corpora, featuring the word "suffering".

[26] The query syntax is actually "bear|stand|take+1,1suffering".

[27] The query syntax is "suffering/NOUN".

This examination of some of the collocates of "suffering" in the BoE supports my intuition that 5.19 is odd. The oddness seems to lie in the fact that "the suffering" is not post-modified in any way by a phrase or clause, such as "the suffering of the children" or "the suffering that you caused me", as well as in the combination of the personal pronoun "I" with the impersonal determiner "the". It might also be relevant that in the BoE the word "suffering" used as a verb occurs 20,647 times,[28] but as a noun it only occurs 3,901 times. (It should be mentioned, however, that the part of speech tagger in the BoE is not 100% accurate, and has by my calculation an error rate of 19% for verbs, and 9% for nouns.[29]) Notwithstanding a degree of subjectivity in judging something as 'odd', I believe that intuitions of 'oddness' can often be substantiated by comparison with a reference corpus, and the concept is frequently a good distinguisher of fabricated suicide notes.

My second example of odd uses of phraseology is shown below (extract 5.20). Here the nominalised "**support**" (as opposed to "supporting") together with the impersonal "the **family**" (rather than "my family") seems odd.

(5.20. Phraseology – Odd)
"I no longer feel that I can be of help in *the Support* of *the family*". [My italics] (Sim30a).

There are no instances of "support" as either a noun or a verb in the "Gens", but there are 3 uses of "support" in the "Hers" and 7 in the "Hims". Only one instance takes the form "the support" and this is in the "Hims" (see below). Yet this does not seem odd: it does not have the second "the" (before "family") and "the support" is preceded by "with" rather than "in" which seems to help in rendering it unmarked for oddness.

---

[28] The query syntax is "suffering/VERB".

[29] Examining 100 concordance lines from the Bank of English for "suffering" tagged as a verb revealed 19 lines where it was actually used as a noun or an adjective. Similarly, examining 100 concordance lines of "suffering" tagged as a noun showed that in 9 lines it was actually used as a verb or an adjective.

(5.21. Phraseology – Not odd)
 "I know with *the support of* family and friends you can cope…" [My italics] (Him: M98080a).

Indeed the whole of the BoE yields only 6 occurrences of "the support of the family". These are shown in Figure 5.9. In none of these instances is the author talking/writing about their own family. Two uses (lines 4 and 6) are in fact referring to compound nouns of which "family" is only a part ("Family Planning Association" and "Family Protection Unit". Only one or two instances appear to refer to a specific family: line 3, from the British book sub-corpus, and therefore probably fictional, and possibly also line 5. Lines 1 and 2 (and maybe 5) are referring to families in general, albeit in the contexts of Asian students, politicians, and what seems to be a court case, respectively.

Searching the BoE for "the + [any single word] + of the family" results in 1,854 lines, most of which either seem to be referring to 'the rest of the family' or are assorted versions of "the death" and "the virtues" of the family. This does not seem to further the case for oddness. However, searching in the BoE for "the support of the" yields 1,577 lines, many of which seem to refer to "the support of" large institutions or organisations such as "the General Assembly". A sample of these concordance lines is given in Figure 5.10. This may partly account for the oddness of the phrase when "family" is used in it – referring to one's own family as if it were an impersonal institution. One can envisage an exception to the perception of "oddness" if the family really were a large organisation: if the writer of the suicide note in question had been a member of the Mafia it would arguably not seem odd.

However, there is another clue in Figures 5.9 and 5.10 to the nature of the oddness in extract 5.20. In almost every case, the noun following "support of" is in a subjective relationship to the "support"; in other words, the person/institution represented by that noun is supporting (or failing to support) the person/institution represented by the noun phrase preceding "support". Strikingly, in 5.20 it is the other way around: "the family" is in an objective relationship to "support", i.e. the writer apparently imagines himself to be expected to support "the family" rather than "the family" supporting him. Further, the verb is very often, or is synonymous with, 'have', 'win' or 'lose'.

```
1. student knows that he or she has the support of the family and the support of
2. simply could not survive without the support of the family." Deputy Quinn, 52,
3.  during the father's illness for the support of the family. None of the
4.           of Femidom that has won the support of the Family Planing Association.
5.   the best hope the child has is the support of the family. We're not blaming
6. they're not so much at risk with the support of the Family Protection Unit. And
```

Figure 5.9. The Support of the Family

```
    be `American" cut her off from the support of the women of her family,
 candidate probably couldn't win the support of the General Assembly, where
 it a program that will sustain the support of the American people? <p>
        The Toon boss has now lost the support of the Board, players and fans in
 to an end. Lloyd George had won the support of the chapels when he had
 a given programme, but to enlist the support of the best minds in formulating a
 and produce a hard spot beyond the support of the terminal sleeve, increasing
 respected qcs, has already won the support of the Democratic Unionists in the
 visited another project run with the support of the Queen Mother of Mafeking.
        but it was pleased to have the support of the powerful Jewish lobby in
 as elected representatives have the support of the people. In the new
 new leader who could govern with the support of the opposition Congress I
        In a blatant attempt to win the support of the people it was also
 appear to be in danger of losing the support of the masses; while communalist
    together ... to make it work." The support of the Australian players for
    BRISBANE players have sought the support of the heavyweight AFL Players
  s radon program, mounted with the support of the Ad Council, was frightening
   confidential and is done with the support of the subject. He concedes,
    intensive. If it weren't for the support of the public and for projects
 than half the seats and needed the support of the Lib Dems to form a majority
```

Figure 5.10. The Support of the

To test this, the BoE was searched for the term "the support of", and a hundred random concordance lines were taken from the resulting 5,001 hits. Fairly conclusively, out of 100 instances of 'X … the support of Y', 93 instances were Y supporting X (subjective genitive). Only 7% of the instances were cases in which X supported Y (objective genitive). Therefore, the oddness may not be related to "family" or any institution or individual, but to the relatively rare use of the objective rather than subjective genitive in the phrase "the support of".

5.7.3. Logic

One group of examples seemed best categorised as 'logic', or more accurately, 'illogic'. This is not logic in the formal sense, but rather that of common sense. There were nine "Sim" notes containing altogether 12 of these instances. Because they are not all as obvious as some of the other instances of oddness described above I have cited a number of them.

The extract below is similar in its oddness to the text (5.12 above) containing "his chosen field". Both texts are illogical in that they are over-specific given their intended addressees. In 5.22 the note was purportedly written to the author's mother who would probably know what the "job" was, how long her son had held it, and what his wife's name was.

> (5.22. Logic – Odd)
> "Friday I lost *the job I have held for the past seven years.* When I told *my wife…*" [My italics] (Sim01b).

Sometimes the surrounding textual context plays a role in rendering the phrase or clause "odd". Here are some examples that are also (like the above) illogical in the context of their addressees, but they could not be described as being over-specific. In the note below the note-writer is asking the police to inform his family. There is no apparent accompanying address or telephone number for the family, and it seems odd that he did not write a note to them himself.

(5.23. Logic – Odd)
"*To the police*. please *tell family* that I love them why say more." [My italics]
(Sim33a).

Example 5.24, whose probable intended meaning is quite clear (i.e. 're-marry after I'm dead'), nevertheless contains the salient juxtaposition of a note addressed to a "wife" with the message of the hope she will marry. (Had the text been 'I hope you will get married again' it would not seem odd.)

(5.24. Logic – Odd)
"Dearest *Wife*:… I hope you will be *married* as soon as it is proper to be,…"
[My italics] (Sim22a).

5.25 is from a note in which the author does not explain his reasons despite claiming that this is the point of his text. This is a context-dependent illogicality, but it is not an addressee-related one.

(5.25. Logic – Odd)
"*I am writing this to explain why* I am going to end it all. I know that this is a cowardly way and I am sorry but I just haven't the will to do otherwise." [My italics] (Sim19a).

Some notes are illogical simply in that they are pragmatically non sensical, even from a lay point of view, considering their pretentions to be suicide notes. In the note below, it seems illogical that the author bothered writing a note at all. He may not want to, or be able to, express his reasons for the suicide, but to say that he does not know is odd.

(5.26. Logic – Odd)
"Dear Mary. *I don't know why I am doing this* unless my reasoning has gone all to pot. Something must have slipped.
Bill" [My italics] (Sim07a).

Other instances of logical oddness included a note saying the author was being unreasonable in killing himself, a note expressing possible doubt about committing suicide, and notes expressing hope that no one would be distressed, or "embarrassed", by the suicide. "Embarrassed" certainly seems to be an odd choice of vocabulary: it is not used in any of the other notes.

Most of these oddnesses cannot be checked against any corpus. They are, to varying degrees, subjective. This makes these types of oddness more difficult to 'prove', yet they remain striking.

5.7.4. Vagueness & Melodrama

Another group of examples, of which there were five, consisted of non-specific items in the texts. I have called this category, 'vague'. It was not that the lexis itself - words used included "things" and "everything" - meant that I considered it vague, but that there was no accompanying amplification of it, nor detail with it. As an example, the note below does not give any indication about what the writer is tired of. This is not an extract, but is the entire note: I consider this to be vague precisely for that reason.

(5.27. Vague – Odd)
 "I'm tired of *it all*. I Love you and God Bless you." [My italics] (Sim32a).

There were five notes containing instances of what could be described as melodramatic prose. Arbeit and Blatt (1973, p.290) mention finding more "embellishments" in fake notes including what they call "high-flown" ones that seem to fit the concept of melodrama.[30] The following extract is an example of some melodramatic text I found in the "Sims".

(5.28. Melodrama – Odd)
 "As I sit here *with this gun in my Hand*, which in a few minutes I will take my life I am thinking of all the wonderful *minutes, Days, years*, I have spent with you." [My italics] (Sim31b).

As with the oddnesses of illogicality, those of vagueness and melodrama can appear more, or less, obvious, and vary in their subjectivity, therefore, making them difficult to substantiate. Nevertheless, they are worth noting.

---

[30] Arbeit and Blatt (1973) consider the phrase, "End my suffering forever" (Sim24a) to be melodramatic, although my reaction to the phrase is that it is not, and I did not count it as "odd" in my results.

5.7.5. Miscellaneous

Following these observations in the "Sims", the texts in the "Gens", "Hers" and "Hims" were also examined for oddness. This involved both keeping an eye out for any further types of oddness, and looking for any similar oddnesses to those already discovered in the "Sims". To ensure gender inclusivity, the 'wife in salutation/addressee' category was naturally extended to cover "wife" or "husband" in salutation/addressee or valediction/signature. All the examples of oddness in the real suicide notes are given in Appendix B.

One additional type of oddness was found in the "Hims": repetition (in two texts). Interestingly, there was also a rare instance of odd usage involving a pronoun and that person's name (in one text). The latter involved the note-writer referring to himself in the third person. Examples are shown below.

> (5.29. Repetition – Odd)
> "*good bye good bye good bye good bye* good b
> L=20   CH=1   J=L NB" [My italics] ([Computer print out, repeated 3 times] Him, M98542).

> (5.30. Name & Pronoun – Odd)
> "Mary has made sure *Bill* can not move in to house, has not revealed all of *his* savings and *his* share's…" [My italics] (Him, M95468g).

Although oddness can exist in real suicide notes, the finding that far more of it seems to occur in simulated suicide notes, and that some of one's intuitions of "oddness" can be corroborated, signals that a good degree of confidence in my results is justified. The results themselves will now be summarised.

5.7.6. Summary of Results

The instances of oddness were counted in two ways. Firstly, oddness was counted on a 'number of notes' basis, that is, counting a maximum of '1' for every text that had any oddness in it, regardless of how many types or instances of oddness it had. Secondly (see below), the total number of occurrences of oddness were counted in order to examine the density of oddness.

Figure 5.11, below, shows the number of notes displaying oddness over the four sub-corpora. This looks very promising: there is a clear difference between the simulated texts and the others.
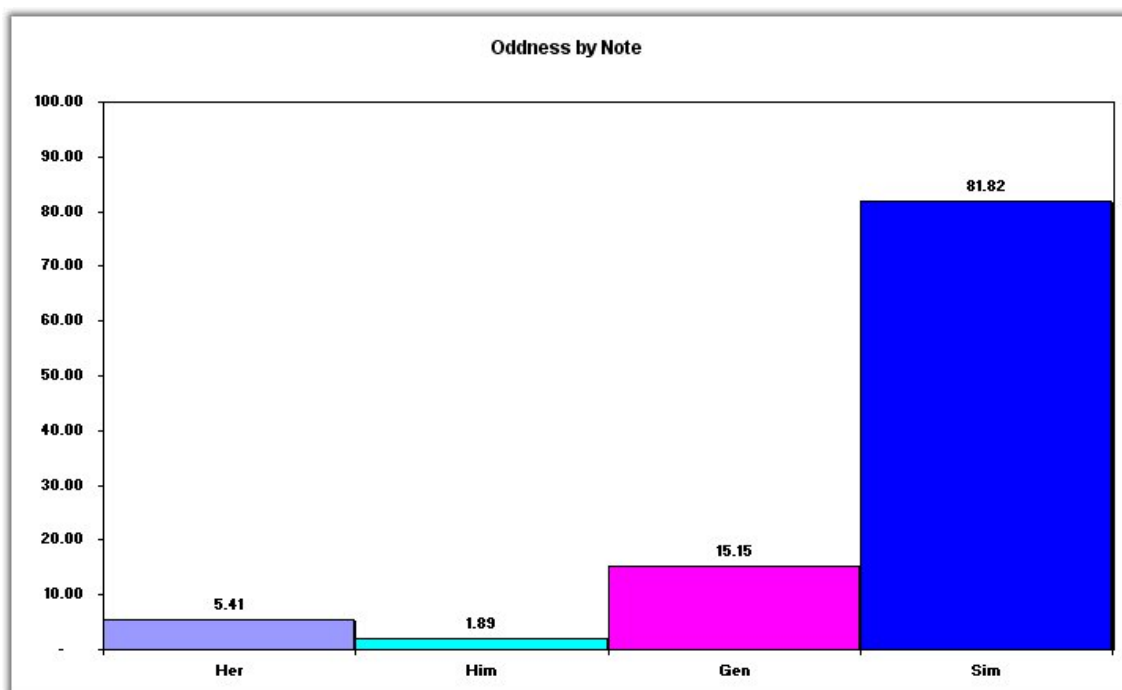
Figure 5.11. Number of Notes showing Oddness, by Sub-corpus

The categories found in the "Sims" were 'naming references' (referring to children and spouses), 'detail of phraseology', 'logic', 'vagueness' and 'melodrama'; whereas the oddnesses found in the real notes, particularly the "Hims", include signatures, repetition and pronoun use as well as the categories found in the Sims.

Looking solely at Shneidman's simulated notes, the odd items seem to be mainly of one type. It seems that in simulated suicide notes, at least in Shneidman's corpus, oddness is not necessarily related to writing technology, length of text, quotations, repetition, ambiguity, or rambling. It does seem mainly to involve the lack of prior knowledge of the addressee, i.e. over-specification, particularly in terms of other people's relationships to the note-writers, but under-specification in terms of identifying the

individuals by name. Yet, with the possible exception of addressing, or signing, a note with the words "Wife" or "Husband", it seems that this tends to occur only when referring to children. Other major categories of oddness seem to be centred on lack of logic, and lack of detail (what I have called vagueness) within the texts. Another category seems best described as detail of phraseology, featuring constructions involving the objective genitive and other marked uses of noun phrases. There are also some instances of melodrama/hyperbole. Oddness is usually context dependent, and this can be in various ways, but sometimes it is phraseological, being merely something that people rarely, if ever, say.

Figure 5.12. Density of Oddness:
"Hers","Hims","Gens","Sims"

Figure 5.12 shows all the oddnesses found, and depicts the extent to which they were spread between the notes. Although the bottom left of the figure has super-imposed some of its icons on top of each other - for example, the first three sets (of 'one oddness per note') each represent all four sub-corpora – it can be seen that most of the notes

155

containing oddness contain only a single instance of it. This is true for the fabricated notes as well as the real ones. Indeed, of the 41 oddnesses in the "Sims" (found in 27 notes), 16 were the only example of oddness in a note, leaving 11 notes containing more than a single oddness. So most of the "Gens", "Sims" and "Hims", and all of the "Hers", had no more than one oddness per note. To clarify the details in Figure 5.12 above, I include Table 5.12, below, which gives the figures concerning the densities. It shows, for example, that one of the "Gen" notes contained four oddnesses, a number not equalled by any of the other sub-corpora. The issue of density of oddness shows that sometimes it can take a single instance of oddness to point to a note being fabricated.

Table 5.12. Oddness Density Table

Instances of Oddness	Her	Him	Gen	Sim
1	4	3	3	16
2		1	1	8
3				3
4			1	
Total	4	5	9	41

The categories in this chapter are not all the same as the ones found in other forensic linguistic cases or studies (see above), although there is clearly some overlap or similarity with some of them. Given that Feichtinger's fake note exists (see above), it would not make sense to postulate that repetition, or any particular type of pronoun use, were not oddities. Only further research with more fake texts might resolve this.

5.8. Conclusion

In this chapter I first tried to characterise the anatomy of a suicide text by examining some genuine suicide notes and comparing them to some simulated ones. In general, it seems that there are features that are more likely to be found in a typical genuine suicide note than in a fabricated note.

Of course, there are likely to be differences between notes written in the 1990s and notes written in the 1940s, and also between notes written in Birmingham, U.K. and

notes written in Los Angeles, U.S.A. Despite these provisos, it seems that real suicide notes may contain postscripts, dates, and addressees, contain more explicit expressions of 'love', 'fault' and 'goodbye' than simulated notes, contain 'who I am' and/or 'where I am' statements, and trivia, mentions of memories and events, and more instructions than simulated notes. Roughly half of them are less than 75 words long (see Chapter 4)). Simulated (or fabricated) suicide notes, on the other hand, seem to contain few instructions, no trivia and no postscripts. All this very broadly concurs with what other researchers have found.

Thus far it is clearly not the case that all real suicide notes, or all fake suicide notes, contain certain properties or have certain characteristics that can invariably distinguish between them. However, it may be that a prototypical real or fake note can be discerned. (Indeed, a couple of prototype (real) notes are given in Chapter 7, Section 7.2, below.) Yet it seems that there will always be some notes whose authenticity cannot be proven or disproven, and typical real note characteristics may be found in fake notes. This suggests, as others have found (see Chapter 3), that the distinction between genuine and simulated suicide notes cannot readily be generalised.

In this chapter I also looked at "oddness" in suicide notes. I began with pure intuition before taking any linguistic approach. In the second part of this chapter (unlike in the first part) I employed the tactic of examining some simulated suicide notes first, and then comparing them to some genuine ones. In general, it seems that oddness tends to be present in fabricated notes, but tends not to be in genuine ones.

Of course, as has been said before, it should be borne in mind that there may be possible differences between notes written in the 1990s and notes written in the 1940s, and between notes written in Birmingham, U.K. and notes written in Los Angeles, U.S.A. Ignoring this, however, one can say that fabricated suicide notes are often 'odd'. But, as has been said before, and must not be forgotten, Shneidman's Simulated notes are not the same as the type of texts that are usually the objects of legal dispute, such as murderers' faked suicide notes. Despite this, any advancement toward being able to define a prototypical suicide note, and to distinguish a fake from a real note, has to be

worthwhile. Interestingly, Leenaars (1988, p.210) (see Chapter 3) gives a prototypical (genuine) suicide note, but in the light of this chapter one might question its prototypicality because it includes 'odd' features such as the words "my brother" although it is addressed to the author's wife, and it has the valediction "Your husband". Here is Leenaars' note.

> Dear Mary,
> I was really in love with you. Remember the Rockies. I guess it was all my fault. I'm no good. I caused all the trouble. I'm sorry for losing my job.
> I can't go on without you. I feel so alone. It would have been easier if you had listened to me. I thought we were happy for all those years.
> I walked outside last night. No one talked to me. You can have the house back. And the girls they will get the insurance money. There is $500,000. Call Mr. Taylor at 000-0000. And my brother always wanted the car.
> I don't want the girls to worry. You said that I always hurt them. They'll be better without me. Tell them that I love them and died in an accident. You don't have to tell my family. It is not important.
> All I have is bad dreams. I don't sleep. I'm so mixed up honey. I can't handle it. Forgive me for I know what I do. This is the only way as long as you feel as you do. Now you can be free. It is just too hard to live.
> Please no service. I'm not worth it.
>   Your husband forever,
>   Bill
>
>   (Leenaars, 1988, p.210.)

Identifying "oddness" is by no means a guaranteed way of distinguishing real from fabricated notes. It is not the case that fabricated notes have oddnesses and real notes do not. Some of the real notes have them, while not all the fabricated notes do. Further, the oddnesses themselves are, to a certain extent, subjective (see below). Thus, although this is a useful pointer, it is not a definitive way of distinguishing between fabricated and real suicide notes. Table 5.13 shows the percentages of odd, and non-odd, Shneidman simulated and genuine notes. This is a useful diagnostic device. Given this matrix, anyone examining suicide notes might find that they had falsely identified 15% of real notes as fakes, and 18% of fake notes as real.

Table 5.13. Simulated vs Genuine Notes

	Gen %	Sim %
+ oddness	15	82
- oddness	85	18

There is a greater and lesser degree of subjectivity in identifying oddness. Oddness is clearer when it involves the naming of family members, and when it involves details of phraseology, which can be checked against a corpus. Oddness is more difficult to prove when it involves logic, vagueness or melodrama. Some things are odd because there are other things that are odd in the environment, such as the subject's personal history or the textual context of other suicide notes they have written. So detecting whether a note is fake or real in these circumstances might require several instances of oddness, not just one. Taking all the various corroborating evidence together, however, I am fairly confident in my results.

Oddness is, on the one hand, one of the more difficult things to identify; but on the other hand, it is one of the most useful things to find. My suggestion is that it might be possible to identify a fake note – although not a genuine note – from its degree of oddness, and that this is useful, even if the identification can never be 100% agreed upon.

6. WHAT ARE SUICIDE NOTES ABOUT?: LEXIS

6.1. Introduction

In this chapter I want to concentrate on the Birmingham corpus and begin to answer the question "What are suicide notes about?", and I want to do this by approaching the general issue of topic in suicide notes in a different way than I do in Chapter 5 by applying some corpus linguistics methodology to the matter. According to Wellisch (1994, p.131) "The term 'aboutness' was coined by Fairthorne…", in the context of library science which was Fairthorne's field of expertise; and the frequently cited reference for the term is Fairthorne (1969). Studies of 'aboutness' are common in library science, where cataloguing various documents and their subsequent retrieval is the primary issue; but there is also a linguistic tradition going back at least as far as Van Dijk (1977, cited by Phillips, 1989) who developed the concept, which was continued by Rayson (2007) and Scott (2007a), among others. However, the 'aboutness' in which I am interested concerns mainly the topics that the note-writers choose. The approach to "topic" here is more quantitative than the more qualitative one applied in Chapter 5. The intuitively obvious relationship between topics and the words used to write about them suggests that an examination of lexis should reveal some frequently-occurring topics. One of the functions of most linguistic computer software is producing word lists, so these should assist the search. Therefore, in order to ascertain what suicide notes are 'about' I began by generating word frequency lists; I then turned to key word lists.

As discussed in Chapter 4 (Section 4.11) I used Wordsmith Tools version 4 (Scott, 2007a) for the basic word and key word lists, and my reference corpus was the British National Corpus (BNC) (see Section 4.7). Additionally, I used 'grep' (Microsoft Corporation, 1987) and 'Readgrep' (which used token counts from Wordsmith Tools version 3) (see Section 4.11). It is worth reminding the reader that the different software applications, or even different versions of the same application, may generate different figures (for example, token counts) from the same input text (see Section 4.11).

Although I did not intend to include gender as an independent variable in my study, I decided to keep the Birmingham "Him" and "Her" sub-corpora separate, as before, to see whether the males and the females tended to write about the same things, or in case any gender differences became apparent in the way they wrote about them.

6.2. Word Frequencies

Using the default Wordsmith Tools settings for this work (see Section 4.11) I produced various types of word lists, which I will describe in this Section. The columns in Wordsmith 4's word lists (going from left to right) show the word, its frequency in the corpus, the latter as a percentage of all the words in the corpus, the number of texts in which the word occurs[31], and the latter as a percentage of the number of texts in the corpus. Thus, in Table 6.1.a below, the first result line shows that in the "Hers", "I" occurred 301 times which amounted to 3.99% of all the tokens in the sub-corpus. It occurred in 61 of the texts in the sub-corpus, and those texts made up 82.43% of all the texts in the sub-corpus. Being able to see what percentage of a corpus as a whole is affected by a word, and also what percentage of a corpus' composite texts is affected by the word might be very useful in beginning to determine the aboutness of a corpus.

Tables 6.1.a and 6.1.b show the most frequent words in the "Hers" and the "Hims" respectively. The '#' shows that a corpus (or sub-corpus) contains numbers in digital form (see Section 4.11). The "Hers" contained 168 such numbers, and the "Hims" contained 481. The 'X' on line 22 in Table 6.1.a indicates the 'X's representing kisses found in some notes, usually at the end. In this case it can be seen that 11 of the "Her" texts contained kisses. (What cannot be seen in the table is that these are produced by only two subjects. One of them wrote 10 notes; the other was responsible for 21 of the occurrences of 'X'.)

---

[31] Leech, Rayson and Wilson's (2001) 'range' is a similar concept, but it is not the same because Leech et al. use chunks of texts of equal size rather than individual texts.

Table 6.1.a. Most Frequent Words in "Hers"

N	Word	Freq.	%	Texts	%
1	I	301	3.99	61	82.43
2	YOU	232	3.08	51	68.92
3	TO	195	2.59	53	71.62
4	THE	193	2.56	55	74.32
5	AND	191	2.53	43	58.11
6	#	168	2.23	25	33.78
7	MY	130	1.72	51	68.92
8	ME	104	1.38	47	63.51
9	FOR	102	1.35	50	67.57
10	HAVE	100	1.33	45	60.81
11	OF	100	1.33	35	47.30
12	A	97	1.29	34	45.95
13	IN	87	1.15	44	59.46
14	BE	71	0.94	38	51.35
15	IT	71	0.94	28	37.84
16	NOT	69	0.91	36	48.65
17	IS	69	0.91	32	43.24
18	ALL	68	0.90	36	48.65
19	WITH	62	0.82	30	40.54
20	THAT	56	0.74	25	33.78
21	BUT	53	0.70	28	37.84
22	X	53	0.70	11	14.86
23	WILL	50	0.66	34	45.95
24	DO	45	0.60	29	39.19
25	LOVE	45	0.60	28	37.84

The next most notable feature of Tables 6.1.a and 6.1.b is that they show that the bulk of the top 25 words, in both sub-corpora, are not lexical words, but grammatical, being mostly a mixture of pronouns, determiners, conjunctions and prepositions. This is not surprising and would be expected in the top 25 of any word list based purely on frequency of occurrence and nothing else (Hunston, 2000). According to Leech, Rayson and Wilson (2001), for example, the top 24 most frequent words in the BNC are: "the", "of", "and", "a", "in", "to", "it", "is", "to", "was", "I", "for", "that", "you", "he", "be", "with", "on", "by", "at", "have", "are", "not" and "this"; and the 25th most frequent occurrence is words having the genitive "'s". Although grammatical words do not have any meaning themselves they are associated with meaning and so it might be

worth looking at some of them to see whether they aid the quest for aboutness. However, before considering this (see Section 6.2.1, below), I want to deal with the words that do carry lexical meaning.

Table 6.1.b. Most Frequent Words in "Hims"

N	Word	Freq.	%	Texts	%
1	I	1,580	5.18	176	83.02
2	TO	1,023	3.35	153	72.17
3	YOU	984	3.22	147	69.34
4	THE	912	2.99	143	67.45
5	AND	657	2.15	127	59.91
6	#	481	1.58	87	41.04
7	MY	445	1.46	126	59.43
8	HAVE	444	1.45	116	54.72
9	OF	434	1.42	116	54.72
10	ME	424	1.39	127	59.91
11	A	403	1.32	114	53.77
12	FOR	395	1.29	132	62.26
13	IN	325	1.06	108	50.94
14	IT	313	1.03	100	47.17
15	THAT	306	1.00	85	40.09
16	IS	301	0.99	102	48.11
17	NOT	284	0.93	103	48.58
18	ALL	277	0.91	112	52.83
19	BUT	261	0.85	103	48.58
20	LOVE	261	0.85	119	56.13
21	THIS	235	0.77	95	44.81
22	WILL	215	0.70	97	45.75
23	WITH	210	0.69	86	40.57
24	BE	196	0.64	90	42.45
25	SO	181	0.59	77	36.32

Both the "Hims" and "Hers" feature the word 'love' in their top 25 most frequent words. This is not surprising either as high frequencies of 'love' were found in previous studies of suicide notes (see Chapter 3). Apart from 'love', the other word most likely to have had some use in a lexical capacity in this context of suicide notes is 'will' referring to 'last will and testament'. However, a check on this (via the 'grep'

command) showed there to be very few such uses of the word. In the "Hims" there was only a single instance of 'last will', but there were two instances each of 'the will', 'a will' and 'my will', making a total of seven uses of 'will' as a noun by the male authors. In the "Hers" there were only three such references and all were of the form 'my will'. The overall low frequency of references to wills seems surprising. Considering all the instructions left in suicide notes (about 50% of notes contain them, see Section 5.4), it could be said that some of these form the basis of a last will, and one would imagine that a few more of the note-writers would use the word 'will' in that sense. Even if some of the note-writers had made wills as documents separate from their suicide notes, one might expect some reference to the former in the latter (such as 'my will is in the top drawer').

There is a fairly similar order of frequency of the words in the "Hers" and the "Hims". Thus, so far, the sub-corpora are, more or less, similar to each other, but the word lists give few clues as to what they are about – other than, perhaps, the authors themselves, the addressees, and 'love'. (Regarding the former two, it is notable that in the top 25 words, 'I' and 'you' are much more frequent, and higher in the wordlist, than they are in either the BNC (see Tables 6.4.a and 6.4.b in Section 6.3 below) or the BoE.[32]) Looking at the next 25 words might show more lexical words, but this may not after all be too useful in the quest for aboutness. One reason for this is that it is difficult to say what it might mean if a word makes up a particular percentage of a corpus. For example, one might find that 'life' accounts for 0.2% of the "Hims", but the remaining 99.8% of the tokens in the "Hims" might, or might not, be synonymous with the word 'life' in the context of the note in question. (This indicates that some type of semantic grouping of words could be profitable (see below).) Another reason is that a very frequent word in a corpus might be so frequent because it is common in the language in general: as such it would not be indicative of any particular genre. (This indicates that some type of standardised corpus should be used for comparison (see below).) Yet another reason is that a word might be used frequently by only a few writers (see Section 6.2.1. below). So rather than proceeding with looking at more words from the simple frequency lists, a more useful word list might be one that is sorted in order of the

---

[32] In the BoE, brbooks, 'you' occurs in 0.72% of the corpus, and 'I' in 1.03%.

number of texts containing a word, that is a list 'by (suicide) note'.

Before proceeding, however, as an aside, one question that a word frequency list poses is whether the corpus from which it is generated has more in common with typical written texts or with typical spoken texts. Leech, Rayson and Wilson (2001) provide a list showing these written-spoken differences as they are exhibited by the BNC: their "Distinctiveness list contrasting speech and writing". Looking at only the first page of Leech et al.'s list reveals the following.

According to Leech et al., in speech (as opposed to writing), along with more hesitations such as "er" and "erm" which emanate from the on-line processing of human speech, contracted forms such as "…n't", slang such as "yeah", informal words such as "got", and responses such as "yes" and "well", there are more frequent occurrences of the pronouns "you", "I", "it", "we" and "they". Speech also features more occurrences of the words "that", "know", "think", "there", "like", "want" and "then" (to mention just a few). Further, phrases such as "I think" and "I know" are more frequent in speech. Leech et al.'s list also shows that in writing there are more occurrences of the words "the", "of", "by", "in", "from", "an" and the pronouns "his", "its" and "her". The long noun phrases that are more frequent in writing typically include the words "the" and "of", for example.

Of the top most frequent words in the "Hers" and "Hims" about half a dozen in each sub-corpus ("I", "you", "have", "it", "that" and "do" ("Hers"), and "so" ("Hims")) are typical of speech, and three ("the", "of" and "in") are typical of writing. In both sub-corpora, of the top 25 word types, it seems that the ratio is about 2:1 in favour of speech. In other words, the Birmingham suicide notes are more like speech than writing in their lexis. Of course, the texts are written and so there are no hesitation phenomena present, but the other types of features characteristic of speech all seem to be represented in the word lists. An analysis of the notes shows that most of the notes are informal and are written to friends, partners and family. Therefore it seems safe to presume that the notes are more like speech because of the informality assumed by the authors in such relationships and the consequent informal style employed by them in

their suicide notes.

Although these basic word lists do not tell us too much about aboutness, then, they do tell us a little about speech and writing, and the Birmingham suicide notes, presumably due to their generally informal nature, appear to be a little more like speech than one might expect from written texts.

6.2.1. Words Frequencies by Note

I now look at which words are contained in the suicide notes with respect to how many texts contain at least one of these words. Wordsmith Tools 4 has the facility to view word lists by what it calls "consistency", that is, by the number of texts in which the word occurs. This is what I refer to in this thesis as 'by note'. As mentioned above, it is evidenced in the "Text" columns of the word list tables. Although one can see the number of texts that contain the most frequent words in Tables 6.1.a and 6.1.b above, it will be more useful to see the frequencies of the words in order of the number of notes that contain them. This is the purpose of Tables 6.2.a and 6.2.b below. (It should be mentioned that not all the information in these lists of words is of importance, and some of the apparent differences can be discounted.)

The first point to make about Table 6.2.a below is about the string 'FQQQ' (on line 18, in 30 of the texts). This represents a number of anonymised female names (see Chapter 4 for a full account of the anonymisation process). Removing this 'word' from the word list (see below) results in all words below it moving up one place, and the word 'NO' appearing in 25th place. 'No' occurs 36 times (0.48%) in the "Hers" in 27 of the texts (36.49%).

Table 6.2.a. Word Frequencies by Note in "Hers"

N	Word	Freq.	%	Texts	%
1	I	301	3.99	61	82.43
2	THE	193	2.56	55	74.32
3	TO	195	2.59	53	71.62
4	MY	130	1.72	51	68.92
5	YOU	232	3.08	51	68.92
6	FOR	102	1.35	50	67.57
7	ME	104	1.38	47	63.51
8	HAVE	100	1.33	45	60.81
9	IN	87	1.15	44	59.46
10	AND	191	2.53	43	58.11
11	BE	71	0.94	38	51.35
12	ALL	68	0.90	36	48.65
13	NOT	69	0.91	36	48.65
14	OF	100	1.33	35	47.30
15	A	97	1.29	34	45.95
16	WILL	50	0.66	34	45.95
17	IS	69	0.91	32	43.24
18	FQQQ	32	0.42	30	40.54
19	WITH	62	0.82	30	40.54
20	DO	45	0.60	29	39.19
21	THIS	42	0.56	29	39.19
22	BUT	53	0.70	28	37.84
23	IT	71	0.94	28	37.84
24	LOVE	45	0.60	28	37.84
25	ON	40	0.53	28	37.84

Most of the words in Table 6.2.a were also present in Table 6.1.a (see Section 6.2 above). Their placement in both tables is also similar, apart from 'will' and 'it' which are reversed, with 'will' now (in Table 6.2.a) being a few places above 'it'. So the most frequent words in the "Hers" are also the words that appear in most (but not all) of the texts. However, Table 6.2.a shows 'this' and 'on' which do not occur in Table 6.1.a, and the 'word frequency' list shows '#' (numbers), 'X' (kisses) and 'that' which are not in the 'text frequency' list. In fact, all the words which appear only in either one of the tables are in 27th to 29th places in the full lists (not shown), with the exception of 'X' which is in 74th place in the "Hers" by note.

Table 6.2.b. Word Frequencies by Note in "Hims"

N	Word	Freq.	%	Texts	%
1	I	1,580	5.18	176	83.02
2	TO	1,023	3.35	153	72.17
3	YOU	984	3.22	147	69.34
4	THE	912	2.99	143	67.45
5	FOR	395	1.29	132	62.26
6	AND	657	2.15	127	59.91
7	ME	424	1.39	127	59.91
8	MY	445	1.46	126	59.43
9	LOVE	261	0.85	119	56.13
10	HAVE	444	1.45	116	54.72
11	OF	434	1.42	116	54.72
12	A	403	1.32	114	53.77
13	ALL	277	0.91	112	52.83
14	IN	325	1.06	108	50.94
15	BUT	261	0.85	103	48.58
16	NOT	284	0.93	103	48.58
17	IS	301	0.99	102	48.11
18	IT	313	1.03	100	47.17
19	WILL	215	0.70	97	45.75
20	THIS	235	0.77	95	44.81
21	BE	196	0.64	90	42.45
22	ON	172	0.56	90	42.45
23	SORRY	126	0.41	88	41.51
24	#	481	1.58	87	41.04
25	DO	166	0.54	87	41.04

As with the "Hers", for the "Hims" the wordlist by overall frequency (Table 6.1.b in Section 6.2) and the wordlist by text (Table 6.2.b) share most of the same words in roughly the same order. The exceptions are: 'love' and '#' whose places are approximately reversed; the words 'on', 'sorry' and 'do' which do not appear in Table 6.1.b; and 'that', 'with' and 'so' which are absent from Table 6.2.b. However, all the words which only appear in one of the tables are in 26th to 29th places in the full lists (not shown), with the exception of 'sorry' which is in 35th place. So, as with the "Hers", nearly all of the most frequent words are also most frequent in terms of their

usage by text.

It is, perhaps, worth mentioning that although the first 25 words of the "Hers" included one anonymised 'name' ('FQQQ' on line 18, see Table 6.2.a above), the first 25 lines of the "Hims" (see Table 6.2.b) did not include any such 'types'. The first anonymised 'name' in the "Hims" ('FNNN') appears in 43rd place in the list by text (not shown here). It occurs in 60 (28.30%) of the "Him" texts. However, this does not necessarily mean that the female note-writers name people more often than do the male note-writers, as should be clear from the discussion of the anonymisation process (see Chapter 4).

Similarly, the frequency figures for the '#' symbol show that there are no digital numbers in the top 25 "Hers" whereas Table 6.2.b (above) shows that digits appear in 87 (41.04%) of the "Him" texts (line 24). In the "Hers" the '#' appears in 28th place in the list (not shown here). Numbers occur in 25 (33.78%) of the "Her" texts, and make up 2.23% of the sub-corpus as a whole. This means that taken all together the females do use more digital numbers than the males, but the males use them in more notes. This could be due to greater use of digits by a relatively small number of female multiple-note writers. It could be that those particular females write in a more detailed fashion. It might also mean that more males than females tend to find something numeric to write about, but not as much as some females do. It is difficult to know what, if anything, either of these scenarios might mean. However, inspection of some of the notes reveals that the numbers occur in addresses, dates, times, people's ages, etc., and so no real relevance can be attached to these findings.

Two other differences between the "Hers" and the "Hims" in the top 25 words by note are the words 'with' in the "Hers", and 'sorry' in the "Hims". 'With' does occur in the "Hims", in 26th place in the list (not shown). It occurs in 86 (40.57%) of the texts. This seems to be a similar pattern of distribution to that in the "Hers" where it occurs in 62 (40.54%) of the texts.

The topic of 'Sorry' has already been mentioned in Section 5.3. 'Sorry', which occurs

in 41.51% (88) of "Him" notes (line 23), occurs in 25.68% (19) of the "Her" notes (line 37, not shown). This would seem to suggest that the male writers are more likely to express regret than the female writers. However, caution is needed before drawing such conclusions, as the mere word 'sorry' does not in itself indicate that the writer is saying that they are sorry: they might be saying that someone else will be sorry, or that they are not sorry. Looking at the concordance lines (not shown), it can be seen that none of the note-writers is in fact using 'sorry' in such ways, and that well over a quarter of the suicide notes are about their authors' being sorry about something they have done. There were no instances, for example, of 'you'll be sorry'. Indeed most of the females and males who used the word 'sorry' were sorry about the suicide itself and the inconvenience it would cause the addressee. This adds a piece to the picture of what suicide notes are about. Here are some examples.

(6.1. Sorry about inconvenience)
"I'm sorry to put you to all this
trouble, I know you have better things to do." (m96335a).

(6.2. Sorry about past event(s))
"I am sorry in [sic] the way I've used you." (f96320e).

Apart from the words I have mentioned, the 25 most frequent words by note are common to both the "Hers" and the "Hims". Their relative order of placement is also similar. One exception to this is the word 'love'. 37.84% of the "Her" notes, and 56.13% of the "Hims", use the word (Table 6.2.a, line 24; Table 6.2.b, line 9). It is not clear why more males than females used the word 'love'. Most cases of 'love' were directed to the addressee, and this was usually as a valediction, or part of one. In some cases the note-writer asked the addressee to pass on the note-writers' love to others. Some instances were indirect love (with respect to the addressee) like the following:

(6.3. Love indirect)
"I love him…" (f97361a).

There were also some expressions about love, such as:

(6.4. Love itself)
"If ever there was proof of love an [sic] Earth…" (m96316c).

Interestingly, the first example above shows 'love' used as a verb while the second uses it as an abstract noun. The above cases expressed positive affect, but a few notes had instances such as:

(6.5. Love, negative affect)
"FNNN DOeS NOT LOVe Me ANY" [sic] (m96382b).

One remarkable finding is that the most common word in both sub-corpora, 'I', only occurs in 85.14%[33] of the "Her" notes and in 83.02% of the "Him" notes.[34] This is surprising because one expects suicide notes to be about the self, yet these figures are far short of being 100%. It is possible that some of the texts that do not use 'I' use instead 'my' and/or 'me'. It is also possible that there are a small number of texts that do not reference their authors' selves at all. I shall now look into these possibilities. Counting and comparing 'grep' lines gives the following information as seen in Table 6.3 below (in which the raw counts are in parenthesis).

---

[33] The difference between the figures of 82.43% in Table 6.2.a and 85.14% in Table 6.3 can be attributed to the different word counts produced by versions 3 and 4 of Wordsmith (see Section 4.11.).

[34] In Appendix C (on the CD) the filename prefixes of the notes without "I" (or "i") are as follows. "Hers" - F96320A, F96320B, F96320D, F96369A, F96369G, F96460D, F96460P, F97158C, F98279A, F98279D and F99123A; "Hims" - M95161H, M95214C, M95214D, M95214F, M95214G, M95214H, M95408, M95420, M95468F, M96262, M96268, M96318B, M96335B, M96335D, M97247, M97474B, M97474C, M97509, M98058, M98207, M98311A, M98329A, M98329B, M98329C, M98329E, M98329H, M98329I, M98354B, M98523C, M98523D, M98542, M99139, M99183A, M99420A, M99433A and M99433B.

Table 6.3. 'I', 'My' and 'Me'

Corpus	Occurrences	Notes	Subjects
i			
her	4.35% (321)	85.14% (63)	100.00% (22)
him	5.70% (1712)	83.02% (176)	88.04% (81)
me			
her	1.41% (104)	63.51% (47)	86.36% (19)
him	1.41% (424)	59.91% (127)	72.83% (67)
my			
her	1.76% (130)	68.92% (51)	77.27% (17)
him	1.48% (445)	59.43% (126)	69.57% (64)
i|me|my			
her	7.53% (555)	94.59% (70)	100.00% (22)
him	8.59% (2581)	87.26% (185)	92.39% (85)

As Table 6.3 shows, when accounting for 'I', 'me' and 'my' together (that is, logically inclusive 'OR'd) a greater proportion of the texts than is indicated by the figures for any of these words individually do reference the self. But with only 94.59% of the "Her" texts, and only 87.26% of the "Him" texts, using these words it seems that there are some texts (at most 4 "Hers" and 27 "Hims") that do not reference the self at all. The latter were written by 4 "Her" and 19 "Him" subjects. All of the "Hers" and 12 of the "Hims" were multiple-note writers.

Looking at the supposedly non-self-referencing notes in the "Hers" reveals two short notes consisting mostly of commands, such as 'Don't come in' and 'Call the police', and two notes that are lists of people their authors wanted to be contacted. Looking at these notes in the "Hims" revealed a similar situation: there were commands giving instructions and lists, but there were also some notes that had managed to communicate without using any of the three pronouns: for example, a note that just said 'Sorry'. It would be incorrect to assume on a semantic level that this did not mean 'I am sorry' and was devoid of self-reference, although it is true that on a lexical level these texts had none of the three 'self' terms. The same can be said of the 'I-me-my-less' notes that are not instructions: nearly half of them are signed. There was yet another type of text in the "Hims" that had been counted as not using 'I', 'my' or 'me'. This was because of spelling mistakes, with two of the note-writers using "Im" instead of "I'm". So

although there are some texts that do not reference the self, there are not quite as many of them as there initially seemed to be. In fact there are 25 "Him" texts by 17 subjects, 11 of whom were multiple-note writers. This look at three personal pronouns does seem to have been enough to confirm that most suicide notes are, at least to some extent, concerned with the self.

It is not so remarkable that most notes do contain 'I'. However, another highly frequent word in the lists is 'you': nearly 70% of the "Her" and the "Him" notes contain it. Because of this, in Section 6.2, it was posited that the notes are partially about the addressee. Given the number of texts containing 'I' some notes must contain both 'I' and 'you', and thus refer to both the author and the addressee.

It is interesting that none of the words is featured in every text. It can be seen that it cannot be said that all suicide notes are about any one common topic when looking at only the first few words, and looking at them in isolation from the whole word list. This further indicates that some kind of grouping of words will be necessary to establish 'aboutness', and that this should include words from the whole word list. The "Hers" and "Hims" have nearly 900 word types in common. However, the "Hers" have over 500 types that are not in the "Hims", and the "Hims" have almost 2,400 types that are not in the "Hers". This implies that any full treatment would be greatly aided by further computer analysis. I pursue word groups, and whole word lists, in Section 6.3.2 below.

However, aside from word groups, there is also the matter of distinctiveness. It would be good to know that suicide notes tend to be about x, but texts of other genres might also tend to be about x. It would be useful to separate out the differences between what suicide notes are about, and what they are peculiarly about. So rather than look at more words from the word frequency lists, I next considered what I (following Scott) have termed 'key words'.

6.3. Key Words

The raw word lists above did not really help too much in the quest for suicide note 'aboutness'. Although I was considering only the first 25 lines of output it was

apparent, because no single word was present in all texts, that a different approach was needed: for example, one that automatically grouped synonymous words together. However, before embarking upon this I wanted to compare the Birmingham sub-corpora with a reference corpus. This would show whether either of the sub-corpora was markedly different from a general (and therefore not specifically suicide-orientated) collection of texts. This would also help determine not just how each of the sub-corpora was different from the reference corpus, but whether they were different from it in the same way, and could be used as a way of comparing them.

These comparisons were carried out using Wordsmith Tools' 'Key Words' option. When utilising key words the assumption is that the texts will be 'about' a restricted range of things, and therefore have something in common in terms of 'aboutness'. Key words in Wordsmith are defined as those words that are more frequent in a text than one would expect when compared to some norm such as a much larger corpus (by default the BNC). They can therefore be indicators of what a text is about. As Wordsmith's creator says, "Key words usually give a reasonably good clue to what the text is about." (Scott, 2007b, p.121).

However Scott was assuming that the subject of investigation would be a single text. Scott's example in the Wordsmith manual concerns comparing a single text with a set of several others, all the texts being newspaper articles, and the single text being one of the set (Scott, 2007b, p.121). But most corpus linguistic work involves corpora consisting of more than one text, and there is nothing to prevent the researcher from running Wordsmith 'Key Words' over several texts at once. Of course, if one uses an entire (sub-)corpus as the subject of investigation then key words can only be indicative of what might be key over all the "Hims", or over all the "Hers", as a whole. Nevertheless this might at least give a sense of aboutness for a sub-corpus.

Wordsmith found 107 key words in the "Hers" and 229 in the "Hims". However, about half of the top 25 key words in the "Hims" and in the "Hers" are nonsense words. This is due to the anonymisation procedure described in Chapter 4. These words mainly represent names of people and parts of addresses (such as names of streets). That so

many of these pseudo-words show up as being 'key' to suicide notes, when compared with the standard (BNC) corpus, is not surprising because they are invented and thus likely to be peculiar to my corpus. Having anonymised the texts in a manner whereby, for instance, FQQQ could really be a mixture of 'Mary's and 'Jane's and many other female names, it is no longer possible to tell whether any particular names might be key. Of course the 'keyness' of a particular name does not have any relevance for this work, but finding that names in general were key *is* important. Others have already found that suicide notes have many references to names (see Chapter 3). Having said that, however, some of the names might be key anyway (in a non-anonymised scenario) because they are rare or unique (as in the invented 'Marypoos'), rather than because they are common in a suicide note corpus but considerably less frequent in a reference corpus.

To get a better picture from which to consider what suicide notes are (or might be) about it was important to reveal some more key words and to omit the anonymised names from the list. To achieve this, a 'stop list' of the anonymised names was created. This was a text file containing the 'words' to be excluded from the key words display. This was done by putting the anonymised names that appeared in the key words, of both the "Hers" and the "Hims", in the file, and then running the key words application again with the stop file activated. The stop list that removed the bulk of the anonymised names contained 68 'words'. The resulting new lists of the top key words are shown in Tables 6.4.a and 6.4.b. In these Tables "RC" refers to the reference corpus used. This was the BNC as supplied by Scott (2007c).

Table 6.4.a. Key Words: "Hers", Stopped

N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P
1	I	301	3.99	732,523	0.74	534.48	0.0000000000
2	MY	130	1.72	146,775	0.15	403.12	0.0000000000
3	YOU	232	3.08	588,503	0.59	394.80	0.0000000000
4	X	53	0.70	7,783		371.95	0.0000000000
5	ME	104	1.38	131,757	0.13	300.38	0.0000000000
6	LOVE	45	0.60	22,224	0.02	209.17	0.0000000000
7	AM	39	0.52	26,042	0.03	158.77	0.0000000000
8	SLAG	13	0.17	186		150.65	0.0000000000
9	B'HAM	8	0.11	5		134.48	0.0000000000
10	PLEASE	28	0.37	13,809	0.01	130.17	0.0000000000
11	DEAR	24	0.32	8,487		127.00	0.0000000000
12	SORRY	24	0.32	10,778	0.01	115.89	0.0000000000
13	LIFE	38	0.50	54,494	0.05	101.03	0.0000000000
14	MOM	9	0.12	244		93.09	0.0000000000
15	LOVED	16	0.21	4,592		91.16	0.0000000000
16	DONT	8	0.11	139		89.69	0.0000000000
17	BLAME	14	0.19	3,045		87.36	0.0000000000
18	HAVE	100	1.33	448,684	0.45	84.25	0.0000000000
19	OVERWELM	4	0.05	0		75.90	0.0000000000
20	THANK	18	0.24	11,714	0.01	74.11	0.0000000000
21	GRIEVE	7	0.09	178		73.28	0.0000000000
22	TRYED	4	0.05	1		70.90	0.0000000000
23	RD	7	0.09	234		69.52	0.0000000000
24	XX	6	0.08	111		66.54	0.0000000000
25	ALL	68	0.90	277,566	0.28	65.88	0.0000000000

In Table 6.4.a (above) 'RD' is an abbreviation for 'Road', part of an address, and 'XX' stands for a couple of kisses. Prior to removing the anonymised names (not shown), the top key words in the "Hers" consisted of the pronouns 'I', 'you', 'my' and 'me', kisses ('X'), 'love', 'am', 'slag', 'B'ham', 'please' and 'dear'. 'Love' has already been mentioned (see Section 6.2.1). 'B'ham' is an unanonymised abbreviation for Birmingham. The presence of this as key is not surprising because my entire corpus consisted of notes found within the jurisdiction of the Birmingham Coroner, and therefore it was not necessary to anonymise the city's name, which frequently appeared in addresses. The appearance of 'dear' as a key word is not surprising because in notes and letters it is a typical salutation (see Section 5.2). That leaves the words 'am', 'slag'

and 'please' as thus far unexplained words that might be key to what this sub-corpus is about. Before considering this further I looked at the key words of the "Hims" (Table 6.4.b).

Table 6.4.b. Key Words: "Hims", Stopped

N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P
1	I	1,580	5.18	732,523	0.74	3,509.91	0.0000000000
2	YOU	984	3.22	588,503	0.59	1,749.61	0.0000000000
3	LOVE	261	0.85	22,224	0.02	1,393.25	0.0000000000
4	MY	445	1.46	146,775	0.15	1,242.60	0.0000000000
5	ME	424	1.39	131,757	0.13	1,229.28	0.0000000000
6	SORRY	126	0.41	10,778	0.01	670.96	0.0000000000
7	MOM	55	0.18	244		604.61	0.0000000000
8	PLEASE	121	0.40	13,809	0.01	577.01	0.0000000000
9	AM	145	0.47	26,042	0.03	566.33	0.0000000000
10	IM	54	0.18	678		488.69	0.0000000000
11	DONT	39	0.13	139		443.91	0.0000000000
12	HAVE	444	1.45	448,684	0.45	429.80	0.0000000000
13	IVE	33	0.11	72		403.24	0.0000000000
14	LIFE	124	0.41	54,494	0.05	282.43	0.0000000000
15	ALL	277	0.91	277,566	0.28	270.66	0.0000000000
16	DEAR	60	0.20	8,487		261.36	0.0000000000
17	HOPE	75	0.25	17,128	0.02	259.07	0.0000000000
18	DAD	53	0.17	5,722		258.29	0.0000000000
19	KNOW	161	0.53	118,820	0.12	229.45	0.0000000000
20	CANT	21	0.07	131		217.75	0.0000000000
21	FORGIVE	32	0.10	1,297		216.82	0.0000000000
22	BYE	30	0.10	1,791		180.62	0.0000000000
23	YOUR	153	0.50	134,393	0.14	177.91	0.0000000000
24	WILL	215	0.70	251,179	0.25	165.73	0.0000000000
25	WISH	48	0.16	11,446	0.01	161.92	0.0000000000

Prior to removing the anonymised names (not shown), the top key words in the "Hims" consisted of pronouns (the same four as in the "Hers"), the words 'love', 'please' and 'am' (as in the "Hers"), three mis-spellings ('im', 'dont' and 'ive', where 'I'm', 'don't' and 'I've' were probably intended), and 'sorry', 'mom', and 'have'. It is not, perhaps, surprising that mis-spellings are indicated as key words. 'Sorry' was discussed in Section 6.2.1 above. That leaves the words 'am' and 'please' (as in the "Hers"), 'mom'

and 'have' as thus far unexplained words that might be key to what the "Him" sub-corpus is about.

With the anonymised names gone, those words that were in the top 25 of the "Hims", but not in the "Hers" ('sorry', 'mom' and 'have') are now also in the top 25 of the "Hers" (Table 6.4.a). And in Table 6.4.b, with the names gone, the word ('Dear') that was in the top 25 of the "Hers", but not in the "Hims" is now also in the top 25 of the "Hims". Along with several mis-spellings there now also appear some other words: 'loved', 'blame' and 'thank' in the "Hers"; 'hope', 'dad', 'know', 'forgive', 'bye', 'your', 'will' and 'wish' in the "Hims"; and 'life' and 'all' in both sub-corpora. Interestingly, over a quarter of the words is both lists are verbs, nearly all in the present tense, but 'is' does not appear in Tables 6.4.a and 6.4.b.

It's notable that 'love' and 'life' are among the top words, but words such as 'death', 'die', 'suicide' and 'kill' are not. It is possible that for some note-writers, because the death is evident from the note and the body, 'death' is a redundant concept, and is referred to by deictics which from the note-writers' point of view are direct references (although the researcher analysing them may consider them to be indirect). The use of deictics (demonstratives, personal pronouns, deictic adverbials and time adverbials) within a text as pointers to the physical world is discussed in Scollon and Scollon (2003, p.31-40). The latter show that the interpretation of texts ought to take account of their physical location or material placement. The linguist, as well as the Coroner's Office, must give due regard to the location of a suicide note, including its location relative to the body (see Sections 2.2.2 and 4.4 above).

Figures 6.1 to 6.17 present a selection of concordance lines from the Birmingham corpus (with the "Hims" and "Hers" combined) showing some of the key words. They are sorted '1R 1L' (i.e. by the immediate right-hand collocate followed by the immediate left-hand collocate) unless stated otherwise. For each set of lines I have indicated the parameters used to compile them with the WordSmith 'random' function, and the total number of lines which WordSmith found.

```
1                                   DENTIST   PLEASE BE KIND  ENOUGH TO I
2                              WILL YOU PLEASE     BE GOOD ENOUGH
3       that I am personally to blame , But please carry out my wishes, if you
4        UUU   OF KIN-->  9 BBCCBB RD  PLEASE        CCDDCC (NEAR HH
5                                   PLEASE DO NOT  GO INTO THE
6   you  all the best for your future and please don't make the same  mista
7     [for]  you were always there for me . Please don't blame   yourselves for
8    you'res to do with as  you see fit . Please don't hate me for this for it i
9    and the only way out is to do this. Please dont waste your life wonderi
10                    My Darling       Please forgive  me I am too much in
11      is that what we had was wonderful, please forgive me for deceiving  you
12      there is a heaven for me to reach . Please give little Mppp a  hug good
13      d  out any problems we had  Please Please go and see her  (Fsss friend
14      ostly)  If you can be strong enough please help  Mppp for me when you
15                                   Please let the following know.    1.
16    but its  about time I did somthing to please me can't face  it any more ju
17     you and hate myself for doing this, please, please forgive me.
18    pen.  Its the will of me to do this so please  respect that will  I'm so sorr
19               NNN   and MPPP.     Please talk to each other.  I am sor
20      of it is missing from my memory.  Please try to go on.  I will write to
21     es and 2nd thoughts  altered that.  Please try hard not to hate me.
22    ut very bad at others. Fnnn will  you please write to Fppp for me  in Thail
```
Random 1 in 7 (of 149)

Figure 6.1. Concordances in "Hers"+"Hims": 'Please'

```
1     bless, Mppp        If there is a way to thank mr. Qqqq,& Dr. Rrrr  possibly
2                              FUUU      THANK U. U TAUGHT ME   HOW T
3      ut  you are the only one I can ask   thank you for all your help the last t
4      d  my ex, nice to now he still cares. thank you
5      shall no longer be with you.     My I thank you most sincerely for  all the
6      all no longer be with you.     May I thank you both from the bottom  of
7    have always been  a faithful friend. I  thank you for that. I  trust you to un
8      om for me and give her  a big kiss. Thank you for looking after me   whil
9                    Dear Mrrr       Thank you for coming round  and try
10    t writing this  I owe you so so much, Thank  you for everything.
11  r me to  put things rite,            Thank you                Mrrr
12      love you more than words can say. thank you for the best 11 months of
13    sickness is  more than I can take.  Thank you all for everything.
14                    Fqqq          thank you for all your help, but  I ha
15                                   Thank you Mccccc,     my little darli
16                                   Thank you both Fsss and Fttt  for th
17                    EST WEEK   Thank    You
18    rn  Back the clock about five  years Thank you for helping me  on my w
19    v.        I am so sorry to upset  you. Thank you so much for all  you have
```
Random 1 in 2 (of 38)

Figure 6.2. Concordances in "Hers"+"Hims": 'Thank'

```
1              S you never WANT TO  LISEN Ime sorry About   the WRITING But I can
2   your ------------------------------- PS SORRY ABOUT THE      STATE OF
3        OU THIS    LETTER PS            IM SORRY  ALLWAY THINK ABOUT
4                                          SORRY  BUT PERHAPS WE ALL
5          n have some peace.  once again im sorry for  the trouble ive caused.  B
6      e good to each  other . Love Rrrr .    Sorry for Inconvience
7        st of time hope  you can understand sorry for letting you down  Back to
8        to her as soon as possible i'm very sorry  for doing this because all my
9         dy. I want you to  know how deeply sorry I am to cause you  this pain. I
10      , they  are something special    I'm sorry, I really am          Love y
11                              MOM   SORRY. I REALLY , REALLY LOV
12      ou  Thanks for our   time together.    Sorry I let you down  You can get o
13        mily    I'm SORRY , I'm so very very sorry!         I hope you all know how
14          king . Remember me  as I useto be sorry if Ive let you down  Fppp hope
15           ART, I CARNT  CARRY ON, IM SO SORRY [?]           Mnnn [?]       [?]
16    the house anyway.                      very sorry
17                        GOOD BYE  FNNN  SORRY  singing in the       rain PR
18   ere for me. Thanks!            I'm So Sorry , Take Care
19         s on my house and yours, but I am sorry.  Tell the kids I loved them al
20      , help, and appreciation. I am really sorry that I cannot repay this in a  b
21       seem to think I'm a week person  so sorry that I've let you down again.
22    I loved them all very much  and I am sorry, they had to be in this       Fss
23                         Fppp & Mqqq  sorry to End this Way  hope you all
24  ----------    good.     I am so very very sorry to  let you down. I did my best
25       have tryed so hard to help  me I am sorry, you have all tryed so  hard
26    ined in the other letters.   Again I'm sorry, you look after yourself, you ar
```
Random 1 in 8 (of 150)

Figure 6.3. Concordances in "Hers"+"Hims": 'Sorry'

```
1              USE WE   COULD NOT TALK ,AND FORGIVE, JUST CRITISE, AND BE
2        he treatment didn't  help. May God forgive
3                        nn            pleas forgive me I  just can not keep on li
4      if there is a god??? Signore  please forgive me - do not want  to be burd
5                     Darling     Please forgive  me I am too much in  Pain.
6     lated taking  such a step.   Please forgive me and  don't think I'm a co
7         what we had was wonderful, please forgive me for deceiving  you all the
8                                   Please forgive me I don't want to live any m
9      wing  how to escape.  Please try to forgive me  but I fear given time I wo
10    ay I hoped  you would find easier to forgive me than  if I had continued t
11        TO FNNN AND ALL       GOD WILL FORGIVE ME
12  Q             DAD xxxx              FORGIVE ME
13     If you can find it in your heart to forgive me, I shall look forward to se
14                 So sorry hope you can forgive  me.  youve both done so m
15                                Please forgive me. I  cant stand anymore
16    ther signs of deterioration.  Please forgive me. I've  enjoyed life and the
17     put you to so much  troubl.  Please forgive me.  Can you find a nice  ho
18                          , Please forgive me.  I'm sorry to  let you all
19    caused.  But i hope in time you  will forgive me.
20      S  Phoning the police on me  I cant forgive you for  that. Ime not DRUnK
```
Random 1 in 2 (of 38)

Figure 6.4. Concordances in "Hers"+"Hims": 'Forgive'

```
1          y.    Mum, Dad and Pppp are not to blame because I never asked them f
2      hip, I realise  that I am personally to blame , But please carry out my wi
3      ", "why didn't he that" and the "well I blame Fnnn, he was  never the sam
4       ot your Life In front of you you must  Blame me fore this But your Quality
5        T  GRIEVE I'ts better this way  Don't blame Mnnn he has made me  so h
6      t I wish I  had never been born. Don't  blame  Mnnn, I knew the score, I lov
7          t  carry on anymore. Please do not  blame Mnnn, for the first time I  felt
8        e a burden to anyone.    There is no  blame on anyone else, in the words
9       not lied to me,    there is no one to  blame only him       and his lies
10            WANT TO  SAY BUT IT'S EASY TO  BLAME   SOMEONE ELSE
11       ay be hard, I know that you want to  blame someone, so  blame me for n
12   ery difficult  at home but take the full  blame.  The silence at  home at tim
13           FRenS TeLL  FTtT + MNNN I DONT   BLAMe Them AND I  Love MsSs A
14    h positions.    I don't want anyone to  blame themselves for what I've done
15       all very much.    Fppp please do not  blame  yourself as I have brought thi
16   Y. The last thing I  want is for you to  blame yourself because it's not  you
17      s.  My only concern is that  you will  blame yourself  for not supporting m
18   hat it is is that I wish that you  don't  blame yourselves. I emphasize you
```
Random 1 in 2 (of 31)

Figure 6.5. Concordances in "Hers"+"Hims": 'Blame'

```
1           ELING WORN OUT OR AGITATED  CANT BE BOTHERED WITH ANYT
2       ression thing's was looking so good cant  bielive this happend its like a
3       have some good times , I think ? , I cant cope  with you moving on with
4    he best for the future.  Im sorry but I cant cope anymore.  it seems that
5                     QQ   I JUST CANT COPE AT ALL LATELY    I A
6                     V    I REALLY CANT COPE MATE!  KNOW THAT
7         REALLY , REALLY LOVE  U, BUT I CANT COPE. I HAVE TO  GO. LO
8       uch and shes another  reson why. i cant do anything  right can i. youve
9        PLUS  Phoning the police on me  I cant forgive you for  that. Ime not D
10  ------------------------------    IF YOU CANT GET HER FOR SOME REA
11  as well as our rascally relations.    I cant help wondering if they have  le
12     ow)  I have thought about this day.  I cant justify to any of you why I've do
13  on myself.            I am sorry but I  cant live without yo
14   e up smoking but          I definatly cant smoke           now    Do go and
15      owever this is one problem even you cant solve nobody  can except the
16                Please forgive me. I  cant stand anymore          Love ca
17     ighbours have put me through  and i cant take a single day more have a
18        NY TROUBLE   I MAY LEAVE      I CANT TAKE ANY  MORE I WAS V
19     ] Feelings But Dont morn  over me I Cant TAKE THE  LYe's no MORE
20           VE TO  GO. LOOK AFTER FNNN. I CANT TELL   U HOW MUCH I WISH
21        I TELL U  HOW MUCH I LOVE U. I CANT.  THERE AREN'T WORDS.
22        HE BALCONY  FNNN , SORRY &  I CANT THINK OF MUCH  TO SAY A
23     TS All I Can say  I Love you to Bits  Cant WRITE NO MORE    AS Ime
```
23 of 23

Figure 6.6. Concordances in "Hers"+"Hims": 'Cant'

```
1          the year, but I know my limitations, I can't be the vicious nasty bastard, it
2            ve talked to me I love her so much I can't ber not being with her its killi
3      l until now. I'm terrible with  people. I can't come to terms with it, can't  fo
4    his letter will answer the 'why' part, I can't do much about  the reactions.
5          retend to be someone I am not and I can't do that    I would like everyone
6        riding on the hard work  of others. I can't do that. I want to be successf
7          ut time I did somthing to please me can't face  it any more just want out
8      people. I can't come to terms with it, can't  focus on the positives, the ne
9        nfidence  Give this plenty of thought. Can't get out of this  depression thin
10            MOM! MY LOVE of my LIFE  BUT I CAN'T HANDLE IT NO  MORE SO t
11      ut I can only  WRITE this letter as I can't  KeeP my EYES OPEN PLUS
12    e had it good here, couldn't we?     I can't leave you anything except my
13    NTED  Fppp           I am very Sorry I can't  see any other way out always
14        on't know  what's wrong with me & I can't seem to understand anything
15    can't sleep feel like Ive let you down  can't take it               Don't Forge
16      t. is she doesn't love me any  more.I can't take the waiting any more. I lo
17    't be  sad . I just want to end all this can't take not being  with Fppp its n
18        me through this past  year . mom i can't think of anything else to say i
19      and enjoy your remaining years.   I can't write any more.   My love to F
20            'S THINGS I WANT TO  SAY BUT I CAN'T. FIRSTLY I  STILL LOVE YO
```
Random 1 in 2 (of 47) sorted 1R 2R 1L (apostrophe is counted as 1R)

Figure 6.7. Concordances in "Hers"+"Hims": 'Can't'

```
1          DOPe FRenS TeLL  FTtT + MNNN I DONT  BLAMe Them AND I  Love
2      ted perant   sorry mom i did my Best Dont Blem Nnnn      KID Dont forget
3                                          DONT COME      IN  CALL THE    P
4          FRESH START AND ENJOY  LIFE (DONT FUCK IT UP!!)          EXCUSE
5      id my Best Dont Blem Nnnn     KID Dont forget me i did love you  But y
6          WAS VERY BAD  LAST NIGHT      DONT GO UP STAIRS  ON THE LA
7      both have  done for Mccc and me    Dont grieve or be upset. To go  whe
8                          DAD    DONT KNOW HOW TO WORD THI
9    SHE IS          LIVING WITH.      I DONT KNOW FQQQS BECAUSE S
10    cc.  You have carried us both,  and I dont know what either of us would
11          WHO HAVE TO DEAL WITH ME  - DONT KNOW YET WHETHER I HA
12      BBBBBB            CCCCCC RD    DONT KNOW THE NUMBER.
13    th fppp being  there is just crawl you dont know  whats shes like. she ma
14    you like this  after all you did for me  Dont let our fr
15          THIS    YOU CAN ALL GET ON    DONT LET ME BRING YOU
16            moment[?]  Feelings But Dont morn  over me I Cant TAKE T
17    t . I happy  to do what I'm about to, I dont  think that any one could help
18      d up and walk through this . People dont  understand how much this ha
19          I HOPE  SO BECAUSE I REALLY DONT WANT TO LIVE ANY  LONG
20    e only way out is to do this.  Please dont waste your life wondering why I
```
Random 1 in 2 (of 47)

Figure 6.8. Concordances in "Hers"+"Hims": 'Dont'

```
1         an't  pull myself up mom dad please don't be sad I'm not  Im happy now l
2      ed. What it is is that I wish that you  don't blame yourselves. I emphasiz
3           OU TOLD ME  LAST NIGHT THAT I DON'T CARE ABOUT MRRR  WHA
4      ves toiling over a job  that 90% of us don't enjoy most of the time, for 45
5       and of course the house is yours.    Don't forget the deeds are with Qqq
6      could not do it sober  all I can say is don't grieve  me & "COMON OLOS"
7      s to do with as  you see fit . Please don't hate me for this for it is more
8           ILLION THINGS BUT I DO NOT.    I DON'T HONESTLY WHAT DAMAGE
9                    TO FQQQ,         I DON'T KNOW WHAT TOO SAY  T
10        happy now love you. Thanks mnnn don't know what  I would have done
11     I don't want to be without  You . You don't know what its been like the la
12       ch a step.    Please forgive me and  don't think I'm a coward.  It has take
13  ----------   Just cancel all Payments.  Don't think I am insane  I am quite s
14        Do not be shocked -  I am in bed Don't touch me at all       Fnnn NOT
15     t I love you both, and am sorry,  you don't under stand, he was my life  a
16      wining you back but its to late now  don't use people in life its not nice
17  ou all into pretty tough positions.    I don't want anyone to blame themsel
18      ll you what was going on  because I don't want to be  a pain any more.
19   t have  they got to look forward to? I don't want children because I canno
20      care every one love you  all so much Don't want to turn in to a sad old git
```
Random 1 in 5 (of 99) sorted 1R 2R 1L (apostrophe is counted as 1R)

Figure 6.9. Concordances in "Hers"+"Hims": 'Don't'

```
1        will sucksead at the highest  level in love and life . I would love  to see
2        already.     Please try to give Mnnn love and understanding,  as well as
3     u , sorry I messed up  by getting ill.  Love for ever.
4  all of this . I tried to hate her but  the love for her is so great. I know peopl
5   l Frrr herself if possible and tell her I  love her beyond comprehension and
6           THE MOON . SHE DOES  KNOW I LOVE HER BUT IT'S FAR  MORE T
7                          Love[?]     I have let you down [?] b
8      lace up a lot.     I send you my fond love. May the days  ahead bring de
9        OOD Night  its no way to make you love me  But im sueisidel Let me sl
10    ot).    God Bless,    Lots and lots of love,      Mijk.
11    all the luck      in the world       love Mqqq   This is No Joke        P
12        e Them AND I  Love MsSs ALL MY  Love Msss I WOS  GOING TO PHO
13  d that this is not a retraction of that  love ; the reason I have taken my o
14      y Fppp & children.  Tell them how I love  them.       From me - Fqqq
15        S OFF, yours Mnnns wife. Give my love to all the family. Give my love t
16    ave a happy life  together       My love to you all.           FRRR xx
17   yyy  like you are doing.       All my love to you & to Myyy        & Hap
18      ry Always  Remember I I will always love you
19    pleasure of Winding up the shop.    I love you and hate myself for doing t
20      pe with you moving on without me  I love you & always  will.           Ha
21   "Tamborine  man" at my funeral      I love you and Dad    So very much
22  irst tedy I bough you       Frrr       I love you      Could you Have[?] Aabc[
23     ch    Faaa you Im so pround of you I love you     Mbbb Do Well Be Happy
24      ays IN MY  THORTS All I Can say  I Love you to Bits  Cant WRITE NO
25  d no right, absolutely no    right       Love You Forever.   My nan will help
26      he message  and when I said wher's love you said you didn't  mean it lik
```
Random 1 in 11 (of 306)

Figure 6.10. Concordances in "Hers"+"Hims": 'Love'

```
1                                     Fzzz dear,        A last word to you of
2                                     Dear Fnnn           Im very sorry  to
3                                     Dear Fnnn , Mppp &  Qqqq I am in
4      ,                  B,HAM 99, 9AA.  Dear Fnnn & Mppp,       I am enclosi
5                                     Dear, Fnnn, Mppp.               so
6                                     Dear fppp        i have often told you
7                                     Dear Fnnn,        I Love you, I am
8                          My Very Dear Friends Fsss and Fttt,       Firs
9                                     Dear Frrr,          You were a grea
10                                 1 Dear Fsss and Mttt ,
11                                   Dear Fsss and Mttt,                I
12                                   Dear Fsss & Mttt,                I ca
13                              My dear Fuuu,        When you receive
14                 12 May 1998   Dear Mdef,               Further to
15                                   Dear mom        by the time you rea
16                                   DeAR MOM: I AM SORRY I HAVe
17                                   Dear Mppp,   Something terrible has
18                                   Dear Mrrr & Fsss        I have let yo
19      th Feb 1996     My Farewell Note   Dear Nnnn and Mppp      Please do
```
Random 1 in 4 (of 84)

Figure 6.11. Concordances in "Hers"+"Hims": 'Dear'

```
1                LIVE   WITHOUT ANY HASSLE AT ALL.    DON'T WORRY ABOUT IT I
2     ant to be happy  sorry for letting you all down              All my love
3            ED TO DO THE BEST I COULD  AS ALL FATHERS ARE SUPPOSED   T
4     n the one I have led.    All my love to all have know
5         y Photos  By the Giro[?] of me you all I deasire a Good[?]  lie in Heaven
6     port from everyone else.    I love you all.
7     --------------------------    I LOVE YOU ALL MORE THAN  YOU'L EVER K
8         een a constant help as he sorts out all my tax and revenue  problems.
9         we had  and dance your hearts out. All my sisters I love you all  What
10          THOUT  YOU.           YOUR WITH ALL
11        thier  Nan and Grandan Cdef + defg  all of the familey  try and get Efgh
12         r knew me can't get  my head round all of this . I tried to hate her but
13     levision and video   To Ssss Tttt     All of my camera equipment and st
14      ese  dark days.         Please thank all our friends.  Ssss Tttt and Mrs R
15      r being doubly incontinent  that was all something that would clear up.  (
16     ost sincerely for  your friendship and all the many  things you have done f
17          nd pans  Back and the beding   all the luck   in the world    love
18     --------------------    I am so sorry for all the trouble I've caused. I can onl
19          OF THE FACT , I LOVE YOU, WITH ALL THE WORTS[?]  AND STILL D
20     have to get out of here. I  don't want all this hassle  and having to move
21             T REAL UNDERSTANDING, WELL ALL THIS  BUILDS UP. AND I END
22          IDAY 14th AUGUST 1998.   TO MY ALL TIME LOVE FNNN.     BY THE
```
Random 1 in 11 (of 345)

Figure 6.12. Concordances in "Hers"+"Hims": 'All'

184

```
1        hears about what I have done, there will a whole rainbow of reactions  an
2                        I know that a last will and testament should  be witne
3        consenting  to be executrices of my will, and all the time and  trouble th
4     ou receive this letter Mqqq and   me will be dead.        Mqqq & Mrrr
5             T  IF MYYY COMES ROUND  SHE WILL BE  IF NOT MZZZ WILL COM
6          K IN  ENVELOPE)     THINGS THAT WILL BE COVERED IF I  DIE. 1/ P
7   d  It all everything seems  diluted. It will come as  a shock to you all to fi
8   t me from forfilling this dream . but I will go to my grave with my soul  sh
9             AT LEAST NOW YOU! AND FTTT  WILL HAVE SOMEWHERE TO LIV
10             O MORE  MEN HURT YOU ,  YOU WILL HAVE  THE MAN YOU ALW
11   jail for something  I did not do. You  will have to live with it for the rest
12    ght       Love You Forever.  My nan will help me now
13             T THiNK TOO BADLY  OF ME  MY WiLL IS WiTH MR PpPP  OVER T
14  rter of your Strength  of Character.  I will miss you but please don't greav
15   ere is a funeral benefit  payable (you will need to mention this to both Mv
16       you anyway  I promise you that you will never go alone.  - She then turn
17   e to look after Frrrn for me because I will  not be able to.  FQQQ HAS N
18          O EXTEND IT  HOPE SMALL PIPE WILL  NOT MAKE MUCH DIFFERE
19   fgh's" at the Ghij and ask for Mijk he will organise  everything, explain tha
20      eet again in the Hereafter, when we will  really know what our grandpare
21   . I am hoping the  sale of the house will recompense them all.    I would
22    The big problem is  that the Council will require the Flat emptied virtuall
23                        Mnnn   You will soon find your  daughter out. all
24     d love with you both, and hope you  will spare a thought for me just now-
25   o do  a great kindness for me.        Will you make sure  I am not breath
```
Random 1 in 10 (of 265)

Figure 6.13. Concordances in "Hers"+"Hims": 'Will'

```
1        nd may you  find in the future all you hope for  and expect.      Your famil
2          rlfriend  FUUU. Best regards to Mvvv hope he's okay.  Theirs only one pr
3     ll be a release and if there is a God I hope I will find  him.    Nobody is to
4     and Mrs Rrrr are  executors.        I hope I've said all  that I should.
5     rry for  the trouble ive caused. But i hope in time you  will forgive me.
6      a gig. My guitars are  worth £500 , I hope it covers the cost.I'm sorry.go
7     ality, again  I'm sorry and I really do hope that we can meet at some futu
8      lf & Fnnn  whom I love very dearly.  I hope  that you can find it in you to
9      after  I hope you will be  happy and I hope  that Fppp gets on  allright . I
10     indeed truly grateful.    I sincerely hope the future will  hold better time
11    e I have been so bad.  I did not no, I hope will  Find that new life. For  yo
12        I leave my love with you both  and hope you wont forget me for a  little
13     s.      I leave my love with you,  and hope you will remember me  someti
14                        FRRR       HOPE YOU  are okay,  see you so
15   Fuuu in any way.              Mttt I hope you  will understand me and al
16          ATED CAN YOU SAY THE SAME I HOPE YOU CAN  LIVE WITH THE
17     eserve as you  are a lovely person. I hope you  recive all you deserve in li
18             E AREA AND ALL GOING  WELL I Hope you will BE HAPPY  AGAIN
19  f.  it is your  turn to be looked after  I hope you will be  happy and I hope
20       e always loved you  very much and I hope you can  survive what I have d
21     more. Sorry if life got boring with me hope you  fined someone whos goo
```
Random 1 in 4 (of 89)

Figure 6.14. Concordances in "Hers"+"Hims": 'Hope'

```
1        d  tell her what I wrote  it is my only wish and  keep in contact with her
2           o choose them , this is my  express wish as I have explained in the othe
3            not see you again.Hate being alone wish I could  get on with it. But Ive
4            signs of being unhappy  sorry Fppp wish I could of gave you more .
5   I hope you do  it ) well she's got her  wish I love you and I hope  our love
6               I WILL BE WITH  YOUR DAD .   I WISH I HAD TIME TO WRITE  A MI
7   whole life into one long nightmare . i wish i could have  done something t
8   for yourself  with your Dad's help.  I wish I could have still been  there fo
9            were doing  was trying to help me i wish i could have told you to  your f
10         citate  me without my clearly stated wish, I  understand I have the right t
11               R ANY OTHER MEANS.    I ALSO WISH MY MOTHER AND  FATHER
12    t have  happened. What it is is that I wish that you  don't blame yourselv
13      e more as the years have gone by  I wish that I possesed a quarter of yo
14     per now that  i am out of the way , i wish that i  could prove you wrong y
15      been yeting wars i did not no  i only wish that i could put it rite but i ju
16    ntly. They've all got great lives and I wish  them all the best in the future.
17               VE U MORE  THAN WORDS. I SO WISH THERE  WAS ANOTHER WA
18   ----------------------  4/   I PLEASE WISH THIS SO SHE  DOE'S NOT
19               HEM FOR MY  ACTIONS.   I ALSO WISH TO APOLOGISE TO  EVERY
20   you wanted & still do  a divrice this I wish to rescind & therefore leave yo
21               FEEL HUMAN AGAIN.  I REALLY WISH U HAD FELT THE  SAME A
22    same  mistakes as me.                I wish we got to  know each other mo
23      you. Mvvv too is the same. I  only wish we got to know each  other (m
24   o love you both and tell  Mrrr I really wish we were  closer and I would ha
25   ----------------     and let you have your wish what youve  wanted for ages. a
26               ER  YOU SPREADING HER LIES. I WISH YOU PUT  MY NAME ON TH
27      a       Have a good life together & I wish you all  the happiness in the w
28            U. - YOU WANT TO TALK. SO     I WISH YO
29      ou will be hearing from  her I do not wish You any  luck. I did everything
30    py just  like us before Mvvv. I do not  wish you best of luck so
```
Random 1 in 2 (of 53)

Figure 6.15. Concordances in "Hers"+"Hims": 'Wish'

```
1        you I couldn't spell either  talk to my mom and dad for me tell them this
2        ult as you know.   Please look after Mom as best as you can . Pension
3            ) Why Did Dad Die  on my Birthday Mom Ballored us lethid[?]  huntingto
4                                        Dear mom        by the time you read this
5                                   Dear Mom & Dad        I am sorry to
6    I know its  a bit of a cheeck help my mom & Dad through this  Ive just go
7    sad old git and can't  pull myself up mom dad please don't be sad I'm no
8    t, I'm far from happiness. I'm scared Mom &  Dad, I'm scared of the futur
9            Ddddd My ashes  to go on Mnnns Mom & Dads  Grave Burried  Simpl
10   t in the yellow, ha-ha.   Say hello to mom for me and give her  a big kiss
11               x.  P.S. look after your Mom for me.
12   xx             x  P.S.  look after your Mom for me.
13   lves         Goodbye           Love Mom. God Bless.
14                                   To Mom,     I could never explain all t
15            SEEMED TO BE ENOUGH  YOUR MOM KNOW YOU LOVE HER  VE
16        S THe BeST.    ITS ALL MY FAULT MOM            LOVe U MOM
17    ve allways wanted to die young  well mom + Mttt i love. And i do Lots of
18                                        Mom 'n' Dad,  It is a beautifull day
19       a good time,              Mom,           Nan.
20     ing it was me  who had died and not Mom.  Suicide is not painless, I kno
21   you it back to pick on  Frrr. but why mom why me. Mttttts  the only one
22      uble shared), the same as with your mom.    With our kids growing up &
```
Random 1 in 3 (of 64)

Figure 6.16. Concordances in "Hers"+"Hims": 'Mom'

```
1          and hopefully re-united with  Mom, Dad and Mqqq etc. and always  re
2       life , I love  him I love you Mum and Dad but  shit happens  Fppp xx
3        like I am    (THE Chldren ) Why Did Dad Die  on my Birthday Mom Ballo
4      good to go out with  as you nor your Dad had  any friends
5                            Dear Mom & Dad          I am sorry to have to
6      in my  account given to my Mom & Dad.  I Tried to get my Book put  in
7              FRRR  I LOVE U   I LOVE DAD   I LOVE MSSS & FTTT       P
8         AUNT YOU I WILL BE WITH  YOUR DAD .    I WISH I HAD TIME TO W
9                          Dear Mom and Dad    If you get this then the worst
10                           Mom 'n' Dad,    It is a beautifull day the SUN
11                         MUM / DAD  (ONLY TO BE OPENED IN E
12      ght  life was moving forward. Mom & Dad please don't be sad  youve give
13      ld git and can't  pull myself up mom dad please don't be sad I'm not  Im
14    an make a life for yourself  with your Dad's help.  I wish I could have still
15            TEENAGE YEARS  FROM YOUR DAD SO I THOUGHT YOU  WOUL
16      G ME DOWN    SORRY              DAD
17      ch a  caring family , and a mum and dad who have  always been there fo
18          N   I HAVE GONE  FOR A WALK    DAD       XX
19          N   I HAVE  GONE FOR  A WALK   DAD    XX
20    great partner.  I love you all       Dad xxx
21        All my Love to you all          Dad               xxxx           xxx
22    piness for your future             Dad                  xxxx
23    um for me.       All my love        Dad               XXXX
24    eet again .      All my love,       Dad          xxxx
25       Lots of LOVE MQQQ               DAD
26            UM , DENYING YOU FROM YOUR DAD  YOU DID IT TO ME.  BY THE
27      ITH ALL MY HEART,               DAD
```
Random 1 in 2 (of 59)

Figure 6.17. Concordances in "Hers"+"Hims": 'Dad'

In order to establish whether any of these patterns of usage were unusual, I concordanced the same node words in the Bank of English and compared the resulting collocates. The collocational patterns for the above words in my corpus nearly all appear in broadly similar form in the BoE (in terms of their relative collocation frequency, distance from the node and position to the right or left of it). However, there are a few differences including the following. The Birmingham notes seem to have more instances of the word 'please' collocating with itself; and 'forgive' collocates with the node 'please' in position R1 in my corpus whereas the BoE pattern shows no instances of it there. Indeed there are only 110 instances of 'please forgive' in the entire BoE. In the BoE 'dear' collocates with itself a few times, but in my corpus this phenomenon is so common that 'dear' is the highest placed collocate of 'dear'.

Instead of looking at all the individual key words it would be better to put them into groups having some semantic commonality (as mentioned above). However, I shall temporarily postpone this enterprise until Section 6.3.2 for reasons that will become apparent in the next Section, 6.3.1. Looking at Tables 6.4.a and 6.4.b, however, one can already see that the keywords suggest some semantic sets. For example: one potential

topic group is 'positive affect' which would include the words 'love' and 'dear'; another is 'responsibility' which would include the words 'blame', 'sorry' and 'forgive'; and another is 'future' which could include the words 'hope' and 'wish'. The last two words might equally form a group called 'want'. There is also the potential for a group about 'certainty' which could include the word 'will'. Alternatively, 'will' could be in the group labelled 'future'. Further, there is the potential for a group about 'absoluteness': this would include the word 'all', which (along with 'love') is featured in the word list of every Table so far in this chapter.

For the very top most key words it can be said that there seems to be relatively little difference between the "Hers" and the "Hims". The three first person pronouns in the top five of each list are the 'I', 'my' and 'me' that were encountered in Section 6.2.1. As discussed there, it seems reasonable to suppose that the relatively high frequency of first person pronouns might be indicative of the writers' concern with themselves. One could add 'am' to the list of 'self-referring' words. The pronoun 'you' and the word 'love' might indicate the writers' concern with their addressees, and affection (or lack of it). But what may now be said is that concern with the self, the addressee and love is not only typical of most of the notes in the "Hers" and the "Hims", but is also not typical of texts in general and thus may be an indicator of the suicide note genre.

It could be postulated that the three high frequency words 'love', 'I' and 'you' indicate that many of the notes fall neatly into Durkheim's (2002) three main categories, notes with 'I' being 'Egoistic' (partly characterised by apathy), notes with 'love' being 'Altruistic' (showing concern for others), and notes with 'you' being potentially 'Anomic' (expressing anger and ascribing blame to the addressee) (see Section 3.3). The following examples support this suggestion:

(6.6. Egoistic – I)
"Since Mom died all I've felt is emptiness…" (m95376a);

(6.7. Altruistic – love)
"I send you my fond love." (f96460l);

(6.8. Anomic – you)
"We even went to relate and look at the lies you tood there." (f99123i).

However, such classifications should take account of context, ideally the whole of a note, and classifying according to the presence of a single word would be naïve, as the following counter-examples show.

(6.9. I – not Egoistic)
"OH. I FORGOT HE'S TO BUSY FUCKING MY WIFE." (m96242b);

(6.10. love – not Altruistic)
"2. cause you dont love me and you've just proved it to me." (f95229);

(6.11. you – not Anomic)
"You are such a good person you do not deserve this…" (m99304).

The author of 6.9 appears to be angry, which is not a feature of Egoistic suicides, who tend to be apathetic, but of Anomic ones. 6.10 does not show the altruism that might be indicated by the author's love for someone or something other than herself; and from this extract alone, 6.10 seems to be a mixture of the Egoistic and Anomic. There is no apparent Anomic anger concerning the addressee 'you' in 6.11; on the contrary, given no further context, this could be part of an Altruistic note.

6.3.1. Key Words by Note

One could continue looking at more and more key words to seek differences and similarities between the "Hers" and the "Hims", but there is a problem. Although it can now be seen that certain words appear to be 'key' regarding the genre, there might be bias due to some words being used in only a small number of texts, or by a small number of subjects. This means that the Birmingham suicide notes are not necessarily about any of the things suggested by the key words – other than (because of the relatively very high frequencies of 'I', 'my' and 'me'), perhaps, their authors' selves.

It is important to check whether any of the above words of potential interest is only key due to some biasing factor. With Wordsmith Tools it is possible to see some 'by note' effects, as shown above (see Section 6.2.1), but Wordsmith cannot be used directly to count any 'by note' or 'by subject' effects in its key words lists. In other words, results

that are skewed because they are, or are largely, due to a small number of texts (notes) or subjects (note-writers) cannot easily be accounted for with Wordsmith[35]. However, by using my own software application 'Readgrep' (as mentioned in Section 4.11) to sort and count the concordance-like lines produced by 'grep', taking account of lines from the same text and subject (from the file names), one can see whether any results are skewed. This is shown in Tables 6.5.a and 6.5.b.

Table 6.5.a. Distribution of "Her" Key Words

	Notes	Subjects
I	85.14% (63)	100.00% (22)
MY	68.92% (51)	77.27% (17)
YOU	68.92% (51)	63.64% (14)
ME	63.51% (47)	86.36% (19)
HAVE	60.81% (45)	63.64% (14)
ALL	48.65% (36)	50.00% (11)
LOVE	37.84% (28)	59.09% (13)
AM	33.78% (25)	54.55% (12)
DEAR	29.73% (22)	50.00% (11)
SORRY	25.68% (19)	50.00% (11)
LIFE	24.32% (18)	50.00% (11)
PLEASE	24.32% (18)	45.45% (10)
THANK	22.97% (17)	27.27% (6)
LOVED	18.92% (14)	22.73% (5)
BLAME	14.86% (11)	18.18% (4)
X	14.86% (11)	9.09% (2)
DONT	9.46% (7)	27.27% (6)
MOM	9.46% (7)	22.73% (5)
GRIEVE	8.11% (6)	18.18% (4)
XX	6.76% (5)	13.64% (3)
SLAG	6.76% (5)	4.55% (1)
RD	5.41% (4)	18.18% (4)
OVERWELM	5.41% (4)	4.55% (1)
B'HAM	4.05% (3)	13.64% (3)
TRYED	4.05% (3)	9.09% (2)

---

[35] One could utilise Wordsmith's concordance line facility which includes the source filename on each line, but it does not facilitate any automatic counting based on parts of filenames.

Table 6.5.b. Distribution of "Him" Key Words

	Notes	Subjects
I	83.02% (176)	88.04% (81)
YOU	69.34% (147)	73.91% (68)
ME	59.91% (127)	72.83% (67)
MY	59.43% (126)	69.57% (64)
LOVE	56.13% (119)	71.74% (66)
HAVE	55.19% (117)	65.22% (60)
ALL	52.83% (112)	68.48% (63)
WILL	45.75% (97)	61.96% (57)
SORRY	41.51% (88)	63.04% (58)
AM	38.21% (81)	47.83% (44)
PLEASE	35.85% (76)	47.83% (44)
KNOW	33.96% (72)	44.57% (41)
LIFE	33.49% (71)	48.91% (45)
YOUR	32.08% (68)	40.22% (37)
DEAR	26.89% (57)	32.61% (30)
HOPE	24.53% (52)	40.22% (37)
DAD	20.28% (43)	29.35% (27)
MOM	18.40% (39)	25.00% (23)
WISH	13.68% (29)	20.65% (19)
FORGIVE	13.21% (28)	26.09% (24)
DONT	12.74% (27)	23.91% (22)
IM	12.74% (27)	19.57% (18)
IVE	8.49% (18)	11.96% (11)
CANT	8.02% (17)	16.30% (15)
BYE	7.08% (15)	11.96% (11)

Comparing Tables 6.4.a and 6.5.a (above), and looking at words featuring in fewer than 9% of the texts, shows that the words 'grieve', 'XX', 'slag', 'Rd', 'overwelm' [sic], 'B'ham' and 'tryed' [sic], that were key words for the "Hers" as a whole (see Table 6.4.a) are actually spread across very few texts, and also very few subjects. Even 'blame' and single kisses ('X'), each in 14.86% of the texts, are used by only 4 and 2 subjects respectively. Similarly, comparing Tables 6.4.b and 6.5.b (above) shows that the words 'Ive', 'cant' and 'bye' that were key words for the "Hims" as a whole (see Table 6.4.b) are each spread across fewer than 9% of texts and 17% of subjects in the "Hims".

Thus, the only key words from the previous section that are not well distributed by text, excluding the mis-spelt words, parts of addresses and kisses ('XX'), are 'slag' in the "Hers" and 'bye' in the "Hims".

It transpired that the word 'slag' was a subject-skewed result: it occurred 13 times, and as a noun, in the "Hers", but only in 5 texts, and all of these texts were written by the same subject. It also occurred once in the "Hims", but that was as a verb ("YOU SLAG ME OFF TO ALL YOUR FRIENDS AND FAMILY…" (m96318d)).

'Bye' has already been mentioned in Section 5.3, where it was noted that over 20% of Shneidman's genuine suicide notes said 'goodbye' in some way that principally involved the words 'goodbye' or 'bye'. As Table 6.5.b shows, nearly 12% of the Birmingham male note-writers used the word 'bye', but this was in only about 7% of the ("Him") notes. There were uses of 'bye' by 13.64% of the Birmingham females, but this percentage represents only 3 subjects, and the uses occurred in only 6.76% (5) of the ("Her") texts.

Having the above lists of key words has satisfied the need to have lists of words that take into account, and are ordered by, their distribution among the texts, and at the same time are indicative of what is peculiar to the sub-corpora, at least in comparison to the BNC. However, there remained the issue of different words with related meanings that could be grouped together – which I shall now address.

6.3.2. Key Word Groups

As already seen, the 'grep' line-reading application can show not only the distribution between texts and subjects of words, but also of groups of words. Because I was moving towards a sense of semantic sets, here I wanted to group together some of the words from the "Her" and "Him" key word lists and put them into categories. The categories were based on themes that seem to suggest themselves when examining the lists; they contained words that seemed to go together in terms of aboutness. Of course there are many different groupings that could be made, but I settled upon the ones shown below as they seemed most salient.

It should be noted that combining the words from the "Hers" and the "Hims" into a single set of categories allowed for the possibility that a word would be included that was key in one sub-corpus but not key in the other. But since I was now seeking

semantic concepts this was acceptable on the basis that the concept itself might be key, and the accumulation of individually less frequent words might contribute to the overall frequency. I decided to include mis-spellings because, for the words considered here, it seemed reasonable to assume what the note-writer had intended: for example, 'I am' instead of 'im'. I also included the only word common to both sub-corpora in their next half dozen keywords (that is, keywords 26-31, excluding the anonymised lexis), namely, 'want'. Having combined the words in Tables 6.5.a and 6.5.b I created the following categories (whose constituent words are shown in parentheses).

Self (i, my, me, am, im, ive)

Others (you, your)

Positive emotion (dear, love, loved)

Responsibility (blame, sorry, forgive)

Want (want, wish, hope)

Parents (mom, dad)

Politeness (please, thank)

The results are shown in Table 6.6.a in which the percentages are given, followed by the raw counts in parentheses. The table shows, for example, that in the "Hers" there were altogether 14 instances of the 'parent' words 'mom' and 'dad', and that 10 of the suicide notes contained at least one occurrence of either the word 'mom' or the word 'dad', but that only 6 of the note-writers had used these words. In the "Hims", however, there were 109 instances of the two words spread over 29.72% of the texts, and 42.39% of the note-writers had used at least one of the words in at least one of their texts.

As mentioned before the software in question utilises sub-corpus token counts acquired from an older version of Wordsmith, and any differences in the percentage figures for overall occurrence between the tables produced by this software (such as Table 6.6.a below) and tables produced by Wordsmith 4 (such as Tables 6.1.a to 6.4.b above) are due to the different token counts calculated by Wordsmith 3 and Wordsmith 4 (see Section 4.11).

Table 6.6.a. Word Groups from Top Key Words

Corpus	Occurrences	Notes	Subjects
Self			
her	8.14% (600)	94.59% (70)	100.00% (22)
him	9.36% (2813)	88.21% (187)	93.48% (86)
Others			
her	3.76% (277)	70.27% (52)	68.18% (15)
him	3.86% (1159)	70.28% (149)	73.91% (68)
Positive Emotion			
her	1.15% (85)	56.76% (42)	77.27% (17)
him	1.15% (347)	64.62% (137)	76.09% (70)
Responsibility			
her	0.60% (44)	44.59% (33)	68.18% (15)
him	0.58% (175)	47.17% (100)	68.48% (63)
Want			
her	0.62% (46)	37.84% (28)	45.45% (10)
him	0.69% (207)	39.62% (84)	55.43% (51)
Parents			
her	0.19% (14)	13.51% (10)	27.27% (6)
him	0.36% (109)	29.72% (63)	42.39% (39)
Politeness			
her	0.62% (46)	37.84% (28)	45.45% (10)
him	0.47% (141)	41.51% (88)	52.17% (48)

Taking Table 6.6.a at face value it appears that suicide note-writers are more concerned with themselves than others; that their texts contain more expressions of positive emotions than terms expressing responsibility; that over a third of the texts, and about half the subjects, appear to want something; that over a third of the texts, and nearly half the subjects, use terms of politeness; and that quite a few of them, particularly the males, reference parents. All of this has already been discovered by other researchers (see Chapter 3 and Section 7.1.1 below); however, it is interesting to be able to corroborate their findings by using such simply composed categories.

I next looked at the whole keyword list to see what other words could be added to the above categories. The list below shows the revised version.

Self (i, my, me, am, im, ive, i'm, ime, myself, i've, self)

Addressee (you, your, youve)

Others (shes, he, his, thay, their, they)

Positive emotion (dear, love, loved, loving, friends, friendship, kindness, bless,

happiness, happy)

Responsibility (blame, sorry, forgive, apologise, fault)

Want (want, wanted, wan't, whont)

Parents (mom, dad, mum)

Politeness (please, thank, thanks)

Absoluteness (all, everyone, everything, nobody)

Future (wish, hope)

Certainty (will, never, allway, allways, always, ever, forever)

Bad (lied, lies, sad, slag, unhappiness, fucked, hate, mess, trouble, overwelm)

Illness (pain, ambulanc, ambulane, dignaty, incontinent, pills, hurt)

Death (depart, die, grieve, cremated, funeral, suffocate)

Negatives (dont, not, nor, cannot, cant, can't, carn't, don't, havent, no, wont)

Most of the categories are now expanded, and 'Others' is split into 'Addressee' and 'Others'. 'Others' includes the only 'negative keywords' (words which are less rather than more frequent compared to a reference corpus - see Section 6.3.4) used in these categories: namely, 'he', 'his', 'their' and 'they' (from the "Hims"). I added the categories 'Absoluteness', 'Future' and 'Certainty' as posited in Section 6.3. To keep the categories mutually exclusive the words 'hope' and 'wish' in the 'Future' category were removed from the 'Want' one. Some totally new categories also emerged, namely, 'Bad', 'Illness', 'Death' and 'Negatives'. Although there were few clues as to the existence of these categories in the top keywords lists, words relating to them were highly salient when examining the entire list.

Although there is great similarity in the usage of the top few key words by the "Hers" and the "Hims", there are differences lower down the list. The "Hers" and "Hims" share 33 positive keywords (and 6 negative keywords), but there are 28 keywords (all positive) that are unique to the "Hers", and 106 positive (and 12 negative) keywords unique to the "Hims". One shared word not yet mentioned in this thesis is 'pain' in the category 'Illness'. Other shared words I have not categorised (or mentioned before) include 'good' and 'letter'. Words in the above categories that are keywords only in the

"Hers" include 'die', 'friends' and 'friendship', 'lies' and 'lied', and 'you'. Keywords unique to the "Hers", but which I have not allocated to any category, include 'believe', 'pension' and 'sincerely'. In the categories, and unique to the "Hims", are keywords including 'always', 'apologise', 'cremated', 'ever', 'fault', 'happy', 'hate', 'pills', 'trouble' and 'wish'. Some unique "Him" keywords not in the categories are 'ask', 'feel', 'forget', 'money' and 'things'.

Only "Hers" keywords were found for the 'Addressee' category, and (apart from 'will') only "Him" keywords were found for the 'Certainty' category. All the other categories were represented in both sub-corpora. Here are some examples of the above categories from the data (with the relevant words in bold).

(6.12. Self – 'my' & 'I')
"Throughout all **my** troubles **I** have always had good reliable friends…" (m95148c);

(6.13. Addressee – 'you' & 'your')
"**You** can take **your** time to sell the house…" (f96369e);

(6.14. Others – 'he' & 'his')
"WHAT WILL **HE** SAY WHEN PEOPLE ASK ABOUT **HIS** DAD" (m96318c);

(6.15. Positive emotion - 'dear' & 'friends')
"**Dear** Family and **Friends**…" (f98279b);

(6.16. Responsibility – 'sorry' & 'forgive')
"So **sorry** hope you can **forgive** me." (m96488e);

(6.17. Want – 'want')
"All I **want** is a life, im now crying for aparent reason…" (f95456a);

(6.18. Parents – 'mom' & 'dad')
"I'm scared **Mom** & **Dad**, I'm scared of the future" (m99223a);

(6.19. Politeness – 'please' & 'thank')
"**Please thank** all our friends." (f97158a);

(6.20. Absoluteness – 'all')
"Yes mate, you have achieved absolutely bugger **all**." (m97157a);

(6.21. Future – 'hope')
"Well I **hope** you both will be very happy…" (f99123i);

(6.22. Certainty – 'will' & 'always')
"I **will always** be watching over you, your loving brother." (m95148c);

(6.23. Bad – 'hate')
"I **hate** you as much as you do me." (f96049a);

(6.24. Illness – 'pills')
"HAVING WRITEN THIS, THE **PILLS** ARE NOW WORKING, I GO…" (m99438a);

(6.25. Death – 'die')
"It breaks my heart to do this, I   thought I would **die** in the comfort of your arms…" (f95302);

(6.26. Negatives – 'don't' & 'cannot')
"I **don't** want children because I **cannot** see a good future anywhere in the world for them…" (m97157a).

Checking against the full word frequency (not keyword) lists shows that many of the words peculiar to one sub-corpus in terms of keyness are present in the other sub-corpus although they are not keywords in it. However, a few words are indeed unique to each of the sub-corpora. 'Depart' and 'suffocate' (in the 'Death' category) occur only in the "Hers", and not in the "Hims", as do 'unhappiness' and 'overwelm'[sic] (in the 'Bad' category). 'Forever' and 'fucked' in the categories 'Certainty' and 'Bad', respectively, and 'incontinent' and 'pills' in the 'Illness' category occur only in the "Hims", and not in the "Hers". Interestingly, 'and' and 'but' are only keywords in the "Hims": 'but' is a positive keyword, whereas 'and' is a negative keyword.

Table 6.6.b shows the results of running 'grep' and 'Readgrep' on the corpus for the revised categories.

Table 6.6.b. Word Groups from All Key Words

Corpus	Occurrences	Notes	Subjects
Self			
her	8.45% (623)	94.59% (70)	100.00% (22)
him	9.92% (2982)	88.21% (187)	93.48% (86)
Addressee			
her	3.80% (280)	70.27% (52)	68.18% (15)
him	3.88% (1166)	70.28% (149)	73.91% (68)
Others			
her	0.69% (51)	27.03% (20)	31.82% (7)
him	0.54% (161)	28.77% (61)	40.22% (37)
Positive emotion			
her	1.86% (137)	63.51% (47)	86.36% (19)
him	1.46% (440)	66.04% (140)	77.17% (71)
Responsibility			
her	0.65% (48)	45.95% (34)	68.18% (15)
him	0.65% (195)	49.06% (104)	69.57% (64)
Want			
her	0.45% (33)	25.68% (19)	40.91% (9)
him	0.42% (127)	25.94% (55)	36.96% (34)
Parents			
her	0.22% (16)	14.86% (11)	31.82% (7)
him	0.43% (128)	30.19% (64)	43.48% (40)
Politeness			
her	0.75% (55)	43.24% (32)	45.45% (10)
him	0.52% (156)	44.34% (94)	54.35% (50)
Absoluteness			
her	1.11% (82)	55.41% (41)	59.09% (13)
him	1.17% (352)	55.19% (117)	71.74% (66)
Future			
her	0.26% (19)	20.27% (15)	31.82% (7)
him	0.41% (123)	31.60% (67)	45.65% (42)
Certainty			
her	1.11% (82)	58.11% (43)	59.09% (13)
him	1.41% (423)	55.66% (118)	67.39% (62)
Bad			
her	0.68% (50)	28.38% (21)	36.36% (8)
him	0.23% (70)	20.75% (44)	31.52% (29)
Illness			
her	0.16% (12)	14.86% (11)	36.36% (8)
him	0.19% (58)	16.98% (36)	28.26% (26)
Death			
her	0.39% (29)	27.03% (20)	36.36% (8)
him	0.15% (45)	10.85% (23)	17.39% (16)
Negatives			
her	2.22% (164)	75.68% (56)	86.36% (19)
him	2.12% (636)	68.87% (146)	81.52% (75)

'Self' now has more occurrences, but is only present in the same number of texts as in the earlier category termed 'Self'. There is much more referencing of the 'Addressee' than of 'Others'. The revised category of 'Positive emotion' shows a slightly increased frequency. The figures for 'Responsibility', 'Parents' and 'Politeness' have also only increased slightly. 'Want' now occurs in about a quarter of the notes, and over a third of the subjects. 'Future'-related words are present in about a fifth of the female notes, and nearly a third of the male ones. Approximately 31% of the female and 45% of the male subjects used words in the 'Future' category. Over half of the notes include references either to 'Absoluteness' or to 'Certainty', or both. Nearly three quarters of the male subjects used words of 'Absoluteness', and over two thirds of the males used words of 'Certainty'. Over half of the female subjects used references to 'Absoluteness' and/or 'Certainty'.

Over a third of the female subjects used words in the categories 'Illness', 'Death' and 'Bad', whereas the males used more words related to 'Bad'ness than 'Illness', but more words related to 'Illness' than 'Death'. There were also fewer male than female subjects referring to any of these three categories. There are similar figures for male and female texts referring to 'Illness'. However, more texts written by females than by males were in the 'Bad' category, and far more texts written by females were in the 'Death' category. 'Negatives' are used in about 70% of all texts, and by over 80% of all subjects.

Because I was increasingly concerned with semantic sets, I wanted to follow up the concept stemming from one of the top most key words that was common to both the "Hers" and the "Hims", which was lexical, spelt correctly, and had a relatively low frequency of occurrence in both sub-corpora, and to see how much difference it made to the figures when some synonyms were added. The word 'mom' fitted these criteria, and accordingly I examine it in the next section.

6.3.3. Mother Terms
One might wonder why 'Mom', the supposed American English equivalent of 'Mum' meaning Mother, appears at all in British suicide notes. The expected usage of 'Mom'

as opposed to 'Mum' is clearly shown in the Bank of English (BoE): the British books have a far greater tendency to use 'Mum' (1,645 occurrences) rather than 'Mom' (350 occurrences), whereas the U.S. books tend to use 'Mom' (1,216 occurrences). In fact, the U.S. books had a total of 58 occurrences of the word 'Mum' and, as the BoE concordance lines show, only 49 of these meant Mother, the remaining 9 meaning 'silent'. The appearance of 'Mom' in my Birmingham corpus is in fact explained by the fact that it is a dialect feature of Birmingham English (Carl Chinn, personal communication).

Aside from the comparative usage of the words 'Mom' and 'Mum', there is the matter of the overall use of terms which refer to mothers. I wanted to investigate a little further the potential clue given by the key word 'Mom', and decided to look at how many of the note-writers mention mothers. Osgood and Walker (1959) had noticed that suicide note-writers often seemed to use 'mother' terms (see Chapter 3). A perusal of the word frequency lists (via Wordsmith) revealed that only three basic 'mother' terms, (along with their inflected forms) existed in the "Hers" and "Hims". These were 'mother', 'mom' and 'mum' (there were no instances of, for example, 'mam' or 'mater').[36] So the mother terms I looked for were largely covered by the lemmas 'mother', 'mom' and 'mum'. Performing wildcard searches allowed for retrieval of words such as 'mothers', 'moms' and 'mummy'. Mis-spellings were also included in the tallies. The words actually found in this way were 'mom', 'mum', 'mummys' (assumed to be a possessive form without the apostrophe), 'mumy' and 'mother' in the "Hers", and 'mom', 'moms', 'mommy', 'mum', 'mother' and 'mothers' in the "Hims". Table 6.7 shows the distribution of for the three maternal lemmas (including all their variant forms) in the "Hims" and the "Hers". (The last two rows of Table 6.7 show the figures for the three lemmas combined by the use of logical 'OR' in 'Readgrep'. The columns show (from left to right): the total occurrences of the lexical items; the total number of notes with one or more instances of any of the target patterns; and the total number of subjects with one or more instances of any of the target patterns.)

---

[36] There was one instance of 'ma' in the "Hims" which was overlooked and would have added one note (and one subject) to the results.

Table 6.7. Mother Terms

```
    Corpus    Occurrences      Notes          Subjects
mother
    her      0.03% (  2)    2.70% (  2)     9.09% (  2)
    him      0.06% ( 17)    6.13% ( 13)    13.04% ( 12)


mom
    her      0.12% (  9)    9.46% (  7)    22.73% (  5)
    him      0.19% ( 58)   19.34% ( 41)   27.17% ( 25)


mum
    her      0.05% (  4)    5.41% (  4)    13.64% (  3)
    him      0.06% ( 19)    5.19% ( 11)   10.87% ( 10)


mother|mom|mum*
    her      0.20% ( 15)   17.57% ( 13)   31.82% (  7)
    him      0.31% ( 94)   28.77% ( 61)   45.65% ( 42)


* Logically 'OR'd
```

It should be remembered that because these categories or collections of words were made using wildcards rather than whole words the figures for 'mom' in Table 6.7 are different from those for the word 'mom' in, for example, Tables 6.5.a and 6.5.b above. For comparison, here are some figures for the (case insensitive) whole word "mom".

```
    Corpus    Occurrences      Notes          Subjects
Mom (whole word)
    her      0.12% (  9)    9.46% (  7)    22.73% (  5)
    him      0.18% ( 55)   18.40% ( 39)   25.00% ( 23)
```

Returning to the distribution shown in Table 6.7, some interesting patterns emerge. In the "Her" corpus there were more than twice as many 'Mom's as 'Mum's, although the former were used in only 7 notes, and by 5 subjects (and the latter in only 4 notes by 3 subjects). The "Him" corpus also had more usage of 'Mom' as opposed to 'Mum', about three times as much. In these two sub-corpora there did not seem to be any great differences between the males and the females regarding these preferences, except for the "Hims" using 'Mom' in nearly 10% more of their notes. What is interesting is that over 31% of the "Her", and 45% of the "Him" subjects used mother terms in their texts.

It should be borne in mind, however, that this discussion of 'Mother' terms does not make any distinction as to *whose* mother is being referred to, or whether the reference is in a positive or negative context.

In about half of the "Hers", forms of 'mother' were used as signatures. The "Hers" also contained 3 instances of mother terms used in direct addressing within the body of the text, rather than as the addressee at the beginning. There was also one reference to someone else's mother, one use of 'mother' in a quotation, one derogatory descriptive use "mummys boy", and one about not applying the term to the addressee. Here are some extracts from the "Hers" illustrating each of these types of reference.

(6.27. Mother term as signature)
"Love Mom." (f98478);

(6.28. Text body addressing)
"but why mom why me." [sic] (f95229);

(6.29. Another's mother)
"Mary was a good dauter[sic] to her mother…" (f96049d);

(6.30. Mother in quotation)
"say, 'Found my Mother in bed…'" (f96369b);

(6.31. Derogatory descriptive)
"you are a mummys boy…" [sic] (f96049a);

(6.32. Rejection of mother term)
"I wont call you mom as you didn't want me to…" [sic] (f95229).

The instances in the "Hims" seem to consist of similar contexts except, of course, there are none used in signatures. This makes the references to 'mother's even more significant in the "Hims". Most "Hims" use the mother terms for the addressee or in direct addressing (see above); several are references to other people's mothers.

One might ask what knowing that 'mothers' are mentioned in 17.57% of the notes authored by females, and 28.77% of the notes authored by males, tells us about suicide notes; or one might ask about the even greater numbers of notes that do *not* mention mothers. The answers seem to be that in finding sets of words that account for a

'reasonable amount' of the total, one could be en route to having a picture of what suicide notes are about, and that no single semantic set of words can give the whole picture. For example, finding 10 or 20 categories that each tended to be represented in, say, 95% of the texts of a corpus would be nice. But finding 100 categories that were present in 100% of the texts would not be very helpful especially if each category made up only 1% of the texts. Thus, finding figures of over 17% and 28% for the texts in the two sub-corpora could be an important piece of the picture of what suicide notes are about (as could finding the figures of over 31% and 45% for the subjects, as mentioned above).

6.3.4. Negative Key Words

For the sake of completeness I now take a very brief diversion and look at negative key words. Negative key words are words in one text, or a set of texts, that are notable for their *low* frequency when compared with a reference corpus. (These are the opposite of the positive keywords described above, whose high frequency compared to a reference corpus made them 'key'.) Tables 6.8.a and 6.8.b show the negative key words for the "Hers" and "Hims" respectively. There are three times as many negative key words in the "Hims" than in the "Hers", but this could be due to the relative sizes of the Birmingham sub-corpora – the "Hers" being smaller than the "Hims" in terms of the total number of words.[37]

Table 6.8.a. Negative Key Words in "Hers"

N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P
1	THE	193	2.56	6,055,105	6.09	-207.60	0.0000000000
2	OF	100	1.33	3,049,564	3.07	-97.15	0.0000000000
3	WAS	20	0.27	863,917	0.87	-43.83	0.0000000000
4	A	97	1.29	2,181,592	2.19	-33.92	0.0000000028
5	BY	10	0.13	513,444	0.52	-30.79	0.0000000258
6	IN	87	1.15	1,946,021	1.96	-29.68	0.0000000480

---

[37] Of course, all the words listed in tables 6.8.a and 6.8.b do still occur in my corpus, even if only with a frequency of 1. It might be pointed out that words which do not appear at all in my corpus are even more "negative" if the comparison corpus contains them in a high frequency.

Table 6.8.b. Negative Key Words in "Hims"

N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P
1	THE	912	2.99	6,055,105	6.09	-625.40	0.0000000000
2	OF	434	1.42	3,049,564	3.07	-345.34	0.0000000000
3	IN	325	1.06	1,946,021	1.96	-151.48	0.0000000000
4	HIS	19	0.06	410,294	0.41	-142.36	0.0000000000
5	A	403	1.32	2,181,592	2.19	-126.33	0.0000000000
6	BY	46	0.15	513,444	0.52	-110.30	0.0000000000
7	HE	61	0.20	593,609	0.60	-109.37	0.0000000000
8	WHICH	28	0.09	366,196	0.37	-91.19	0.0000000000
9	THEIR	14	0.05	254,603	0.26	-80.27	0.0000000000
10	AN	34	0.11	338,743	0.34	-64.09	0.0000000000
11	WERE	35	0.11	308,363	0.31	-49.77	0.0000000000
12	THEY	50	0.16	376,289	0.38	-47.41	0.0000000000
13	WHERE	4	0.01	103,471	0.10	-38.96	0.0000000000
14	WAS	173	0.57	863,917	0.87	-36.85	0.0000000000
15	NEW	9	0.03	124,409	0.13	-32.38	0.0000000097
16	BETWEEN	4	0.01	90,620	0.09	-32.13	0.0000000115
17	AND	657	2.15	2,624,341	2.64	-29.98	0.0000000407
18	SAID	23	0.08	195,580	0.20	-29.97	0.0000000409

A couple of interesting observations can be made from Tables 6.8.a and 6.8.b. The first couple of words in each table are the words most indicative of written, as opposed to spoken, texts (see Section 6.2 above). The inclusion of 'was', and also of 'were' in the "Hims", suggests reference to the present (rather than the past) because they are both negative key words (although there is no 'is' in the positive keywords either (see Section 6.3)). Further, the 'said' in Table 6.8.b indicates that there may be relatively little reported speech in the "Hims". However, using 'grep' (and 'Readgrep') on the words 'said', 'saying', 'told' and 'telling' gives figures for usage 'by note' of 12.16% and 16.04% for the "Hers" and "Hims", respectively, and also shows that 22.73% of the "Her" and 25.00% of the "Him" subjects used at least one of these words.

These words were worth checking just in case any peculiarities were present. For example, one might suppose that words such as 'funeral' and 'death' would be frequently used in suicide notes. As has been seen, however, they are not. But had such words appeared in a negative key word list that might suggest something quite

unexpected. It could suggest that suicides go to some lengths not to talk about their suicide. However, Tables 6.8.a and 6.8.b seem to show nothing outstanding.

6.4. Most Frequent Clusters

I next take a brief look at the most frequent clusters of four, and of three, words, in case any unusual groups exist in the corpus. A "cluster" is simply a multi-word unit, not necessarily a grammatically well-formed phrase or clause. This work was carried out with the aid of Wordsmith 4 (Scott, 2007a), but I also used Wordsmith 3 (Scott, 1998a) to check it (because of the problems with the former (see Section 4.11)). Wordsmith has a feature which enables the user to generate automatically a list of all the clusters which occur more than once in a text. To compare the frequencies of the various groups of words against a general corpus I used the Bank of English (BoE). Tables 6.9.a to 6.10.b show the results.

I have shown the clusters in descending frequency order, cutting off some of the lower frequency clusters for manageability of the output. I decided to limit the number of clusters (per table) to a maximum of about 25. Each table contains enough lines to fulfil this limit, and also includes all the lines which show the lowest frequency in it. For example, Table 6.9.a ends at line 23 and includes all the clusters occurring 4 times, but lines 24 and 25 are not shown because they are 2 of 34 lines of 4-word clusters each occurring 3 times in the "Hers". Thus, the different numbers of word clusters shown in each table reflect appropriate cut-off points for a brief analysis.

In Tables 6.9.a to 6.10.b the BoE figures are raw numbers of occurrences. It should be noted, however, that unlike the Birmingham corpus where an "X" represents a kiss, in the BoE most of the "X"s seem to represent space, perhaps tabulation characters in tables. (In Table 6.9.a, for example, 185 of the 204 "X X X X"s are in the U.S. academic sub-corpus.)

Table 6.9.a. Top 4-Word Clusters: "Hers"

N	Cluster	Freq.	%	Texts	%	BoE
1	X X X X	29	0.38	1	1.35	204
2	# # # #	7	0.09	3	4.05	-
3	I SHALL NO LONGER	6	0.08	6	8.11	4
4	MAY I THANK YOU	6	0.08	6	8.11	5
5	SHALL NO LONGER BE	6	0.08	6	8.11	11
6	HAVE DONE FOR ME	5	0.07	5	6.76	19
7	LETTER I SHALL NO	5	0.07	5	6.76	0
8	THIS LETTER I SHALL	5	0.07	5	6.76	2
9	GOOD THINGS IN LIFE	4	0.05	4	5.41	91
10	I THANK YOU MOST	4	0.05	4	5.41	4
11	IN LIFE AND SUFFOCATE	4	0.05	4	5.41	0
12	LIFE AND SUFFOCATE IT	4	0.05	4	5.41	0
13	NO LONGER BE HERE	4	0.05	4	5.41	2
14	ON THE OTHER SIDE	4	0.05	3	4.05	4921
15	RECEIVE THIS LETTER I	4	0.05	4	5.41	1
16	THANK YOU MOST SINCERELY	4	0.05	4	5.41	1
17	THE GOOD THINGS IN	4	0.05	4	5.41	93
18	THINGS IN LIFE AND	4	0.05	4	5.41	44
19	WHEN YOU RECEIVE THIS	4	0.05	4	5.41	6
20	WITH A LOVING THOUGHT	4	0.05	4	5.41	0
21	YOU HAVE DONE FOR	4	0.05	4	5.41	44
22	YOU MOST SINCERELY FOR	4	0.05	4	5.41	1
23	YOU RECEIVE THIS LETTER	4	0.05	4	5.41	6

Table 6.9.b. Top 4-Word Clusters: "Hims"

N	Cluster	Freq.	%	Texts	%	BoE
1	# # # #	20	0.07	11	5.19	-
2	I LOVE YOU ALL	11	0.04	9	4.25	53
3	I AM SO SORRY	9	0.03	9	4.25	84
4	WHAT I HAVE DONE	9	0.03	8	3.77	201
5	X X X X	8	0.03	2	0.94	204
6	AS SOON AS POSSIBLE	7	0.02	7	3.30	3943
7	I DO NOT KNOW	7	0.02	6	2.83	1188
8	I HOPE YOU WILL	7	0.02	7	3.30	318
9	I LOVE YOU AND*	7	0.02	7	3.30	188
10	BUT I DO NOT	6	0.02	6	2.83	584
11	I HOPE YOU CAN	6	0.02	6	2.83	126
12	I STILL LOVE YOU	6	0.02	5	2.36	36
13	I WILL ALWAYS LOVE	6	0.02	6	2.83	86
14	THE LAST # MONTHS	6	0.02	6	2.83	4082
15	TUES # JAN #	6	0.02	6	2.83	0

* In Wordsmith 3 this line was "PTO --", & "I LOVE YOU AND"
  was on line 23.

At first sight there are no outstanding clusters. For some, such as "on the other side"
(line 14) in Table 6.9.a, "as soon as possible" (line 6), "I do not know" (line 7) and "the
last # months" (line 14) in Table 6.9.b, this is not surprising. They are not infrequent in
the Bank of English. However, looking at the BoE figures shows that some clusters
occur very rarely in the context of a general corpus although they appear to be

commonplace phrases to a native English speaker. "May I thank you" (line 4), "thank you most sincerely" (line 16) and "when you receive this" (line 19) in Table 6.9.a occur only 5, 1 and 6 times, respectively, in the whole of the BoE. Of course, a corpus of personal letters might contain more of these items than the BoE does, but time restraints did not permit such enquiries.

"I will always love" (line 13) in Table 6.9.b occurs 86 times in the BoE, but many of these occurrences are references to a popular song (namely, "I will always love you" by Dolly Parton, covered by Whitney Houston). Excluding the song references makes it even less frequent. However, despite recourse to the Bank of English, the cluster does not strike one as being particularly rare. Another cluster that does not 'sound' unusual is, "shall no longer be" (line 5) in Table 6.9.a which tends to be used in formal or official texts in the BoE, and occurs only 11 times there.

Another observation is that there are some clusters that are related. For example, in Table 6.9.a, "good things in life" (line 9) and "the good things in" (line 17) when input to the BoE as a five-word cluster "the good things in life" occurs 78 times. Concordances of "things in life and" (line 18) in the BoE show that most occurrences have a word with positive affect as the first word to the left of this cluster. Indeed, lines 9, 17 and 18, along with "in life and suffocate" (line 11) and "life and suffocate it" (line 12) all come from several texts written by one subject. One version of a sentence in which they are used is:

> (6.33. A multi-note writer's contribution to clusters)
> "Unhappiness can overwelm [sic] all the good things in life and suffocate it."
> (f96320h).

I shall now consider the 3-word clusters shown in Tables 6.10.a and 6.10.b.

Table 6.10.a. Top 3-Word Clusters: "Hers"

N	Cluster	Freq.	%	Texts	%	BoE
1	X X X	33	0.44	3	4.05	344
2	# # #	23	0.30	9	12.16	–
3	I LOVE YOU	9	0.12	7	9.46	2211
4	I THANK YOU	9	0.12	8	10.81	345
5	NO LONGER BE	7	0.09	7	9.46	3011
6	THERE IS NO	7	0.09	7	9.46	45720
7	BE WITH YOU	6	0.08	6	8.11	683
8	HAVE DONE FOR	6	0.08	6	8.11	409
9	I AM NOT	6	0.08	5	6.76	9157
10	I SHALL NO	6	0.08	6	8.11	9
11	MAY I THANK	6	0.08	6	8.11	8
12	SHALL NO LONGER	6	0.08	6	8.11	23
13	THIS LETTER I	6	0.08	6	8.11	54
14	A LOT OF	5	0.07	5	6.76	89748
15	AND ALL THE	5	0.07	5	6.76	8038
16	DONE FOR ME	5	0.07	5	6.76	145
17	DON'T BE SAD	5	0.07	5	6.76	3
18	FQQQ X X	5	0.07	5	6.76	–
19	I AM SORRY	5	0.07	3	4.05	789
20	LETTER I SHALL	5	0.07	5	6.76	4
21	OVER THE YEARS	5	0.07	5	6.76	9044
22	PLEASE FORGIVE ME	5	0.07	5	6.76	79
23	THE OTHER SIDE	5	0.07	3	4.05	9173
24	TO YOU AND	5	0.07	5	6.76	2116
25	WANT TO BE	5	0.07	5	6.76	15189
26	YOU HAVE DONE	5	0.07	5	6.76	886
27	YOUR LOVED ONES	5	0.07	5	6.76	133

Table 6.10.b. Top 3-Word Clusters: "Hims"

N	Cluster	Freq.	%	Texts	%	BoE
1	I LOVE YOU	63	0.21	47	22.17	2211
2	# # #	53	0.17	27	12.74	–
3	ALL MY LOVE	24	0.08	22	10.38	102
4	I DO NOT	23	0.08	19	8.96	8271
5	I HOPE YOU	22	0.07	18	8.49	2213
6	I HAVE BEEN	18	0.06	14	6.60	9692
7	I CAN NOT	16	0.05	9	4.25	89
8	I AM SO	15	0.05	14	6.60	1182
9	I AM SORRY	15	0.05	15	7.08	789
10	I WOULD LIKE	15	0.05	12	5.66	7812
11	I KNOW YOU	14	0.05	12	5.66	2851
12	I WOULD HAVE	14	0.05	13	6.13	10444
13	LOVE YOU ALL	14	0.05	11	5.19	72
14	I HAVE DONE	13	0.04	11	5.19	1478
15	MY LOVE TO	13	0.04	11	5.19	75
16	PLEASE FORGIVE ME	13	0.04	12	5.66	79
17	I DID NOT	12	0.04	11	5.19	5572
18	I HAVE TO	12	0.04	9	4.25	14357
19	HOPE YOU CAN	11	0.04	11	5.19	191
20	I DON'T WANT	11	0.04	10	4.72	18
21	I HAVE NOT	11	0.04	9	4.25	2685
22	LOVE YOU AND*	11	0.04	11	5.19	303
23	TO GO TO	11	0.04	9	4.25	20155

* In Wordsmith 3 this line was "TO GO TO", & "LOVE YOU AND"
  was on line 36.

The picture given by Tables 6.10.a and 6.10.b is very similar to that given by Tables 6.9.a and 6.9.b: some clusters are not unusual, but a few are. "I love you" (lines 9 and 1) in Table 6.10.a and 6.10.b), "there is no" (line 6), "I am not" (line 9), "a lot of" (line 14), "over the years" (line 21) and "want to be" (line 25), for example, in Table 6.10.a, and "I do not (line 4), "I hope you" (line 5), "I would have" (line 12) and "I have to" (line18), for example, in Table 6.10.b are not infrequent in the Bank of English.

However, the clusters which are not so frequent in the BoE compared to the Birmingham corpus are a little surprising. These include "may I thank" (line 11) and "don't be sad" (line 17) in Table 6.10.a (occurring 8 and 3 times, respectively, in the BoE); along with "I don't want" (line 20), "love you all" (line 13) and "my love to" (line 15), in Table 6.10.b (occurring 18, 72 and 75 times, respectively, in the BoE). To some extent this can also be said of "please forgive me", which occurs in both the "Hers" and the "Hims" (line 22 in Table 6.10.a, and line 16 in Table 6.10.b) and occurs 79 times in the BoE.

Very few 3- or 4-word clusters were used to any significant extent in the Birmingham corpus. The exceptions among the 4-word clusters of the "Hers" (see Table 6.9.a) are perhaps kisses, although these are used in only one text; and "I shall no longer", "may I thank you" and "shall no longer be" (lines 3, 4 and 5 respectively) which each appear in 8.11% of the texts (in addition to being rather infrequent in the BoE). Exceptions in the 4-word clusters of the "Hims" (see Table 6.9.b) are perhaps the (digital) numbers which appear 20 times in 11 texts, although this is only 5.19% of the "Him" texts.

There are several 3-word clusters that are present in over 8% of the "Her" texts, but, as noted above, most of these are quite frequent in the BoE. However, it is worth pointing out that over 12% of the "Her" texts contained 3-digit numbers, and over 10% of them said "I thank you" (lines 2 and 4 in Table 6.9.a) which is moderately frequent in the BoE. The most striking clusters in the "Hers" are "I shall no", "may I thank", "shall no longer" and "this letter I" (lines 10-13 in Table 6.9.a) because they are not too common in the BoE, and yet are present in over 8% of the "Her" texts.

The 3-word clusters in the "Hims" (see Table 6.10.b) that are notable are "I love you" (line 1) which features in over 22% of the texts and digital numbers (line 2) which occur in over 12% of them. Additionally worth noting are "I do not" (line 4) and "I hope you" (line 5) in Table 6.10.b which also occur in over 8% of the "Him" texts. However, all four of these clusters are quite frequent in the BoE, whereas "all my love" (line 3 in Table 6.10.b) is only fairly frequent in the BoE (occurring 102 times), and yet occurs in over 10% of the "Him" texts.

Thus, certain formulations of proffering thanks or love; writing about the note itself (as in 'this letter'), and saying 'I shall no longer be…' are, to some degree, key clusters in the Birmingham suicide notes.

6.5. Conclusion

The "Hers" and the "Hims" are both closer to spoken texts than they are to written texts. Both sub-corpora are also similar in their main key words when compared to a standard reference corpus, the BNC. Also, nearly a third of the "Her", and nearly half of the "Him", subjects in some way referred to mothers in at least one of their suicide notes.

Since suicide notes are an 'occluded genre' (see Chapter 1), it is notable that in the Tables of the top most frequent word and keyword lists (and in many of the Figures in this thesis), the results for the "Hers" and the "Hims" are remarkably similar. This suggests that there is a cognitive structure present, since the note-writers are using similar structure and language although they can have no awareness that they are doing so.

The above word and key word lists only compare individual words. Their lexical results can give clues, as with the key word 'mom', but they are limited in their usefulness for showing the kinds of things that suicide note-writers write about. Looking at sets of words grouped according to some semantic concepts would be a far more profitable route to take. Rather than attempt to expand on the small sets of words developed above (Section 6.3.2) I wanted to employ an already existing and established

software application which would not only aid this process but do so automatically. Software applications that achieve this are known as semantic taggers and the process is called semantic tagging: this will be the focus of the next chapter.

7. WHAT ARE SUICIDE NOTES ABOUT?: SEMANTICS

In this chapter, to order to extend my investigation of what suicide notes are about, I examine some semantically-orientated, content analytic lists with the aid of a semantic tagger (see Section 6.5 above). I used Wmatrix (Rayson, 2007) for the semantic tagging (see Section 4.11), and the BNC as a reference corpus (see Section 4.7). I also used 'grep' to aid my searches of the corpus (see Section 4.11).

The results of the semantic tagging should suggest a description of the average, or typical, suicide note. Bakhtin (1999) describes linguistic genres as follows, where his use of the term "speech" includes the written as well as the spoken.

> "A speech genre is… a typical form of utterance; as such the genre also includes a certain typical kind of expression that inheres in it. … Genres correspond to typical situations of speech communication, typical themes, and, consequently, also to particular contacts between the *meanings* of words and actual concrete reality under certain typical circumstances." [Bakhtin's italics.] (ibid., p.129).

Extracting some typical semantic features from the corpus should lead to an understanding of the suicide note genre.

7.1. Semantic Tagging

Wmatrix (Rayson, 2007; see Section 4.11) has 21 major semantic categories, and many minor sub-categories. According to Archer, Wilson and Rayson (2002), there are 232 categories altogether. The major ones are shown in Table 7.1 which is reproduced from Archer et al. (2002, p.2). The contents of categories are not always mutually exclusive and it is possible for words to fall into more than one of them (Archer et al., 2002, p.2).

The semantic concepts used for the categories are based upon general meaning, rather than on any psychological interpretation. According to Archer et al. (2002, p.2) the 1981 <u>Longman Lexicon of Contemporary English</u> was used for Wmatrix's first semantic categories, and although Wmatrix has evolved somewhat during its lifetime

(ibid., p.2), the basis for its categories has not changed. This can be contrasted with the General Inquirer (Stone et al., 1966) which tends to be used with dictionaries based on psychological concepts, such as the "Harvard II Psycho-sociological Dictionary", (Ogilvie, Stone, and Shneidman (1966, p.528).

Table 7.1. Wmatrix Major Categories

A general and abstract terms	B the body and the individual	C arts and crafts	E emotion
F food and farming	G government and public	H architecture, housing and the home	I money and commerce in industry
K entertainment, sports and games	L life and living things	M movement, location, travel and transport	N numbers and measurement
O substances, materials, objects and equipment	P education	Q language and communication	S social actions, states and processes
T Time	W world and environment	X psychological actions, states and processes	Y science and technology
Z names and grammar			

(Reproduced from Archer et al. (2002, p.2))

7.1.1. Aboutness

In the quest to find out what suicide notes are about I decided to use Wmatrix's semantic tagging facility, and to see what the most frequent semantic domains were (in what follows I use the terms 'domain' and 'category' as synonyms). It turned out that the threshold of 0.5% of the total text proved to be a convenient cut-off point, above which the categories displayed were potentially useful ones. Moreover, using this threshold yielded similar numbers of categories in the "Hers" and the "Hims": 29 and 28 respectively. These are shown in Tables 7.2.a and 7.2.b.

In order to use Wmatrix to discover what semantic categories there are in one's own corpus one can simply upload the appropriate file to the Wmatrix website, run the tagger and view the output. However, although the resulting display will show the semantic category codes (e.g. "A7+"), it will not contain the corresponding names (e.g.

213

"Likely"). To obtain a list that includes the category names it is necessary to use one's own corpus also as the reference corpus, and thus to compare the corpus with itself. This results in a list that has duplicate columns of raw frequencies and their percentages because the two corpora are actually identical. The list also includes a column of the Log Likelihood calculation of the 'keyness' resulting from the comparison of the two corpora, which in this case will be zero for every row. Performing a 'cut and paste' operation to save the file to a computer disk, and removing the duplicated and keyness columns, yields a list such as those shown in Tables 7.2.a or 7.2.b below.

In these Tables, 'Item' is the Wmatrix semantic category code, 'O1' is the frequency of occurrence of the category within the corpus (in this case the "Hers" and "Hims" respectively), '%1' is the 'O1' figure given as a percentage of the whole corpus, and the unnamed column on the far right gives the names of the semantic categories (and sub-categories where appropriate). The numbers on the category codes indicate sub-divisions; the signs indicate whether the words within the category are semantically positive (+) or negative (-), and to what degree ('++' meaning very positive, and '---' meaning extremely negative, for example) (Archer et al., 2002, p.1-2). For example, the categories 'bad', 'worse' and 'worst' have the tags A5.1-, A5.1-- and A5.1---, respectively. These positive and negative signs appear relate to the assumed inherent nature of the words and the categories themselves according to Wmatrix. For example, if a text includes the phrases 'I'm not unhappy' and 'this is not a bad thing', the words 'unhappy' and 'bad' get assigned to categories whose tags have negative signs attached to them, and the 'not' does not change this. Similarly, phrases such as 'I do not love you' still end up with 'love' being assigned to a positively tagged category. However, any tag could potentially be rendered either positive or negative, for example A7+ or A7- for 'probable' and 'improbable' respectively.

In Tables 7.2.a and 7.2.b. about half way down, the category 'Entire: maximum' can be seen. This (as Table 7.3 will show) includes the word 'all' which features as a keyword in Tables 6.4.a and 6.4.b (see Section 6.3). 'All' is also high in the keyword distribution lists of Tables 6.5.a and 6.5.b.

Table 7.2.a. Semantic concepts >= 0.5%: "Hers"

```
Item          01      %1

Z5          1579     23.00      Grammatical bin
Z8          1178     17.16      Pronouns
Z99          387      5.64      Unmatched
A3+          213      3.10      Existing
Z6           146      2.13      Negative
A7+          110      1.60      Likely
Z4            98      1.43      Discourse Bin
E2+           96      1.40      Like
Z1            94      1.37      Personal names
N1            93      1.35      Numbers
A9+           88      1.28      Getting and possession
T1.1.3        85      1.24      Time: Future
A1.1.1        85      1.24      General actions / making
N5.1+         80      1.17      Entire; maximum
M6            73      1.06      Location and direction
X7+           65      0.95      Wanted
M1            64      0.93      Moving, coming and going
A13.3         64      0.93      Degree: Boosters
S4            63      0.92      Kin
T1.3          56      0.82      Time: Period
B1            53      0.77      Anatomy and physiology
X2.1          50      0.73      Thought, belief
X2.2+         50      0.73      Knowledgeable
Q2.2          44      0.64      Speech acts
A14           39      0.57      Exclusivizers/particularizers
B3            38      0.55      Medicines and medical treatment
L1+           38      0.55      Alive
E4.1-         37      0.54      Sad
A5.1+         37      0.54      Evaluation: Good
```

Table 7.2.b. Semantic concepts >= 0.5%: "Hims"

```
Item          01      %1

Z5          6783     23.45      Grammatical bin
Z8          5403     18.68      Pronouns
Z99         1579      5.46      Unmatched
A3+          975      3.37      Existing
Z6           629      2.17      Negative
A7+          503      1.74      Likely
A9+          425      1.47      Getting and possession
Z4           415      1.43      Discourse Bin
M6           358      1.24      Location and direction
Z1           350      1.21      Personal names
E2+          351      1.21      Like
N5.1+        344      1.19      Entire; maximum
A1.1.1       341      1.18      General actions / making
A13.3        316      1.09      Degree: Boosters
N1           312      1.08      Numbers
S4           297      1.03      Kin
M1           290      1.00      Moving, coming and going
T1.1.3       257      0.89      Time: Future
Q2.2         247      0.85      Speech acts
X7+          229      0.79      Wanted
X2.2+        222      0.77      Knowledgeable
X2.1         215      0.74      Thought, belief
T1.3         187      0.65      Time: Period
O2           170      0.59      Objects generally
Z3           161      0.56      Other proper names
A14          160      0.55      Exclusivizers/particularizers
A5.1+        159      0.55      Evaluation: Good
Q2.1         157      0.54      Speech: Communicative
```

As the tables above show, the six most frequent semantic concepts are the same, and occur in the same order, in the "Hers" and the "Hims". '**Pronouns**' (see Chapter 3 and Sections 6.2 and 6.3) is fairly self-explanatory; '**Unmatched**' words mostly consist of anonymised names and mis-spellings (see Chapter 4 and Sections 6.2.1 and 6.3.

Because Wmatrix is not a psychological tool, the presence of '**Negative**' as a major concept (5th most frequent in both sub-corpora) does not necessarily mean that the "Hims" and "Hers" have a negative outlook on life. Others, such as Osgood & Walker (1959), have found that suicide notes tend to portray a positive outlook (see Section 3.1.2, Chapter 3). Further, according to Houston (2003, p.1) a medical study conducted in the Republic of Ireland, by Feeley, Gafoo, Connolly and Brophy, which recovered 145 notes from 772 suicide cases for the year 2000, and examined their "emotional content", found "positive emotion in a majority of notes...".[38]

The presence of 'Negative' as a frequently-occurring concept merely means that the "Hims" and "Hers" tend to include a large quantity of words and morphemes such as 'not', 'no', 'nothing', 'none' and 'n't' which Wmatrix classifies as negative. Using Wmatrix to examine the concordance lines where these words were found in the "Hims" did indeed show that most of the uses of the words in the 'Negative' category were actually negative. However, as mentioned at the beginning of this section, Wmatrix cannot prevent entities such as "Please don't be sad I 'm not Im happy now…" (M96488a) from being assigned to its 'Negative' category. The tagger has no way of knowing when a human speaker/hearer would use the negatives to 'cancel each other out'.

As mentioned before (see Chapter 2), negatives are important. They always presuppose an expected positive (Coulthard 1992, p. 252), and so they reveal something about expectations. For example, 'Please don't be sad' assumes that the addressee might be expected to be sad, and 'I'm not sad' presupposes that the addressee might expect the writer to be sad. Table 7.3 presents summary explanations of some more of the

---

[38] Houston's piece, being a newspaper article rather than an academic paper, does not provide a proper reference for the study being cited here, which makes it impossible to locate the original source.

categories, particularly the less self-explanatory ones. The descriptions are mostly paraphrased or taken directly from Archer, et al. (2002, p.3-36).[39] Some are compiled from seeing the Wmatrix results directly.

Table 7.3. Some Wmatrix Categories

Category	Explanation
Grammatical bin	Prepositions, some adverbs and conjunctions such as 'for', 'subject to,' and 'but' (Archer et al., 2002, p.36).
Existing	"General or abstract terms relating to being or existing" (ibid., p.5) such as 'consist' and 'exist'.
Likely	"Abstract terms of modality (possibility, necessity, certainty, etc.)" (ibid., p.6) such as 'can' and 'made certain'.
Discourse Bin	Various discourse markers such as 'please', 'sorry', 'you know' and 'you see' (ibid., p.36).
Like	Positive affect. Already mentioned in Sections 3.1.7 and 5.3.
Getting and possession	Includes words such as 'accrued' and 'acquire' (ibid., p.7).
Entire: maximum	Maximum quantity terms (ibid., p.20), entities such as are conveyed by the words 'all', 'any' and 'each', for example.
Wanted	Terms about wanting, desire and aspiration such as 'ambition' (ibid., p.34).
Degree: Boosters	"Intensifiers that amplify to a high degree (but not the upper extreme)" (ibid., p.8) and include terms such as 'more and more' and 'deeply'.
Knowledgeable	Words such as 'acquainted' and 'aware' (ibid., p.32).
Speech acts	Terms used in speech acts, such as 'accuse' and 'addressing' (ibid., p.24).
Speech: Communicative	"Terms relating to spoken communication" (ibid., p.23), such as 'chat' and 'conversation'.
Exclusiveizers/particularizers	Terms that foreground or focus some entity. They include words such as 'overall' and 'purely' (ibid., p.9).
Alive	About "Life and living things" (ibid., p.16.) and includes words such as 'live' and 'life'.
Other proper names	Includes products and companies, and acronyms and words such as 'Nike' and 'Airport' (ibid., p.36).

---

[39] The category names given in Archer et al. can differ from the ones produced by Wmatrix, particularly when the category tags can have plus or minus signs suffixed to them. For example, the 6th row of Table 7.2.a shows tag 'A7+' as the category named 'Likely'. In Archer et al. (2002, p.6) this category (A7) is called 'Definite (+ modals)', but more to the point, it includes negative examples (i.e. examples of items that are not likely) such as 'contentious' which would be given the tag 'A7-' (and a name different from 'Likely'!).

The "Hers" and the "Hims" share 25 of the semantic concepts shown in Tables 7.2.a and 7.2.b. Here are examples of some of them. (Examples of the remaining common concepts are given in Section 7.1.2 and beyond, where they re-emerge in the light of further analysis.) The words pertinent to the category are in bold typeface. One can see from these examples that the tagging, as with all such attempts to automate a human process, is not always perfect.

(7.1. 'Grammatical bin')
"Im sorry but I cant cope **anymore**." (m95021);

(7.2. 'Getting and possession')
"I would like you to **have** my briefcase." (f96320c);

(7.3. 'Location and direction')
"…I could not express what was **inside of** me…" (m95054);

(7.4. 'Personal names')
"Letter Rack - **TV.Licence BROWN** ENVelope Look." (f96369g);

(7.5. 'General actions / making')
"I no that i have **made** the bigest mistak of all…" (m95468e);

(7.6. 'Numbers')
"I have had my **three score** years and ten…" (f96460c);

(7.7. 'Moving, coming and going')
"A lot of letters **come** for old tenants…" (m95161e);

(7.8. 'Speech acts')
"**Tell** the kids I loved them all very much…" (f99123f);

(7.9. 'Wanted')
"ALL I EVER **WANTED** WAS FOR US TO BE A FAMILY…" (m96318e);

(7.10. 'Thought, belief')
"I have **thought** about this long and hard." (f98279c);

(7.11. 'Time: Period')
"It is a beautifull **day** the SUN is OUT…" (m98222c);

(7.12. 'Evaluation: Good')
"You make a **good** lyer…" (f96049a).

There are 4 concepts in the "Hers" that are not in present in the "Hims" (given the 0.5% cut off point). These are 'Anatomy and physiology', 'Medicines and medical treatment', 'Alive', and 'Sad'. Conversely there are 3 concepts in the "Hims" that are not present in the "Hers". These are 'Objects generally', 'Other proper names' and 'Speech: Communicative'. This suggests that females tend to write more than males do about bodies and health, the fact of their being alive, or perhaps their quality of life, and being unhappy; and that males tend to write more than females do about objects and conversations. Here are some examples.

"**Hers**"

(7.13. 'Anatomy and physiology')
"A **Nerve** Specialist is coming my house see me…" (f96369b);

(7.14. 'Medicines and medical treatment')
"I do not want to be **resuscitated**…" (f96460c);

(7.15. 'Alive')
"My **life** is such a mess and I can only see things getting worse." (f97361a);

(7.16. 'Sad')
"It **breaks my heart** to do this…" (f95302);

"**Hims**"

(7.17. 'Objects generally')
"Use my flat - you have **keys**…" (m95161c);

(7.18. 'Other proper names')
"I WRITE THIS WITH TOTALLY UNSOUND **MIND**!" (m99483b);

(7.19. 'Speech: Communicative')
"It is real malicious **gossip** ." [sic] (m99510d).

219

7.1.2. Distinctiveness

I now want to look at the semantic categories that are 'key' with respect to a non-specifically suicide-orientated standard reference corpus. Although the Birmingham suicide notes have some commonalities with speech (see Section 6.2), they are actually written texts, and so I have used the BNC written sub-corpus as my reference corpus. The top twenty semantic concepts for the "Hers" and the "Hims" are shown below in Tables 7.4.a and 7.4.b respectively.

As before, in Tables 7.4.a and 7.4.b 'Item' is the Wmatrix semantic category code, 'O1' is the frequency of occurrence of the category within the user corpus (the "Hers" and "Hims" respectively), and the unnamed column on the far right gives the names of the semantic categories and sub-categories. Now, however, the 'LL' column for the Log Likelihood statistic gives meaningful information, since we are now making genuine comparisons between different corpora. The LL statistic shows the degree of 'keyness' of the user corpus compared to the reference corpus.

Table 7.4.a. Top 20 Key Semantic concepts: "Hers"

Item	O1		LL	
Z8	1178	+	628.36	Pronouns
E2+	96	+	261.72	Like
Z99	387	+	235.43	Unmatched
L1+	38	+	220.19	Alive
Z6	146	+	95.37	Negative
Z4	98	+	91.42	Discourse Bin
E4.1-	37	+	62.80	Sad
S1.2.4+	15	+	54.06	Polite
A5.2-	21	+	53.60	Evaluation: False
T1.1.3	85	+	52.17	Time: Future
S1.2	12	+	50.77	Personality traits
S3.2	23	+	45.76	Relationship: Intimacy and sex
X2.2+	50	+	44.10	Knowledgeable
S3.1	30	+	35.12	Personal relationship: General
B3	38	+	34.65	Medicines and medical treatment
S4	63	+	30.95	Kin
W2-	3	+	29.74	Darkness
N5.1+	80	+	26.59	Entire; maximum
A13.3	64	+	20.61	Degree: Boosters
W2	2	+	19.82	Light

Table 7.4.b. Top 20 Key Semantic concepts: "Hims"

Item	O1	LL	
Z8	5403 +	3306.01	Pronouns
Z99	1579 +	874.35	Unmatched
E2+	351 +	823.99	Like
L1+	102 +	457.76	Alive
Z6	629 +	414.36	Negative
Z4	415 +	376.59	Discourse Bin
S3.2	116 +	251.91	Relationship: Intimacy and sex
T1.1	48 +	210.94	Time: General
X2.2+	222 +	204.16	Knowledgeable
S4	297 +	178.08	Kin
A13.3	316 +	147.64	Degree: Boosters
X2.6+	93 +	123.77	Expected
N5.1+	344 +	117.60	Entire; maximum
N6+++	65 +	113.02	Frequent
A7+	503 +	90.02	Likely
A3+	975 +	77.05	Existing
S1.2.4+	32 +	73.21	Polite
T1.1.3	257 +	68.03	Time: Future
W2	9 +	63.72	Light
A14	160 +	59.73	Exclusivizers/particularizers

As the tables above show, the six most key distinctive semantic concepts are the same, albeit in a slightly different order, in the "Hers" and the "Hims". The "Hers" and "Hims" each have 13 categories in common with those shown in Tables 7.2.a and 7.2.b respectively. Each sub-corpus has 7 semantic categories which are revealed here to be 'distinctive' but which did not appear as high frequency items and as such, indicative of the 'aboutness' of the sub-corpora. I refer the reader to Section 7.1.1 above for details of the categories that featured as indicators of 'aboutness' and have therefore already been described.

Both the "Hers" and the "Hims" share the distinctive semantic concept of 'Alive' (words such as 'life' and 'live') as their 4th most key concept; and in 6th most key position, both sub-corpora feature the 'Discourse Bin'.

Thus, in comparison with a standard norm, the BNC, both the "Her" and "Him" suicide notes contain many pronouns, and may be about liking something or someone; and about life, and being, or not being, alive. They may both also be about, or at least contain in their texts, the concepts of politeness, future time, intimate relationships, knowledge, kinship or family relations, 'Entire; maximum' entities (such as are conveyed by the words 'all, 'any' and 'each'), 'degree: boosters' (such as 'more and more' and 'deeply'), and 'Light'. Here are some examples of three of the categories

that are common to both sub-corpora. (Examples of the others will be given before the end of this Section.)

> (7.20. 'Relationship: Intimacy and sex')
> "I send you my fond **love**" (f96460l);
> "Give your Mum a farewell **kiss** from me." (m96081);
>
> (7.21. 'Polite')
> "**Thanks** for the messages you gave me…" (f99123h);
> "I want to say firstly how **grateful** I am to the both of you for everything you have done for me." (m96316c);
>
> (7.22. 'Light')
> "This must be better than any **light** and safer than a know way." (f96460b);
> "ALL GAS AND POILOT **LIGHTS** TURNED OFF" (m98354b).

It is worth noting that the Wmatrix concept of 'Light', in 20th and 19th place in the "Hers" and "Hims" key concepts respectively, is a useful reminder that being key may bear little relationship to being high-frequency. 'Light' only occurs twice in the "Hers", and nine times in the "Hims", and so cannot be any kind of distinguishing feature of suicide notes. (Indeed, it has been shown in Section 7.1.1 that 'Light' is not an indicator of aboutness.) The same can be said of 'Darkness' in the "Hers": it is fairly distinctive but has a low frequency and cannot therefore be a characteristic of many of the notes. A more useful distinguishing feature would be a set of words that occurred in most of the texts. Potentially, 'Time: Future' is such a set (see the discussion of 'will' in Section 6.2 and of the 'Future' set of lexis in Section 6.3; and the 'Key Words' Table, 6.4.b). 'Time: Future' occurs 85 times in the "Hers" and 257 times in the "Hims", and considering that each of these figures is greater than the total number of texts in the sub-corpus concerned, there is the possibility that future time is mentioned in every text. Whether or not this is actually the case remains to be seen. 'Time: Future' does appear to be a strong candidate as an indicator of aboutness, at least for the corpus as a whole if not for individual texts.

The concepts common to both the top 'aboutness' categories and the top key ones are those which best define suicide notes in terms of what they are about and which also

distinguish them from other texts. Logically 'AND'ing the 'aboutness' and 'distinctive' concepts from the "Hers" and the "Hims" together results in the following ten categories: 'Pronouns', 'Like', 'Unmatched', 'Negative', 'Discourse Bin', 'Time: Future', 'Knowledgeable', 'Kin', 'Entire: maximum' and 'Degree: Boosters'.

Pronouns and unmatched items have already been discussed; I now provide some examples from the other eight semantic categories that show both aboutness and distinctiveness. In each case I have put the word (or phrase, or part of word) that signifies membership of the category in bold typeface.

(7.23. 'Negative')
 "I have an incurable abdominal condition. The treatment did**n't** help."
(f97050b.txt);
"There is just **nothing** to live for any more." (m99510c.txt);

(7.24. 'Discourse Bin')
"**Please** tidy up. The washing will need drying and the flowers will need watering." (f97158b.txt);
"I had no quality of life **anyway** so I decided to end it." (m96335a.txt);

(7.25. 'Degree: Boosters')
"I never know you hated me **so** much" (f99123i.txt);
"I refused to accept things were **really** that way." (m95054.txt);

(7.26. 'Entire: maximum')
"I hope I've said **all** that I should." (f97158a.txt);
"CANNOT BE BOTHERED TO DO **ANY** HOUSEWORK" (m96318a.txt);

(7.27. 'Kin')
"We have been **married** for 52 years." (f97158a.txt);
"I know youre calling me a stupid **Bastard** but Ive had enogh" [sic]
(m96488g.txt);

(7.28. 'Knowledgeable')
"It has been a great pleasure to **know** you." (f96460f.txt);
"Try to **remember** me with nice thoughts,…" (m97145a.txt);

(7.29. 'Like')
"I would **like** you to have my briefcase." (f96320c.txt);
"TO SAY I **LOVE** HER SEEMS SO INADEQUATE ." (m96395a.txt);

(7.30. 'Time: Future')
"Who knows? Perhaps **one day** we will meet again." (f96460f.txt);

"This I am sorry to say **will** be a shock." (m96263c.txt).

One could say that suicide notes are distinctly about being affectionate, referencing the future (including the discovery of the suicide), things their authors know, and their authors' relatives; and they may be further distinguished by their greater than normal use of pronouns, people's names, mis-spellings, negatives, discourse markers, maximum quantity terms, and intensifiers.

The "Hers" and the "Hims" each have 6 of the top 20 key distinctive semantic concepts which are peculiar to them. In other words, there are 6 concepts in the "Hims" top 20 that are not present in the "Hers" top 20, and vice versa. As the above tables show, for the "Hers" these include 'Sad' (which includes words such as 'unhappiness', 'suffering' and 'misery') and 'Evaluation: False' (which includes words such as 'lied' and 'liar'), and for the "Hims" they include 'Expected' (which includes words such as 'hope' and 'expect') and 'Likely' (which includes a large number of modals: words such as 'can' and 'might'). However, with one exception, these concepts are all key in both sub-corpora, but they fall below 20th place in their relative keyness. The exception is 'Medicines and medical treatment' which is the 15th most key word in the "Hers" but is not key at all in the "Hims".

One could speculate that the differences between the "Hers" and the "Hims", when compared to the written BNC corpus, might indicate some degree of gender difference. The females write more than the males do about sadness; tellers of falsehoods; 'Personality traits' such as kindness; general personal relationships such as friendships; medicines and medical treatment; and darkness. The males write more than the females do about time in general; expectations (using words such as 'hope'); frequent occurrences or frequency (for example, 'always'), likely events (using words such as 'probably'), existing situations (using words such as 'is', 'real' and 'events'), and 'Exclusiveizers/particularizers' (using words such as 'just' and 'only'). One could argue that while the females tend to write about people, the males tend to talk about objects, which would make the "Hers" and "Hims" somewhat stereotypical in terms of their authors' gender. Here are some examples.

"**Hers**"

(7.31. 'Sad')
"I am so sorry to **upset** you." (f97158b);

(7.32. 'Evaluation: False')
"She has been telling you **lies** about me but you would not listen." (f96049b);

(7.33. 'Medicines and medical treatment')
"The **treatment** didn't help." (f97050b);

(7.34. 'Personality traits')
"I want you to do a great **kindness** for me." (f97008);

(7.35. 'Personal relationship: General')
"I have always considered myself very lucky to have you as my **friend**."
(f96320i);

(7.36. 'Darkness')
"Your steadfastness and integrity, your uncritical acceptance of your friends and acquaintances in their **darkness** and light, is an example to us all." (f96460h);

"**Hims**"

(7.37. 'Expected')
"My guitars are worth £500 ,[sic] I **hope** it covers the cost." (m95054);

(7.38. 'Likely')
"I TOLD YOU I **COULD** NOT LIVE WITH OUT YOU AND I MENT IT."
(m98362);

(7.39. 'Time: General')
"This **ever** growing risk I pose to children is just too scarry to ignore ."
(m99483a);

(7.40. 'Frequent')
"I have **always** enjoyed your company ." (m96263c);

(7.41. 'Existing')
"I**'M** SORRY THIS LETTER **IS** ALL OVER THE PLACE BUT I'VE - WRITTEN IT HOW I FELT." (m99424a);

(7.42. 'Exclusivizers/particularizers')
"But Ive **just** missed the boat." (m96488g).

One reason for using a reference corpus, as has been said before, is to facilitate a

linguistically valid comparison of the "Hers" and "Hims". Comparing the "Hers" and "Hims" to each other directly (that is, using each as a reference corpus to the other) however, led to a potentially important piece of information about Wmatrix that must be acknowledged, and borne in mind when interpreting its results. Comparing the "Hims" with the "Hers" (that is, using the "Hers" as a reference corpus in Wmatrix) resulted in the most key semantic concept being 'Information technology and computing'. This seemed implausible. Inspection of the concordance lines in Wmatrix revealed that capitalised versions of the word 'it' were being assigned to this category, as if the letters stood for "Information Technology". Thus, as with most software, one should not consider the output of Wmatrix as infallible. Below, I have conducted some further investigation of some of the above categories.

7.2. Semantic Distribution

Wmatrix does not provide any read-out for individual texts within a corpus: it only deals with a corpus as a whole entity. I therefore worked with Wmatrix to obtain an overall picture of my sub-corpora; but then, in order to check whether any part of that picture was skewed due to a small number of individual texts, or their authors, I used a different approach. I now copied from Wmatrix the word lists of any interesting semantic categories, and use my 'grep' line-based application (see Section 6.3 above) to search my suicide notes and obtain a read-out by note and by subject. In Section 7.2.1 I consider a single category; then in Section 7.2.2 I examine a selection of categories.

7.2.1. Sadness in female-authored notes

The Wmatrix semantic category 'sad' is the most key concept, being in seventh place, that 'distinguishes' the "Hers" from the "Hims". The 'sad' words found by Wmatrix in the "Hers" are 'grieve', 'sad', 'upset', 'unhappiness', 'suffering', 'crying', 'misery', 'depressed', 'unhappy', 'grieved', 'pity' and 'regret'. Wmatrix also handles multi-word units, and it found the phrase 'breaks_heart' (where '_' is effectively a wild card) in addition to the individual words. As Table 7.4.a, above, shows, there are 37 occurrences of sadness in the "Hers". In other words these 13 lexical items, including the phrase, collectively occurred 37 times in the "Hers". Since there are 74 texts in the "Hers" it would be interesting to know whether the number of texts containing this

'sadness' lexis might amount to anywhere near the figure of 50% of the sub-corpus (37 being half of 74).

The 'grep' line reader cannot currently deal with phrases, but there was only one occurrence of 'breaks_heart'. Ignoring phrases and searching for only the 12 'sad' words produced the following results (including raw counts in parentheses). That there are 37 occurrences shown with the phrase excluded is due to 'Readgrep' finding one instance of 'sad' lexis that had eluded Wmatrix.[40]

```
        Corpus   Occurrences   Notes          Subjects
sad: her   0.50% ( 37)   29.73% ( 22)   36.36% (  8)
```

What this shows is that, indeed, nearly 30% of the notes were involved, making 'sad' words indicative of a text's probably being a suicide text, and possibly being a suicide text authored by a female. It also shows that over 36% of the female note-writers used 'sad' words. So although the total number of occurrences comprising 'sad' terms in the sub-corpus is small, the concept is arguably important, and it can be said that about a third of female-authored suicide notes are at least partially about sadness.

Perhaps not surprisingly, there were no notes that contained all the 'sad' words. On inspection the 'sad' occurrences seemed to be of 4 main types. 12 were words to the effect, 'I am sad'; but 15 were about the sentiment, 'I don't want you to be sad'; 7 were concerned with saying 'I am not sad'; and there was a miscellaneous set comprising the other 3 instances which consisted of one apologising for the fact the addressee was sad, one criticising the government and one recounting someone else being sad. Here is an example of each type.

    (7.43. I am sad)
    "All I want is a life, im now crying for aparent reason I just feel so lost" [sic]

---

[40] This might be because of electronic glitches during the uploading of the corpus to Wmatrix, but it could equally be because of its construction which was "pain/Misery". For each of the 'sad' words Wmatrix found no difference between the total number of occurrences in the corpus and the number of tagged instances.

(F95456a).

(7.44. You, don't be sad)
"Please don't grieve me [sic] I am not worth it I am nothing…" (F97361a).

(7.45. I am not sad)
Nor am I unhappy. On the contrary I am a very happy woman…" (F96460b).

(7.46. Other sad)
 "'Misery government' will be pleased NO PENSION" [sic][41] (F96369b).

7.2.2. A Selection of Categories

I will now examine in some detail the distribution 'by note' of some of the semantic categories from the previous section that were found to indicate both distinctiveness and aboutness. I will exclude the pronouns, unmatched and negative items and the discourse 'bin' and shall instead concentrate on the remaining six semantic categories, namely 'Like', 'Time: Future', 'Knowledgeable', 'Kin', 'Entire: maximum' and 'Degree: Boosters'.

Unlike the 'sad' category, above, these categories included words that also occurred in other categories. For example, for its category 'Like' Wmatrix found 27 word and phrase types in the "Hims". These included the word 'like', of which Wmatrix found 76 instances. However, for its 'Like' category only 34 of these tokens were included. An example of an excluded instance is:

(7.47. A non-'Like' category 'like')
 "I hate leaving everything in a mess like this but I have to." (m95161a.txt).

This meant that every instance had to be checked to ensure that the results which follow were accurate, or at least consistent with what Wmatrix had found. Wmatrix allows the user to view all the concordance lines within a category, but because it views each sub-corpus as a whole, and cannot distinguish individual texts, I carried out the checking via

---

[41] It is, of course, possible that the word "misery" in this example is the writer's mis-spelling of "miserly". This would render the sentence more grammatical, and seems to fit with the general content of this part of the note.

a mixture of Wordsmith's Concordancer and 'grep'. Ultimately, 'Readgrep' was used
to count the various sets of notes and subjects. (I elaborate on this below.)

Figures 7.1 to 7.4 illustrate the distribution of the six categories. The vertical axes
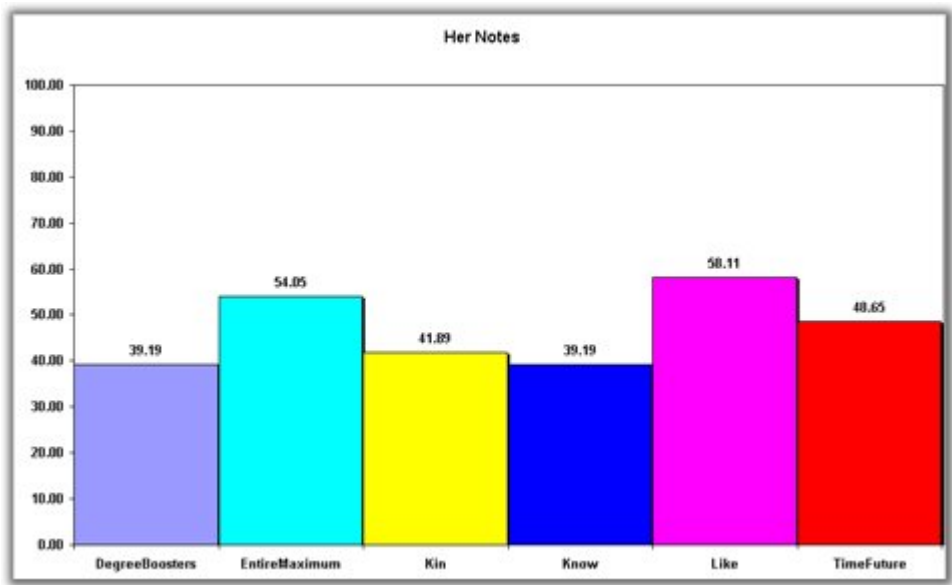represent the percentages of notes containing them in each case.
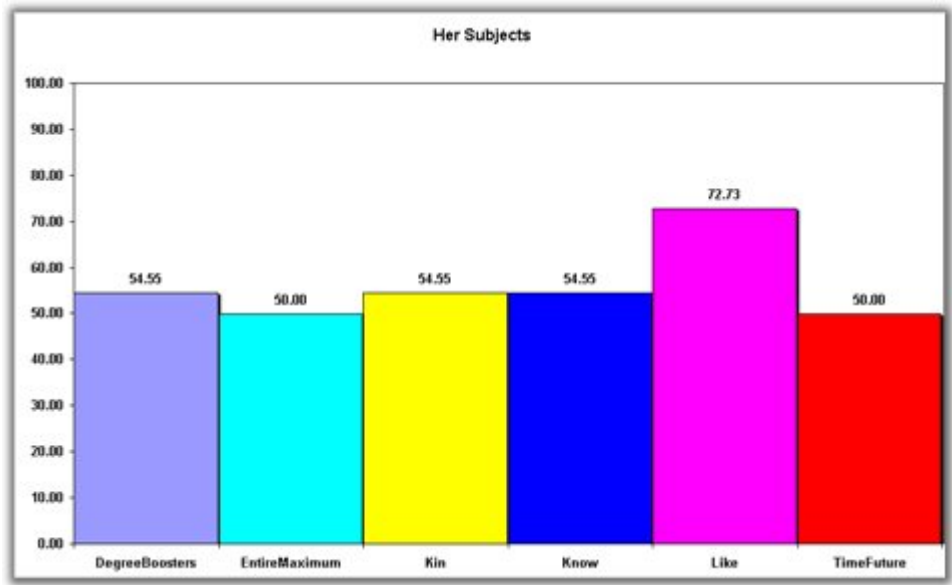

Figure 7.1. Distinctive Aboutness by Note: "Hers"

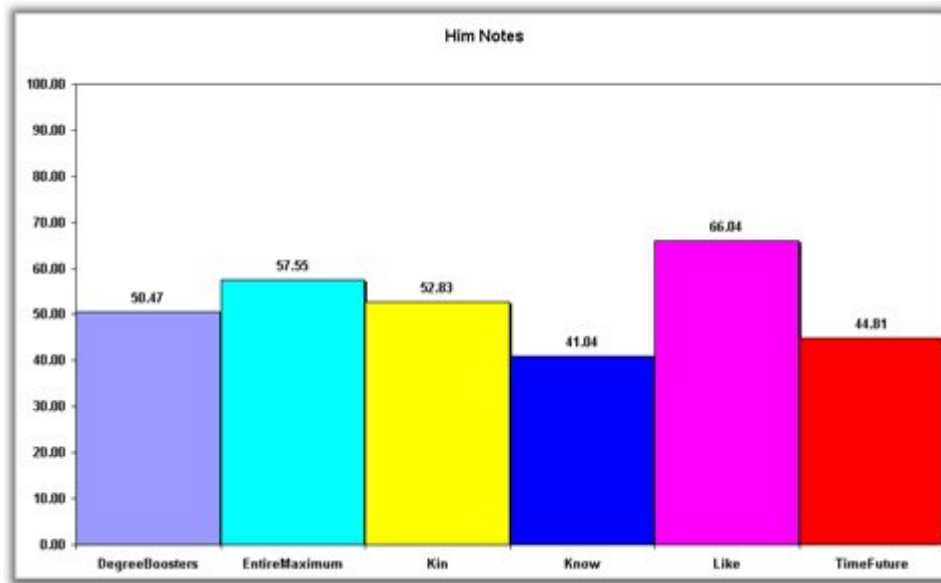**Figure 7.2. Distinctive Aboutness by Subject: "Hers"**
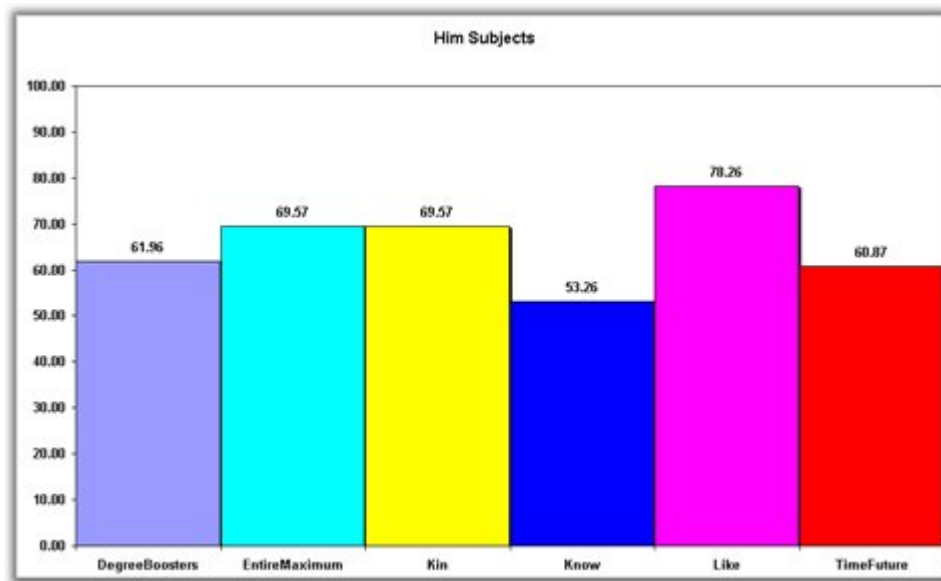
Figure 7.3. Distinctive Aboutness by Note: "Hims"

Figure 7.4. Distinctive Aboutness by Subject: "Hims"

As can be seen in Figures 7.1 and 7.3, in both the "Hers" and the "Hims" the most used category, by note, is 'Like'. The relative proportions of notes using the different categories are similar in the "Hers" and the "Hims", except for ' Time: Future' which appears a little more frequently in the "Hers" (48.65%) than the "Hims" (44.81%). Figures 7.2 and 7.4 show that each category is used by at least 50% of all the subjects.

In both the "Hers" and the "Hims" the most used category, by subject, is 'Like': more subjects used this than any of the other categories.

The 'Like' category incorporates all the instances of the key word 'dear' and most of the occurrences of the key word 'love' in my corpus (see Section 6.3). This lexis is commonly found in the salutations of openings and the valedictions of closings in personal letters and is also present in a proportion of my suicide notes (see Section 5.2). Just as Chapter 6 showed strong similarity between the "Hers" and "Hims" for both the top frequency words and the key words, there is also a remarkable similarity between the "Hers" and the "Hims" in these structural features (see Section 5.2).

Figures 7.1 to 7.4 do not show whether more than one of the six categories co-occur in any single text, or if so, to what degree the categories overlap within texts. In order to obtain this information, it was necessary firstly to produce concordance lines from Wmatrix for all the "hits" for each of the categories under investigation in turn; then I used Wordsmith and 'grep' to search on each of these concordance lines and acquire the text code for the suicide note in question. Having reunited each concordance line with its corresponding text code, I created a set of files containing only the text codes, and sorted them in order of text code.[42] I next used a tool called the Duplicate Lines Finder (Hristov, 2008) to compare the lists of text codes for two categories at a time and produce statistics for the percentages of texts which contained one or both of the categories. Finally I generated Venn Diagrams (Hulsen, 2008) to represent these percentages in a graphic form. Figures 7.5 to 7.8 are the result of this somewhat complicated process.

I have divided the six categories into two groups of three to facilitate the presentation of the results. Three seems to be the optimum number of circles for Venn diagrams as a reasonable compromise between clarity and completeness. The first (arbitrarily selected) group consisted of 'Degree: Boosters', 'Entire: maximum' and 'Knowledgeable', identified in Figures 7.5 and 7.6 as DB, EM and KW respectively.

---

[42] At this point I used the software tool TED Notepad (Medvedik 2006) to make the lines of text codes unique in each file.

The second group consisted of the remaining three categories, namely 'Kin', 'Like', and 'Time: Future' identified in Figures 7.7 and 7.8 as KI, LI and TF, respectively.

Using Figure 7.5 as an example of what the inscribed percentage figures mean, it can be seen that 54.05% of all the texts in the "Her" sub-corpus each contained at least one instance of EM, 29.73% of all the texts contained both DB and EM instances, and 17.57% of all the texts contained instances of DB, EM and KW. This figure of 17.57% is also, coincidentally, the percentage of texts that contained instances of both DB and KW. This is because the texts were counted cumulatively: the count of texts containing DB and KW terms did not exclude any that also contained EM terms. Indeed the DB+KW texts are the same individual texts as the DB+KW+EM ones. Clearly, Figures 7.5 to 7.8 should be viewed as conceptual visualisations rather than accurate representations.
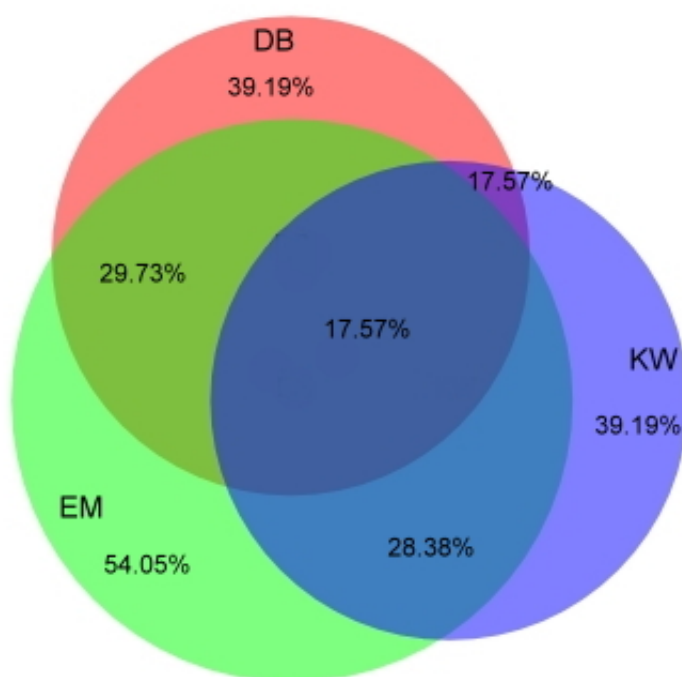
Figure 7.5. 'Degree: Boosters'(DB), 'Entire: maximum'(EM) and 'Knowledgeable'(KW) Categories by Note: "Hers"
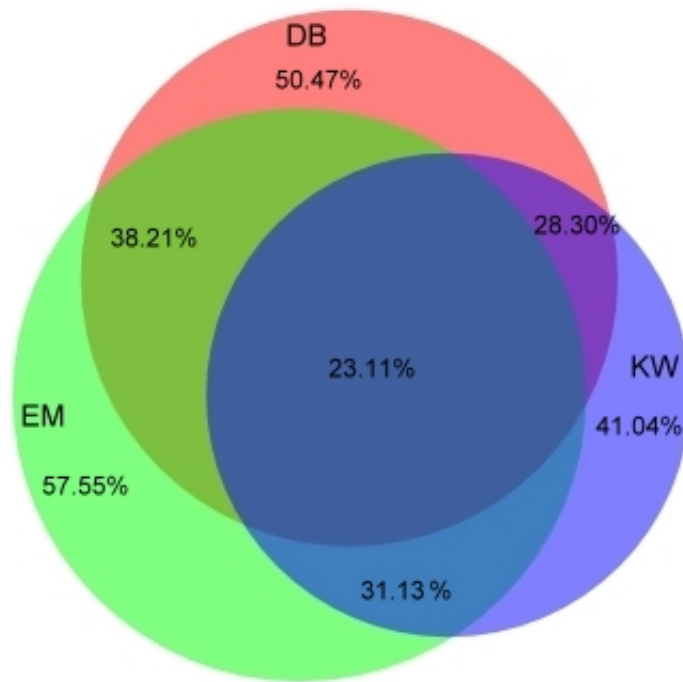
Figure 7.6. 'Degree: Boosters'(DB), 'Entire: maximum'(EM)
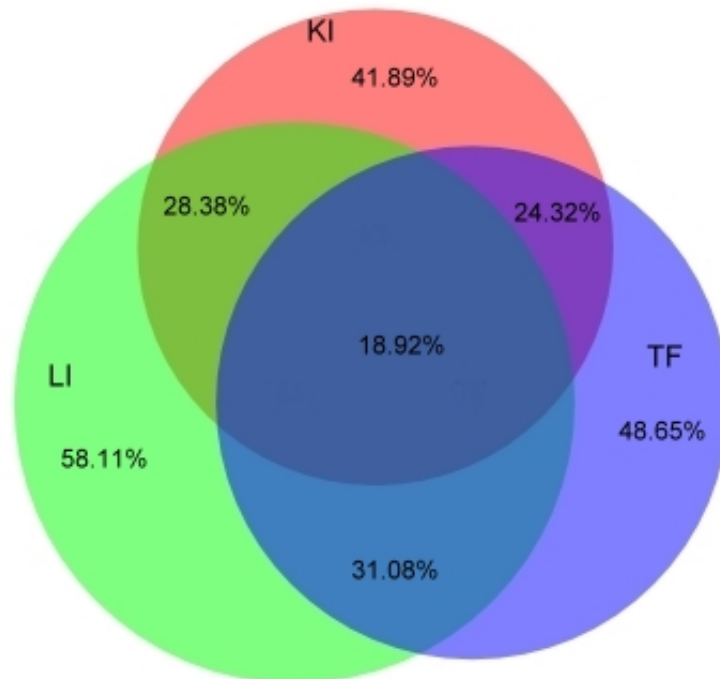and 'Knowledgeable'(KW) Categories by Note: "Hims"

Figure 7.7. 'Kin'(KI), 'Like'(LI) and 'Time: Future'(TF)
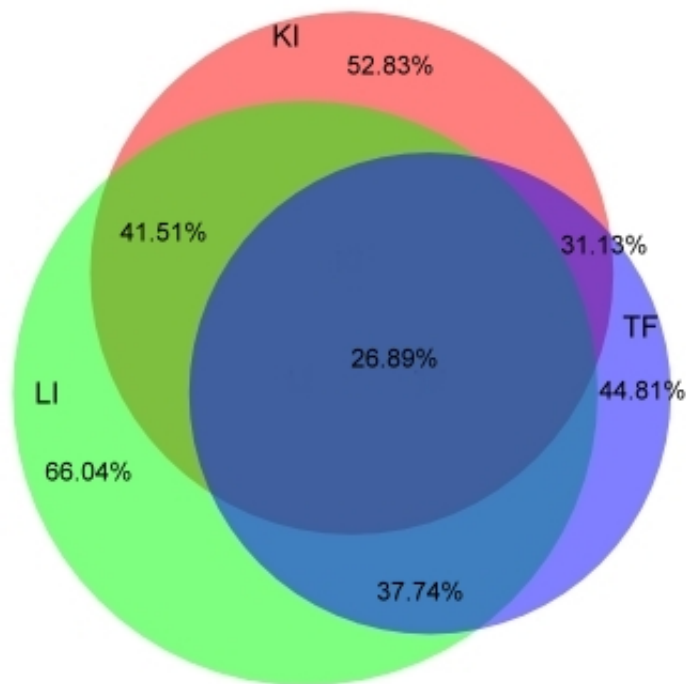Categories by Note: "Hers"

Figure 7.8. 'Kin'(KI), 'Like'(LI) and 'Time: Future'(TF) Categories by Note: "Hims"

Notes containing all of these six categories are, in a sense, exemplars of the typical suicide note. In the next sub-section I consider the issue of prototypicality.

7.3. Prototypicality

In trying to draw some conclusions about what one can expect of a genuine suicide note, I turned to Rosch's work on 'prototypicality' (Rosch 1973; Rosch and Mervis 1975). Rosch was concerned with how we recognise entities as being members of particular categories, and showed that we achieve this mostly by using prototypes of them (see Akmajian et al., 1992, p.379 and p.383, for some discussion).

I should point out that, in the following discussion of prototypicality, (because I have been using the terms 'category' and 'concept' in different contexts and fairly liberally throughout this thesis, and because most of the descriptions below involve these terms) I intend the words 'concept' and 'category' to refer not just to semantic concepts/categories, but to the typical suicide note and the genre of 'suicide notes',

respectively.

Akmajian et al. (1992) give some succinct definitions. A prototype is

"A typical or representative instance of a concept." (ibid., p.489),

and a concept is

"A way of categorizing things, events, and the like, into sets." (ibid., p.480).

Eysenck and Keane (1993) give a useful summary of points about prototypes:

"Concepts have a prototype structure; the prototype is either a collection of characteristic attributes or the best example (or examples) of the concept. …

"There is no limiting set of necessary and sufficient attributes for determining category membership…

"Category boundaries are fuzzy or unclear…

"Instances of concepts can be ranged in terms of their typicality; that is, there is a *typicality gradient* which characterises the differential typicality of examples of the concept. …

"Category membership is determined by the similarity of an object's attributes to the category's prototype…" [their italics].

(ibid., p.264.)

The ideal prototype of a suicide note, then, will have all the most typical characteristics of suicide notes, although it is possible that no single instance of a suicide note will match the prototype exactly.

Of course, a prototypical note should include all the top common concepts that were both frequent and key. This includes the concepts of 'Unmatched', 'Pronouns', 'Discourse Bin' and 'Negative' (see Section 7.1.2). The 'Unmatched' category signifies that a note contains, for instance, anonymised names or mis-spellings. The 'Discourse Bin' is where most of the 'please's and 'sorry's end up. These are keywords in Tables 6.4.a and 6.4.b (see Section 6.3), and are part of my 'politeness' keyword set in Section 6.3.2. Clearly, however, they are not major contributors to the Wmatrix 'polite' category shown in Tables 7.4.a and 7.4.b (see Section 7.1.2) which contains terms of

gratitude such as 'thanks'. Here are some examples that show all the common top semantic concepts.

(7.48. 'Degree: Boosters')
"…but some of it was **so** horrid I could not say it…" (m95161a);
"…they probably won't be **much** use now." (m97157c);

(7.49. 'Entire: maximum')
"Since Mom died **all** I've felt is emptiness…" (m95376a);
"Im so Proud of you **all**…" (m97394e);

(7.50. 'Kin')
"The strong **family** threads have never been broken…" (f96460l);
"…you wanted a **divorce** from me…" (m98036a);

(7.51. 'Knowledgeable')
"…because she **knows** that I am going to be there with her…" (m95214b);
"I don't **know** what's wrong with me" (m98196);

(7.52. 'Like')
"I can't **live with** Fqqq and I can't live without her…" (m96242a);
"I want you to know I **love** you both." (m99223a);

(7.53. 'Time: Future')
"Well I hope you both **will** be very happy…" (f99123i);
"DON'T WORRY ABOUT IT I **WO**N'T HAUNT YOU…" (m99424a);

(7.54. 'Unmatched')
"Give my love to **Ssss** and **Tttt**." (m96335c);
"This ever growing risk I pose to children is just too **scarry** to ignore ." (m99483a);

(7.55. 'Pronouns')
"I tried to find **something** to replace what I'd lost…" (m95054);
"Dear mom by the time **you** read this letter…" (m96480a);

(7.56. 'Negative')
"I wont call you mom as you did**n't** want me to call it you." (f95229);
"I can't ber **not** being with her…" (m96488a);

(7.57. 'Discourse Bin')
"AND YET - **HONESTLY** FNNNA I AM COMPLETELY CALM…" (m95214a);
"But **please** carry out my wishes…" (m97145b).

However, I felt that the categories 'Unmatched', 'Pronouns', 'Discourse Bin' and 'Negative' were linguistically marginal, and decided to continue to work with the six categories discussed in Section 7.2.2 above. Comparing the files of concordance lines generated as described in that section for all six categories (BD+EM+KW+KI+LI+TF) together produced the following results.

	Notes	Subjects
her	4.05% (3)	13.64% (3)
him	16.04% (34)	28.26% (26)

These 3 "Her" and 34 "Him" texts could therefore perhaps be described as 'prototypical' suicide notes, and the other notes could be said to fall into a continuum of prototypicality (Rosch and Mervis 1975). Where they are placed on the scale would then depend on the number of relevant semantic categories they contain.

By the analysis given above, each of the 37 notes mentioned above is equally 'prototypical' because each of them contains at least one instance of all six categories which are the only ones common in both over 0.05% of the corpus and the top key concepts, and also both sub-corpora. (Incidentally, all 37 notes also contain 'Unmatched' items and 'Pronouns'; only one note (in the "Hims") contained no 'Discourse Bin' items; and all except two contained 'Negative' concepts).

My next decision was which of the candidate notes to pick for illustration. A prototype note based on 'distinctiveness', i.e. possessing the highest number of key concepts from all the top key concepts (see Tables 7.4.a and 7.4.b), including the above six, is one possibility. The 37 notes each contain at least 11 of these categories, while the maximum number of categories exhibited is 19 out of 20. Alternatively, focusing on 'aboutness', i.e. picking the highest number of semantic concepts from those that are most frequent in the whole corpus (see Tables 7.2.a and 7.2.b) yields a range of 18 to 28 concepts across the 37 notes. (Combining 'distinctiveness' with 'aboutness' does not help, because no single note contains the highest number of features of both.) Further, the more concepts they contain the larger the notes tend to be. One note with 27 corpus-

frequent semantic concepts and 17 top key semantic concepts is over 1,000 words long. Indeed, in some ways this could be said to be the most typical note, but its length makes it highly atypical (see Section 4.10).

Instead of showing a note which contains as much as possible, as Leenaars (1988) did with his fabricated note (see Section 5.8 above), I am concerned here with showing the common semantic concepts described above, with the intention of keeping my illustrations clear and simple. Additionally, as mentioned above, another aspect of prototypicality is note length. 89% of all the "Hers" and 77% of all the "Hims" are under 200 words long (see Section 4.10). Excluding notes of 200 words and over to control for any distorting effects, and as a means of shortening the candidate list, reduced the list to one female and four male notes. The male note I decided to show (see below) was chosen on the basis that it contained both more 'about' (frequent) and 'distinctive' (key) semantic concepts than the other three. (Incidentally, the female-authored note shown below also has eight of the corpus-key semantic concept categories as key concepts in itself, and the male-authored note has nine.) So, by this process of combining the various aspects of what is typical in the notes, the two shown below are, arguably, the most prototypical notes in the corpus (despite the fact that the female-authored note does not contain any 'Negative' category terms).

Here are the two notes with the items representing the six chosen semantic categories, along with items from the 'Discourse Bin' and 'Negative' categories, in bold typeface. For clarity I have not put the 'Unmatched' or the many 'Pronouns' in bold. Apart from the anonymised names I have altered only the layout.

(7.58. **A female-authored note**)

"My dearest Bill,
      When this letter reaches you I **shall** no longer be here.
  May I **thank you** and **all** your **family** too for what you and they have done for me and my **parents** over the years. The strong **family** threads have never been broken in spite of a tug or two now and again on one side or the other.
    Who **knows**? Perhaps **one day** we may meet again in the **Hereafter**, when we **will really know** what our **grandparents** were like as well as our rascally relations.
    I cant help wondering if they have learned to live **more** peacefully on the

238

Other Side than on this - somehow I **rather** doubt it. But in **any** case, I am sure they must have livened the place up a lot.

I send you my **fond** love. May the days ahead bring deepest happiness and contentment to you and **all** those you **love**.

Again, many many thanks for everything
Mary" (f96460l).

(7.59. **A male-authored note**)

"To **Mom** and **dad**,

Well as I'm **not** here to speak I would **like** to tell you whats gone on.
I went back to my old ways in September do**n't know** why (maybe I thought I would hit it big but **no** I did**n't**. I just seem to think I'm a week person **so sorry** that I've let you down again.

I Hope you **will** forgive me for **any** pain and hurt I have caused you

all my love Bill
xx.

Maybe see you in the next life if there is one

P.S. if they havent found me I'm in Forest WOODS if I've got the nerve to do it
love Billy" (m98222a).

The words 'family', 'parents' and 'grandparents' in the female-authored note, and 'Mom' and 'dad' in the male-authored note are references to 'Kin'. The word 'fond' in the female note, and some of the instances of 'love' and 'like' in both notes, refer to the category 'Like' (see below). All instances of 'knows' and 'know' are in the 'Knowledgeable' category. The 'Degree: Boosters' are 'really' and 'more' in the female note and 'so' in the male one. Every instance of 'all' and 'any' in the notes is in the 'Entire: maximum' category. In the female note 'shall', 'one day', 'Hereafter' and 'will' are references to 'Time: Future'. In the male note 'will' is the only such reference. The items 'thank you', 'rather' (female note), 'no' and 'sorry' (male note) are in the 'Discourse Bin'; 'not' and the 'n't's (male note) are in the 'Negative' category. The 'Unmatched' items include the anonymous 'Bill', 'Mary' and 'Forest' (see stage 'b' in Section 4.5), and the apostrophe-less 'havent' in the male note. The 29 and 27 'Pronouns' in the female and male note, respectively, include the words 'my', 'you', 'I', 'what', 'who' and 'those'.

An important point is that the Wmatrix categories are quite subtle. In the female-

239

authored text above, the non-bold "love" is not in the category 'Like'. It is, in fact, to be found in another category, namely "Relationships: Intimacy and sex". In the male-authored text above, there is another non-bold "love" which is also in this category. One confusing feature of Wmatrix is that, while a difference of degree (the "-" and "+" parts of the codes referred to in Section 7.1.1) generally results in a change in the category name, this is not always the case. Codes E2+ and E2+++, for instance, are both labelled as 'Like'. The 'Like' category used and referred to in this Section is E2+ (see Figures 7.2.a, 7.2.b, 7.4.a and 7.4.b). Therefore, the word "dearest" in the female-authored note above is not shown in bold type because it is in category E2+++. Further, in the same female-authored note, other words may also appear to have been erroneously excluded from the set of emboldened words. However, these are not errors: one might expect the words "never" and "everything" to concern certainty or absoluteness (see Section 6.3.2, "Key Word Groups", where these words were assigned in this way), and to be in the 'Entire; maximum' category; but in fact "never" falls into the category 'Time', and "everything" appears in the category of 'Pronouns'. Additionally, one might expect the word "deepest" to be one of the 'Degree: Boosters', but it is actually in a category entitled "Long, tall and wide".

Interestingly, in the male-authored note, the word 'will' (in bold type) is in the category 'Time: Future'. It was seen (Section 6.3) that 'Future' was one of the possible ways of classifying it; although 'will' was ultimately (Section 6.3.2) used in a category about certainty. One could create alternative dictionaries for the semantic tagger, but the subtleties already exposed by the default dictionary in Wmatrix indicate important aspects of what suicide notes are about.

Another way of describing what suicide notes are about comes from Durkheim (2002) and was mentioned in Chapter 6. Although employing key words alone to classify the two notes above in terms of Durkheim's types of suicide is not appropriate (see Section 6.3), here we have two full texts, and one can attempt to match them to Durkheim's types, which I outlined in Section 3.3. The female-authored note can be termed Altruistic because of its concern for the addressee, together with its lack of melancholy, apathy, fatalistic or angry expressions. The male-authored note is a mixture of mainly

the Egoistic, with some Fatalistic and Altruistic components. It expresses some depression and sadness, and a degree of acknowledging the inevitable, but without any anger. It also shows some concern for others in acknowledging wrongs done to the addressees and asking their forgiveness. From this viewpoint, Durkheim's types can be thought of as describing sub-genres of suicide notes. However, whether adding together the features of all his types could be considered to define 'the suicide note genre' is another matter which I shall not pursue here.

In a somewhat circular argument, if a corpus of suicide notes can provide a prototype, a prototype suicide note should provide clues about the essence of the suicide note genre. According to Swales (2007, p.6), prototypes provide "Variable links to the centre [of a genre]", and are one of a half dozen metaphors that may help in the categorising of texts into genres (ibid., p.149). Along with 'category' and 'concept' (see above), 'genre' is another term I have used liberally in this thesis, but I shall now say just a little more about how prototypes relate to it.

While Bakhtin (1999, p.129) emphasises the 'typical', ("typical form of utterance… typical kind of expression… typical situations… typical themes… and… typical circumstances": see the beginning of this chapter), others extend this to the 'prototypical'. Touchon (1999, p.18) for example, provides a succinct definition of genre, although he is writing about the use of prototypes in the context of disciplines, teaching and learning:

> "A genre may be defined as a prototype that thematizes a type of utterance
> through links based on similarities."

Another reference (also, coincidentally, from the viewpoint of didactics) that links prototype and genre comes from Donahue (2009) who cites J. M. Adam's influential 1992 work, "Les textes". According to Donahue (2009, p.426), "Adam's genre and text type work, in the 1970s and 80s, was key to the entrenched relationship between text types and writing instruction in France. Adam single-handedly set the parameters for describing, teaching, and thinking about text types and prototypes for years." Donahue also states that Adam went on to focus upon

"… the flexibility we need when talking about text typologies and the importance of seeing a text as only more or less typical of the prototype in a particular genre..." (Donahue 2009, p.427).

So, at least one major European state makes practical use of prototypes for determining genre, seeing prototypes as crucial to understanding it.

7.4. Conclusion

What determines aboutness is difficult to pin down. It has been seen that there are difficulties in determining what proportion of a corpus may convey aboutness. Percentage frequencies in a corpus cannot have any great meaning unless one can ascertain that they are not peculiar to only a small proportion of the constituent texts. Further, determining what proportion of texts can be said to convey aboutness is largely a matter of judgement. Results pertaining to two texts in a couple of hundred are clearly not representative, but finding something common to 99% of texts probably is. The judgement involved is where to draw the line between 2 and 99. This concern regarding bias can be extended to subjects where a corpus contains more than one text per author. Similar judgements must be made when considering what proportion of subjects should be considered to be representative of aboutness.

But there is another facet to aboutness. If some semantic concept is found to be common across most of the texts in a corpus, is it reasonable to conclude that suicide notes are, at least partially, about that concept? I conclude that it is. However, extending concepts to single words, (rather than categories of words) is more problematic (see Chapter 6).

Simple word lists, as used in Chapter 6, can reveal something about aboutness, and they also give some information about speech and writing. However, word lists, including key word lists can be large and therefore cumbersome to deal with. When it comes to the pursuit of finding out what people write about, automatic semantic tagging is a better method, and Wmatrix is a very useful tool. The value of a standard reference corpus, such as the BNC, is also indisputable: it can help distinguish items that may be

peculiar to the corpus under investigation. There are, however, some problems and limitations with Wmatrix, particularly with the user interface and with the format required for its input, as described in Section 4.11. As mentioned in that Section, Wmatrix is not ideal because one cannot see the distribution of any findings with respect to the individual texts in a corpus, or to their authors. By combining software tools, such as simply counting 'grep' lines of the word lists produced by Wmatrix, it is possible to acquire a more accurate picture of what suicide notes are about.

Suicide notes are about many things, some of which make them distinctive from other texts. The things they are about that do not distinguish them include the following 15 Wmatrix categories: 'Grammatical bin', 'Existing', 'Likely', 'Personal names', 'Numbers', 'Getting and possession', 'General actions / making', 'Location and direction', 'Wanted', 'Moving, coming and going', 'Time: Period', 'Thought, belief', 'Speech acts', 'Exclusivizers/particularizers', and 'Evaluation: Good'.

Ten Wmatrix categories cover what suicide notes are about and also make them distinctive. These are: 'Pronouns', 'Like', 'Unmatched', 'Negative', 'Discourse Bin', 'Time: Future', 'Knowledgeable', 'Kin', 'Entire: maximum' and 'Degree: Boosters'. Suicide notes are distinctively about being affectionate, referencing the future, things their authors know, and their relatives. They may be further distinguished by their greater than normal use of pronouns, people's names and mis-spellings, negatives, discourse markers, maximum quantity terms, and intensifiers. These latter have not been investigated in any detail within the limitations of this thesis, but could prove fertile ground for further research.

8. CONCLUSION

8.1. Questions and Answers

In this chapter I review the aims of this thesis and re-present my research questions, as noted in Chapter 1, and in summarising the conclusions of my individual chapters I describe how I answered them.

The main aim of this thesis was to investigate the language of suicide notes, and to find out what kinds of things suicidal people see as important to write about. I was also interested in finding clues that would go some way to help solve the problem of how to distinguish fabricated suicide notes from genuine ones.

My main research questions (see Chapter 1) were:

1. What factors are involved in building a corpus of suicide notes?

2. What are the characteristics of the Birmingham corpus (from a mainly qualitative viewpoint)?

3. Are there ways to distinguish between the Shneidman sub-corpora of genuine and simulated suicide notes, in addition to those used by other researchers?

4. What are suicide notes about (from a mainly quantitative point of view)?

5. Is the Birmingham corpus demonstrably more like the Shneidman genuine sub-corpus or the Shneidman simulated sub-corpus?

The remainder of this Section describes how I answered the research questions.

1. What factors are involved in building a corpus of suicide notes?
I answered this question in Chapters 2, 3 and 4 where I described the confusing legal situation regarding suicide and suicide notes, provided a literature review and described

my data collection, respectively.

In Chapter 2 I gave some legal background to the subject of suicide notes. I discussed why defining suicide and suicide notes is problematic. I also looked at Coroners' Courts where (in England and Wales) the verdict of suicide is usually determined, and why such verdicts are problematic. Although having sympathy with the wide, all-inclusive, view, for the purposes of analysis I opted to accept (and count only) coroner's court verdicts of suicide. Further, I looked at the confusing reality of what happens to a suicide note following a suicide, and the problems of access to suicide notes. (In Chapter 2 I also examined the use of the term "suicide note" as it appeared in some legal judgments from non-coroner courts.)

In Chapter 3 I reviewed some of the most often cited studies of suicide notes. This review made it clear that collecting a corpus of fake suicide notes would not be tenable. However, it also pointed me towards the existing published set of such notes (Shneidman's), using a grounded research methodology, semantic tagging software, and to the need for a corpus of suicide notes.

In Chapter 4 I described the data collection process, and the problems I encountered when building my corpus. In the context of having data in my possession I reprised the thorny questions (from Chapter 2) of what should be counted as a suicide and what should be counted as a suicide note. I went on to explain the problems I had with the transcription process, and stated the various decisions that had to be made in this regard. I then described my data, and also looked at the subject of suicide note length.

2. What are the characteristics of the Birmingham corpus (from a mainly qualitative viewpoint)?
I answered this question in Chapter 5, although aspects of it were also addressed in Chapters 6 and 7.

In Chapter 5 I looked at the anatomy, the characteristics, of suicide notes. For each set of features I found in the (real) notes, I proceeded to look at the sub-corpus of simulated

notes, seeking similarities and differences. Within this remit, I looked at how the notes begin and end, some categories that were mainly (from a lexical viewpoint) salient in the notes, and some other categories, culled from various sources, that I termed "meta categories", since they were less salient (from a purely lexical viewpoint).

I found certain features tended to be present in real notes, and tended not to be present, at least to such a great extent, in simulated notes. Some of the characteristics of real suicide notes are that they may contain postscripts, dates, explicit expressions of love, fault and goodbye, 'who I am' statements, and 'where I am' statements. They are also likely to contain mentions of events and memories, and instructions and trivia. However, real notes are unlikely to contain expressions of not wanting to be a burden. Fake suicide notes, on the other hand, seem to contain few instructions, no trivia, and no postscripts.

3. Are there ways to distinguish between the Shneidman sub-corpora of genuine and simulated suicide notes, in addition to those used by other researchers?
I answered this question in Chapter 5. With respect to the levels of discourse, and most of the category concepts (if not their actual details) that were counted, the answer is negative. However, although it would not be true to say that the method I employed was wholly original – certainly its various facets were not – I believe that it has not been applied in the same way before.

In Chapter 5 I have looked at "oddness" in suicide notes. Beginning with intuition, rather than any particular linguistic method, I first looked at simulated suicide notes, and then compared them to genuine notes. The types of "oddness" I found prevalent in the simulated suicide notes involved roles and naming references, peculiarities in phraseology, illogical semantics, vagueness, and melodrama. Indeed, I found that a generalisation can be made that oddness will occur in fake suicide notes, as opposed to genuine notes. However, this is only a generalisation because it is possible to find fake notes in which there is no oddness and genuine notes in which there is oddness. Further, the definition of certain types of oddness is quite subjective; and although finding several instances of oddness in one text would be convenient, there may be only

one in any given text. Thus oddness cannot be used as the perfect tool for distinguishing real from fake suicide notes, although it should be possible to use it as a 'rule of thumb'.

There are actually several findings in this thesis that could be considered to be 'rules of thumb': there are several features whose presence or absence will tend to indicate genuineness or fakeness without any one of them being definitive. What everybody, especially the forensic linguist, desires is a litmus test for genuineness or fakeness. I have not found a litmus test because there is probably no absolute means of distinguishing real from fake suicide notes. From an overall perspective – not just that of oddness – it seems that the difference between genuine and simulated notes can only be expressed in relative terms and probabilities.

4. What are suicide notes about (from a mainly quantitative point of view)?

I answered this question in Chapters 6 and 7, although aspects of it were also addressed in Chapter 5.

In this thesis I have shown a worthwhile application of corpus linguistics. In Chapter 7 I established what typical suicide notes look like, and how prototypicality is one way of exemplifying the suicide note genre. In Chapters 6 and 7 I considered four aspects of 'aboutness', namely, lexical content, keyness, semantic categories and genre. I began by considering various word lists and key word lists in Chapter 6, and created some context-free sets of words from the lists. I then looked at semantic categories and prototypicality in Chapter 7, where I performed some semantic tagging with the Wmatrix software. (Wmatrix was able to produce categories of semantic concepts which took account of word and phrase context.) With the aid of additional software I was able to obtain a more reliable picture of 'aboutness' by seeing how words and concepts were distributed amongst the individual texts, and between the individual authors of those texts, within the corpus. Chapter 7 shows the cumulative significance of these different approaches: for example, it describes a semantic category containing lexis common to the beginnings and endings of the texts (investigated in Section 5.2) which reflects some structural aspects of suicide notes.

There are indeed some things that suicide notes are 'about', and these can distinguish them from other texts. There are ten Wmatrix categories, namely, 'Pronouns', 'Like', 'Unmatched', 'Negative', 'Discourse Bin', 'Time: Future', 'Knowledgeable', 'Kin', 'Entire: maximum' and 'Degree: Boosters'. These reflect both what suicide notes are about and also what makes them distinctive from other texts. Suicide notes are 'distinctively' and 'about' affection, references to the future, their authors knowledge, and their authors relatives. Additionally, suicide notes are distinguished by their greater than normal use of pronouns, people's names and mis-spellings, negatives, discourse markers, maximum quantity terms, and intensifiers. In Chapter 7 I presented two prototypical (and real) suicide notes, from the Birmingham corpus, which illustrate the major semantic categories I found.

5. Is the Birmingham corpus demonstrably more like the Shneidman genuine sub-corpus or the Shneidman simulated sub-corpus?

Chapter 5 answers this question, in two different ways. From the second part of Chapter 5 it is clear that the Birmingham corpus is indeed more like Shneidman's sub-corpus of genuine suicide notes than his sub-corpus of simulated notes. On the basis of "oddness" there is a clear distinction. However, in many other respects the Birmingham "His" and "Hers" sub-corpora are more similar to each other than they are to either of the Shneidman sub-corpora, and this is borne out by the first part of Chapter 5.

It cannot be ruled out that the latter observation might be due to the fact that one corpus is composed of texts written in the 1940s whereas the other is composed of texts written in the 1990s; or that one corpus is made up of texts written by people in the U.K. whereas the other consists of texts written by people in the U.S.A. There might also be various social and cultural differences between Birmingham, England and Los Angeles, America that affected the note-writers' texts. Additionally, it could be argued that there might be differences between the corpora that are related to gender because the Shneidman notes were written exclusively by males, whereas the Birmingham notes were written by females as well as males. Further, it must be noted that differences could be due to one corpus being much larger than the other.

One implication of all this is that cultural differences exist and are of vital importance when finding language differences between such corpora. Another is that, somehow, fake suicide notes should be created and collected so that future researchers can use them. Clearly, ethical as well as practical considerations would need to be taken into account in creating such a corpus; but the difficulties should not be insuperable.

8.2. More Questions and Future Work

As indicated in Chapter 1, my work raises questions for both the forensic linguistics and the corpus linguistics communities concerning the consistency of texts across time and across space and the homogeneity of corpora, particularly regarding the relative sizes (in word length) of their constituent texts. I have also raised questions about how to deal with some difficult categories such as 'topic' in situations such as suicide note corpora. And I have discussed the importance of analysing phenomena in corpora on the basis of how diversely they are spread across the component texts.

It seems to be easier to define the central and typical suicide note rather than to establish the borderline between the real and the simulated. However, this raises questions in respect of forensic linguistics in general, such as how this may relate to distinguishing real from fake in other (non-suicide note) areas of research. One of my original aims (see Chapter 1, research question 3) was to see whether there were "ways to distinguish between the Shneidman sub-corpora of genuine and simulated suicide notes". Distinguishing real from fake is something that forensic linguists, such as Picornell (2003), often engage in. Forensic linguists commonly deal with disputed texts and, when available, genuine texts for comparison. Examples of texts that are subject to forensic linguistic study in this context include disputed confessions and police statements.

One of the things my thesis has demonstrated is the extent to which this discriminating could be done, and the extent to which it cannot be done. It may be possible, for example, to express a view about the central and typical genuine statement and the central and typical fabricated statement, although it will probably never be possible to

distinguish accurately between the two given a single instance. However, for future work the items used by Picornell (2003) to distinguish truth from fiction, including combinations of function words and words expressing anger, could be applied to suicide notes.

Future work on suicide notes could also look at the types of texts I found in the Coroner's Office that were classified under verdicts of "Other… suicide". These notes may, or may not, differ from the notes classified as "Killed Himself" or "Killed Herself" in some linguistic way. Future work could also look at transcriptions of suicide notes made on audio or video tapes. These might contain interesting differences from written suicide notes. An even more interesting exercise would be comparing notes from "Open" verdicts - which may or may not be genuine suicide notes - with undisputed suicide notes, and notes which are known to be simulated or fake, because it might even be possible to suggest to the coroner that some "Open" verdicts should be revisited.

In view of the similar results for the males and females in the most frequent word and key word lists (see Chapter 6), and given that suicide notes are an 'occluded genre', the existence of a cognitive structure should be investigated, as mentioned in Section 6.5. Additionally, the way that the suicide itself is referred to by the note-writers is worthy of future research as was suggested in Section 6.3.

The issue of intent has been mentioned several times in this thesis. It would be worth investigating exactly what written intent looks like, and the extent to which suicide notes deal with it, both directly and indirectly. If the live writer projects to the position of post-dying, does this signal intent to die? If such projections manifest themselves in instructions, then what are the implications with respect to multiple-note writers? A comparison of multiple-note writers with single-note writers is another area where future research could be fruitful. In addition to leaving more instructions than single-note writers (see Section 5.4), multiple-note writers clearly have a way of dealing with the issue of multiple audiences (see Section 5.7.1). The linguistic cues signalling single, as opposed to multiple, addressees are themselves also worth further research.

Future work on suicide notes could also look at how suicide notes are used by their addressees, compare the results of different semantic taggers, and investigate texts written by people in supposedly similar states of mind to suicides, such as informal wills; texts from the terminally ill and death-row prisoners; and texts from people planning acts of terrorism or murder as well as suicide (Blackwell, Meijs and Shapero 2007). Forensic linguistics would benefit from studies of a far wider range of suicide note corpora than that used by existing studies. Forensic linguistics also needs more research into the probabilities of various phenomena occurring, and this includes the probabilities of suicide notes being real or fake.

Further to the above, future work could also investigate correlations with non-linguistic variables by looking at the distribution of any linguistic findings between various combinations of suicide note-writers' methods, ages, reasons, etc. While this is not a novel suggestion, modern database software would make it relatively simple to include many such attributes as a matter of course.

APPENDIX A: Oddness in Simulated Suicide Notes

*The asterisked items appear within the text of Chapter 5.

A.1. Naming References (10 occurrences in 9 notes)
"I am very proud of *our son,…*" [My italics] (Sim02b, Shneidman & Farberow, 1957e).*

"Have *the kids* remember me,…" [My italics] (Sim09a).

"Be happy and all my love always to you and *our three*." [My italics] (Sim13b).

"Continue with your will to live, fill any emptiness with your love for *our children,…*" [My italics] (Sim16b).

"My continued living,… would be a continuing depressant, because of my condition, upon you and *the kids*." [My italics] (Sim18a).

"My undying love to you and *the kids*." [My italics] (Sim18a).

"Tell *the kids* I…" [My italics] (Sim21b).

"To *my kids*- This…" [Sic] [My italics] (Sim27a).

"With money from our Life Insurance program you will be able to give *the children* and yourself…" [My italics] (Sim28b).

"I am sorry I was such a trial to you and *the children,…*" [My italics] (Sim29a).

A.2. Naming References – Wife/Husband (7 wife +2 husband occurrences in 8 notes)
"Goodbye dear *wife*." [My italics] (Sim08b).

"Dear *Wife*;…" [My italics] (Sim17a).

"My dear *wife*:…" [My italics] (Sim18a).

"Dearest *Wife*." [My italics] (Sim19a).

"Dearest *Wife*:" [My italics] (Sim22a).

"Dear *wife*." [My italics] (Sim28b).

"Dear *Wife*." [My italics] (Sim30a).

"Your Desperate *Husband*, Bill Smith" [My italics] (Sim23b).

"Your ever loving *husband*" [My italics] (Sim28b).*

A.3. Phraseology (2 notes)
"I cannot stand *the suffering* any longer." (Sim08b).*

 "My health has broken and I no longer feel that I can be of help in *the Support* of *the family…*" (Sim30a).*

A.4. Logic (12 occurrences in 9 notes)
"Friday I lost *the job I have held for the past seven years.* When I told *my wife…*" [My italics] (Sim01b).*

"I am very proud of our son, and his high potential in *his chosen field* for which he has real talent." [My italics] (Sim02b).*

"I am tired of living so *I decided to end it all, hope this will not distress anybody.*" [My italics] (Sim03b).

"*I don't know why I am doing this* unless my reasoning has gone all to pot." [My italics] (Sim07a).*

 "As we both might reasonable [Sic] recognize *this is not the right way to solve any situation. …* The will is gone, *reason is gone*, there is only one answer." [My italics] (Sim16b).

"I am sorry to cause you this *embarrassment* but I can't *seem to* stand life this way." [My italics] (Sim17a).

"*I am writing this to explain why* I am going to end it all. I know that this is a cowardly way and I am sorry but I just haven't the will to do otherwise." [My italics] (Sim19a).*

"Dearest *Wife*:… I hope you will be *married* as soon as it is proper to be,…" [My italics] (Sim22a).*

"*To the police.* please [Sic] *tell family* that I love them why say more." [My italics] (Sim33a).*

A.5. Vagueness (5 occurrences in 5 notes)
"*Things* have become so *uncertain* and unbearable," [My italics] (Sim09a).

"I've taken care of *everything*." [My italics] (Sim14b).

"I can see no reason to battle *the elements of life* any longer…" [My italics] (Sim25a).

"*Things* are not going right and don't look as though they ever will." [My italics] (Sim26b).

"I'm tired of *it all.*" [My italics] (Sim32a).*

A.6. Melodrama (8 occurrences in 5 notes)

"*Not being of sound mind* I have decided to leave this world *by electrocuting* myself." [My italics] (Sim05a).

"*This is the end. Ive had enough. Cant take anymore*." [Sic] [My italics] (Sim10a).

"I am Sorry Mary But i just Can't Stand *Life Any Longer*." [Sic] [My italics] (Sim11b).

"*This world is too cruel* for me. I am in search of peaceeternal [Sic] peace…" [My italics] (Sim23b).

"Mother meant good, but *she drove me to my grave*." [My italics] (Sim23b).

"As I sit here *with this gun in my Hand*, which in a few minutes I will take my life I am thinking of all the wonderful *minutes, Days, years*, I have spent with you." [Sic] [My italics] (Sim31b).*

APPENDIX B: Oddness in Real Suicide Notes

*The asterisked items appear within the text of Chapter 5.

B.1. Naming References
"Please be good to little Betty, our daughter, I love her so." (Gen07b, Shneidman & Farberow, 1957e).*

"I love you so & my Fppp & children." (Her, F95389a).

B.2. Naming References – Wife/Husband
"I leave everything which has all been acquired since we were married to you my darling wife-Mary Smith…" (Gen05b).

"My Dearest Wife:…" (Gen07b).

"To my wife Mary:…" (Gen24b).

"Darling wife, Mary Helen Smith…" (Gen28a).
"So darling this is your divorce my darling wife Mary. " (Gen28a).
"I wish you get the rings back my dear wife. " (Gen28a).
"Goodby my dear wife Mary. " (Gen28a).

"Your husband, William H. Smith" (Gen31a).

B.3. Logic
"I can't explain why I have done this…" (Her, F98504).

"Sorry about this, but you are the only one I can ask…" (Her, F99123b).

"please help me?? Im so numb[?]" (Him, M98523c).

B.4. Signature
"I really   Six Boots   R   WORD" (Her, F95456a).

B.5. Repetition
"I LOVE YOU ALL MORE THAN YOU'L EVER KNOW, IV WORKED ALL MY LIFE, IV ALWAYS TRIED TO DO MY BEST, BUT IT NEVER SEEMED TO BE ENOUGH… BUT SHE HAS TREATED ME LIKE A LUMP OF SHIT…
[+ 5 unique 'sentences' in 3 paragraphs]
I LOVE YOU ALL MORE [?] YOU'L EVER KNOW, IV WORKED [?] MY LIFE, BUT YOUR MOM [?] TREATED ME LIKE A LUMP OF SHIT...

I LOVE YOU ALL MORE THAN YOU'L EVER KNOW, I HAVE REALY TRIED TO DO THE BEST I COULD…
[+ 2 unique 'sentences' in 1 paragraph]
I HAVE ALWAYS TRIED TO TO BEST I COULD , YOUR MOM HAS OLYaYS

TEREAED LIKE A LO" (Him, M97301).

"good bye good bye good bye good bye good b
L=20   CH=1   J=L NB" ([A computer print out, repeated 3 times] Him, M98542).*

B.6. Name & Pronouns
"Fnnn has made sure Mqqq can not move in to house, has not revealed all of his savings and his share's…" (Him, M95468g).*

"Fnnn has made sure that Mqqq can not get out of this mess he is in… he is on the street" (Him, M95468g).

BIBLIOGRAPHY

1. General References

Adam, J. M. (1992). <u>Les textes: Types et prototypes</u>. Paris: Nathan.

Akmajian, A., Demers, R.A., Farmer, A.K. & Harnish, R.M. (1992). Linguistics, An Introduction to Language and Communication. (3rd Ed.) London & Cambridge, Massachusetts: The MIT Press.

Allen, Richard. (2000). "Statistics of Deaths Reported to Coroners: England and Wales 1999". <u>Home Office Statistical Bulletin</u>. Issue 8/00, 20 April, 2000. London: Government Statistical Service, The Home Office.

Allison, Rebecca. (2003). "Suicide verdict on man shot by police". <u>The Guardian</u>. [Home Pages.] May 10, 2003. p.2.
http://web.lexis-nexis.com/professional/... [Accessed May 16, 2003.]

Anderson, Clive. (2002). "Unreliable Evidence". <u>BBC Radio 4</u> Broadcast Tue. May 21st, 2002, 9.30pm (& 9am). [Guest participants: Dr. William Dolman (coroner), Lord Justice Sir Anthony Clark (judge), Tom Luce (senior civil servant), & Samantha Leek (barrister).]

Anon. (n.d.). "A Brief History". <u>The Coroners' Society of England and Wales</u>.
http://www.coroner.org.uk/public/history.asp [Accessed 16 July, 2004.]

Anon. (1996). <u>The Work of The Coroner: Some Questions Answered</u>. [A Home Office Guide.] 100M/3/96. U.K.: The Home Office.

Anon. (2000). "Notes from the other side". <u>The Sunday Telegraph</u>. ['Review'.] Sun. January 23, 2000. p.1-2.

Anon. (2006a). "Q&A: The Deepcut mystery". <u>BBC NEWS</u>. 29/03/2006, p.1-2.
http://news.bbc.co.uk/go/pr/fr/-/2/hi/
uk_news/3179985.stm [Accessed September 1, 2006.]

Anon. (2006b). <u>Draft Coroners Bill</u>. [Contents.] 12 June 2006. p.1-2. dca (Department for Constitutional Affairs).
http://www.dca.gov.uk/legist/coronersreform.htm [Accessed September 1, 2006.]

Anon. (2007). "The Court Structure of Her Majesty's Court Service (HMCS)".
http://www.hmcourts-service.gov.uk/aboutus/structure/index.htm [Accessed August 28, 2010.]

257

Arbeit, S.A. & Blatt, S.J. (1973). "Differentiation of simulated and genuine suicide notes". Psychological Reports. 33(1), p.283-297.

Archer, D. Wilson, A. & Rayson, P. (2002). "Introduction to the USAS Category System". Rayson, P. (2007). Wmatrix: a web-based corpus processing environment. Computing Department, Lancaster University. http://ucrel.lancs.ac.uk/ucas/ucas%20guide.pdf [Accessed August 23, 2007.]

Atkinson, J. Maxwell. (1978). Discovering Suicide: Studies in the Social Organization of Sudden Death. London & Basingstoke: The Macmillan Press Ltd.

Bakhtin, M.M. (1999) "The Problem of Speech Genres". In Jaworski, A. & Coupland, N. (Eds.) The Discourse Reader. London & New York: Routledge.

Bank of English, 450 million word version, jointly owned by HarperCollins Publishers and the University of Birmingham.

Barraclough, Brian and Harris, E.Clare. (2002). "Suicide preceded by murder: the epidemiology of homicide-suicide in England and Wales 1988-92". Psychological Medicine. Vol.32, p.577-584. [Online journal. Accessed May 2005.]

Bell, Allan, (1984) "Language Style as Audience Design". Language in Society. Vol.13, No.2, June 1984, p.145-204. http://www.jstor.org/stable/pdfplus/4147516.pdf [Accessed 2 June 2009.]

Bird, Steve. (2000). "999 call operator heard gun suicide on phone". The Times. Thursday November 30, 2000. p.9.

Black, Stephen T. (1993). "Comparing Genuine and Simulated Suicide Notes: A New Perspective". Journal of Consulting and Clinical Psychology. Vol.61, No.4, p.699-702.

Black's Law Dictionary with Pronunciations. (6th Edition; Centennial Edition, 1891-1991.) St. Paul, Minn.: The Publisher.

Blackwell, Susan; Meijs, Willem; and Shapero, Jess. (2007). "Texts of Murder and Martyrdom". In Kredens, Krzysztof & Goźdź-Roszkowski, Stanisław. (Eds.) Language and the Law: International Outlooks. p.395-420. [Łódź Studies in Language 16.] Frankfurt am Mein: Peter Lang.

Blatchley, Nick. (2006). Suicide and Open Verdict Deaths in the UK Regular Armed Forces 1984-2005. National Statistics, Statistical Notice. p.1-8. Bath, U.K.: DASA (Defence Analytical Services Agency). http://www.dasa.mod.uk/publications/pdfs/suicide/suicideMar06.pdf [Accessed September 1, 2006.]

Bourne, S.R. (1983). The UNIX System. London: Addison-Wesley.

Bresler, Scott; Scalora, Mario J.; Elbogen, Eric B.; & Moore, Y. Scott. (2003). "Attempted suicide by a cop: a case study of traumatic brain injury and the insanity defense". Journal of Forensic Sciences. 48(1): 190-194. http://journalsip.astm.org/JOURNALS/FORENSIC/PAGES/4015.htm & http://journalsip.astm.org/DOWNLOAD/JFS2001045_481.28322.pdf [Accessed February 7, 2003. (Pages mapped 1:190 etc.)]

Buvac, Vanja & Stone, Philip. (2001). "The General Inquirer User's Guide". p.1-7. http://www.wjh.harvard.edu/~inquirer/ (manual.pdf) [Accessed September 15, 2003.]

Capstick, A. (1960). "Recognition of emotional disturbance and the prevention of suicide". British Medical Journal, 1, p.1179-1182.

Casey, C. (1993). "Making Murder look like Suicide". Police Review. 6th August, p.16-18.

Chomsky, N. (1972). Language and Mind. Enlarged Edition. New York: Harcourt Brace Jovanovich. In Laurence (2003).

Collins Concise English Dictionary. (1993). (3rd Edition.) Harper Collins Publications.

Coulthard, Malcolm. (1992). "Forensic Discourse Analysis". In Coulthard, M. (Ed.) Advances in Spoken Discourse Analysis. Ch.12. p.242-258. London: Routledge.

Coulthard, Malcolm. (1994). "Powerful evidence for the defence: an exercise in forensic discourse analysis". In Gibbons, John (Ed.) Language And The Law. Ch.17, p.414-427. London: Longman.

Coulthard, Malcolm & Johnson, Alison. (2007). "The work of the forensic linguist". An Introduction to Forensic Linguistics: Language in Evidence. Ch.6, p.121-143. London & New York: Routledge.

Cummings, H.W., & Renshaw, S.L. (1976). SLCA-II: A computerized technique for the analysis of syntactic language behaviour. Minneapolis: Burgess.

Cummings, H.W., & Renshaw, S.L. (1979). "SLCA-III: A metatheoretic approach to the study of language". Human Communication Research. Vol.5, p.291-300.

Darbonne, A.R. (1969). "Suicide and age: A suicide note analysis". Journal of Consulting & Clinical Psychology, 33, p.46-50.

Davis, T. (1994). "ESDA and the analysis of contested contemporaneous notes of police interviews". Forensic Linguistics: The International Journal of Speech, Language and the Law. Vol.1, No.1, p71-89.

Dollard, J. & Mowrer, O.H. (1947). "A method of measuring tension in written documents". J. abnorm. soc. Psychol. 42, p.3-32.

Donahue, Tiane. (2009). "Genre and Disciplinary Work in French Didactics Research". In Bazerman, C., Bonini, A. & Figueiredo, D. (Eds.) Genre in a Changing World. Ch.21, p.424-441. West Lafayette, Indiana: Parlor Press. Colorado State University,
http://wac.colostate.edu/BOOKS/GENRE/genre.pdf [Accessed 26 October, 2010.]

Durkheim, É. (1897). Le Suicide: Étude de sociologie. Paris.

Durkheim, É. (2002). Suicide: A Study in Scociology. Translation of Durkheim (1897). Translated by Spaulding, John A. & Simpson, George. Edited with an introduction by George Simpson. London & New York: Routledge.

Edelman, Ann M. & Renshaw, Steven L. (1982). "Genuine versus Simulated Suicide Notes: An Issue Revisited Through Discourse Analysis". Suicide and Life Threatening Behaviour. 12(2), 103-113.

Etkind, Marc. (1997). ... Or Not To Be: A Collection of Suicide Notes. The Berkley Publishing Company, New York: Riverhead Books.

Eysenck, M.W. & Keane, M.T. (1993). Cognitive Psychology: A Student's Handbook. Hove & London: Lawrence Erlbaum Associates Ltd.

Fairthorne, R.A. (1969). "Content analysis, specification and control". Annual Review of Information Science and Technology. 4 1969, p.73-109.

Ford, Richard. & Moseley, A. L. (1963). "Motor Vehicular Suicides". Journal of Criminal Law, Criminology and Police Science. 54(3), p.357-9.

Fox, Gwyneth (1993). "A comparison of 'Policespeak' and 'Normalspeak': A preliminary study". In Sinclair, John M., Hoey, Michael & Fox, Gwyneth (Eds.) Techniques of Description: Spoken and Written Discourse. A festschrift for Malcolm Coulthard. p.183-195. London and New York: Routledge.

Frederick, Calvin J. (1969). "Suicide notes: A survey and evaluation". Bulletin of Suicidology. 2, p.17-26.

Gibbons, John. (2003). Forensic Linguistics: An Introduction to Language in the Justice System. Oxford, U.K.: Blackwell Publishing Ltd.

Giles, S. (2007). The Final Farewell: Using a Narrative Approach to Explore Suicide Notes as Ultra-Social Phenomena. November 2007. Ph.D. thesis. University of Liverpool, U.K.

Glidewell, Iain (Right Honourable Lord Justice). (1986). "Preface". In Matthews & Foreman.

Goffman, E. (1981). Forms of Talk. Philadelphia: University of Pennsylvania Press.

Gottschalk, L.A. & Gleser, G.C. (1960). "An analysis of the verbal content of suicide notes". British Journal of Medical Psychology, 33. p.195-204.

Gregory, Adam. (1999). "The Decision to Die: The Psychology of the Suicide Note". In Canter, D. & Alison, L. (Eds.) Interviewing and Deception. Offender Profiling Series, Volume 1, Ch.7, p.129-156. Aldershot: Ashgate Publishing.

Grice, H.P. (1975). "Logic and Conversation". In Cole, P. & Morgan, J.L. (Eds.) Syntax and Semantics. Vol.3, Speech Acts, p.41-58. New York: Academic Press. [Reprinted in Jaworski, A. & Coupland, N. (Eds.) (2004). The Discourse Reader. p.76-88. London: Routledge.]

Houston, Dr Muiris. (2003). "Final Notes found to be far from despairing." In Irish Times, 8th March, 2003, p.1.

Hristov, Vladislav P. (2008). "Duplicate Lines Finder". [Software application.] WonderWebWare.com.

Hulsen, Tim. (2008). "Welcome at VennDiagram.tk!: Create". http://www.venndiagram.tk/. [Accessed 11th May 2008.]

Hunston, Susan. (2000). "Summing Up". In Hunston, Susan (Ed.) Pattern Grammar. A corpus-driven approach to the lexical grammar of English. Ch.9, p.247-274. London: John Benjamins Publishing Co. http://site.ebrary.com/lib/bham/Doc?id=5000193 [Accessed September 23, 2008.]

Jacobs, Jerry. (1967). "A Phenomenological study of suicide notes". Social Problems, 15, p.60-72.

Jefferson, Gail (Ed.) (1989). Harvey Sacks Lectures 1964-1965. Dordrecht, Boston, London: Kluwer Academic Publishers.

Johnson, Alison. (1997). "Textual Kidnapping - a case of plagiarism among three student texts?" Forensic Linguistics. 4.2, p.210-225.

Kelly, E. & Stone, P.J. (1975). Computer Recognition of English Word Senses. North-Holland Publishing.

Kilgarriff, Adam. (1997). "Using Word Frequency Lists to Measure Corpus Homogeneity and Similarity between Corpora". Proceedings of the 5th ACL Workshop on Very Large Corpora, Beijing and Hong Kong, August 1997. p.1-15. http://www.itri.brighton.ac.uk/techreports/ITRI-97-07_abs

[Accessed October 11, 2002.] &
link from http://www.itri.bton.ac.uk/techreports/
to ftp://ftp.itri.bton.ac.uk/reports/ITRI-97-07.ps.gz [Accessed May 2, 2003.]

Laurence, S. (2003). "Is Linguistics a Branch of Psychology?". In Barber, A. (Ed.) (2003). <u>Epistemology of Language</u>. Oxford University Press. p.69-106. http://www.philosophy.dept.shef.ac.uk/papers/IsLingPsych.pdf [Accessed January 19, 2011.]

Leech, G., Rayson, P., Wilson, A. (2001). <u>Companion Website for: Word Frequencies in Written and Spoken English: based on the British National Corpus</u>. http://www.comp.lancs.ac.uk/ucrel/bncfreq/ [Accessed December 25, 2007.]

Leenaars, A.A. (1988). <u>Suicide Notes: Predictive Clues and Patterns</u>. New York: Human Sciences Press.

Lester, D. & Leenaars, A.A. (1988). "The moral justification of suicide in suicide notes". <u>Psychological Reports</u>. 63(1), p.106.

Linsley, K.R.; Schapira, K; & Kelly, T.P. (2001). "Open verdict v. suicide - importance to research". <u>British Journal of Psychiatry</u>. 178, p.465-468. http://bjp/rcpsych.org/cgi/reprint/178/5/465 [Accessed March 25, 2006.]

Longman Lexicon of Contemporary English. (1981). McArthur, Tom. Harlow: Longman.

Luce, Tom. (2002a). "The Public Inquest". In Luce, Tom. <u>Review of Coroners and death certification in England, Wales and Northern Ireland</u>. [The Consultation Document.] Chapter 5. www.coronersreview.org.uk/docs/Coroner_Report.pdf [Accessed May 9, 2003.]

Luce, Tom. (2002b). "The Judicial Investigation of Deaths". In Luce, Tom. <u>Review of Coroners and death certification in England, Wales and Northern Ireland</u>. [The Consultation Document.] Chapter 4. www.coronersreview.org.uk/docs/Coroner_Report.pdf [Accessed May 9, 2003.]

Martin, J.R. & White, P.R.R. (2005). "Engagement and Graduation: Alignment, Solidarity and the Construed Reader". Ch.3, p.92-160. In <u>The Language of Evaluation: Appraisal in English</u>. Basingstoke: Palgrave Macmillan.

Matthews, Paul & Foreman, J.C. (1986). <u>Jervis on the Office and Duties of Coroners: with forms and precedents</u>. 10th Edition. London: Sweet & Maxwell. [1st Edition by Jervis, Sir John published 1829.]

McArthur, T. (1981). <u>Longman Lexicon of Contemporary English</u>. Longman, London.

McClelland, L., Reicher, S. & Booth, N. (2000). "A Last Defence: The Negotiation of Blame Within Suicide Notes". Journal of Community & Applied Social Psychology. Vol.10, p.225-240.

Medvedik, J.S. (2006). "TED Notepad". [Software application.] http://jsimlo.sk.

Microsoft Corporation. (1987). "MEGREP.EXE". Appendix B.1, Support Programs for the Microsoft Editor. (p.111-112). In Microsoft Editor for MS OS/2 and MS-DOS Operating Systems: User's Guide. [In Microsoft C 5.1 Optimizing Compiler: CodeView and Utilities, Microsoft Editor, Mixed-Language Programming Guide.] Microsoft Corporation.

Mowrer, O.H. (1953). Psychotherapy: Theory and Research. New York: The Ronald Press Company.

O'Donnell, Ian; Farmer, Richard; & Catalan, Jose. (1993). "Suicide Notes". British Journal of Psychiatry. 163: 45-48.
http://bjp.rcpsych.org/cgi/reprint/163/1/44 [Accessed October 8, 2007.]

Ogilvie, D.M., Stone, P.J. & Shneidman, E.S. (1966). "Some characteristics of genuine versus simulated suicide notes". In Stone, P.J., Dunphy, D.C., Smith, M.S., Ogilvie, D.M. and Associates. The General Inquirer: A Computer Approach to Content Analysis. Cambridge, Massachusetts: MIT Press.

Olsson, John. (2004). Forensic Linguistics: An Introduction to Language, Crime and the Law. London & New York: Continuum.

Osgood, C.E. & Walker, E.G. (1959). "Motivation and language behaviour: A content analysis of suicide notes". Journal of Abnormal & Social Psychology, 59, p.58-67.

Palermo, George B. (1997). "The Berserk Syndrome: A Review of Mass Murder". Aggression and Violent Behaviour. Vol.2, No.1, p.1-8. [Online journal. Accessed May 2005.]

Peek, Laura. (2003). "First 'suicide by cop' verdict on man shot in siege". The Times. [Home news.] May 10, 2003. p.1. http://web.lexis-nexis.com/professional/... [Accessed May 16, 2003.]

Pestian, J.P.; Matykiewicz, P.; Grupp-Phelan, J.; Lavanier, S.A.; Combs, J.; & Kowatch, R. (2008). "Using Natural Language Processing to Classify Suicide Notes". BioNLP. June 2008, p.96-97.

Phillips, Martin. (1989). Lexical Structure of Text. Discourse Analysis Monograph No.12, English Language Research. Birmingham U.K.: The University of Birmingham.

Picornell, I. (2003). "It is Not What We Say but How We Say It". <u>The White Paper</u>. May/June 2003. Austin, Texas: The Association of Certified Fraud Examiners. http://www.qed.info/whitepaper.html [Accessed 22 September, 2009.]

Rayson, P. (2007). Wmatrix: a web-based corpus processing environment. Computing Department, Lancaster University. <u>http://www.comp.lancs.ac.uk/ucrel/wmatrix/</u>

Rayson, P., Archer, D., Piao, S. L., McEnery, T. (2004). "The UCREL semantic analysis system". In <u>Proceedings of the workshop on Beyond Named Entity Recognition Semantic labelling for NLP tasks in association with 4th International Conference on Language Resources and Evaluation (LREC 2004), 25th May 2004, Lisbon, Portugal</u>. pp. 7-12. http://www.comp.lancs.ac.uk/computing/users/paul/publications/usas_lrec04ws. pdf [Accessed 30 November, 2010.]

Rosch, E. (1973). "On the internal structure of perceptual and semantic categories". In Moore, T. (Ed.) <u>Cognitive development and the acquisition of language</u>. New York: Academic Press.

Rosch, E. & Mervis, C. (1975). "Family resemblance studies in the internal structure of categories". <u>Cognitive Psychology</u>. 7, p.575-605.

Salib E.; Cawley S.; Healy R. (2002). "The significance of suicide notes in the elderly". <u>Aging and Mental Health</u>. May 1, 2002. Vol.6, No.2, p.186-190. Carfax Publishing.

Salib, Emad & Maximous, Joseph. (2002). "Intimation of intent in elderly fatal self-harm: do the elderly who leave suicide notes differ from those who do not?" <u>International Journal of Psychiatry in Clinical Practice</u>. September 1, 2002. Vol.6, No.3, p.155-161. Martin Dunitz.

Scollon, R. & Scollon, S.W. (2003). <u>Discourses in Place: Language in the material world</u>. London & New York: Routledge.

Scott, Kirsty. (2003). "Woman shot by police left suicide note". <u>The Guardian</u>. Tuesday August 19, 2003. http://www.guardian.co.uk/gun/Story/0,2763,1021515,00.html [Accessed 14 October, 2003.]

Scott, Mike. (1998a). <u>WordSmith Tools. Version 3</u>. Oxford University Press. http://www.lexically.net/wordsmith [Accessed circa 2001.]

Scott, Mike. (1998b). <u>Wordsmith Tools Manual. Version 3.0</u>. Oxford University Press. http://www.lexically.net/wordsmith [Accessed circa 2001.]

Scott, Mike. (1998c). "Written [sub-corpus]". <u>British National Corpus (BNC)</u>. http://www.lexically.net/wordsmith [Accessed circa 2001.]

Scott, Mike. (2007a). WordSmith Tools. Version 4, Demonstration version. http://www.lexically.net/wordsmith [Accessed January 2008.]

Scott, Mike. (2007b). Wordsmith Tools Manual. Version 4.0. Oxford University Press. http://www.lexically.net/wordsmith [Accessed January 2008.]

Scott, Mike. (2007c). "BNC_World.lst". British National Corpus (BNC). http://www.lexically.net/wordsmith [Accessed circa 2007.]

Scott, Mike. (2008). WordSmith Tools. Version 5.0, Demonstration version. http://www.lexically.net/wordsmith [Accessed September 2010.]

Shneidman, E.S. (1976). "Suicide Notes Reconsidered". In Shneidman, E.S. (Ed.). Suicidology: Contemporary Developments. Ch.8, p.253-278. New York: Grune and Stratton.

Shneidman, E.S. (1979). "An Overview: Personality, Motivation, and Behaviour Theories". In Hankoff, L.D. & Einsidler, Bernice (Eds.). Suicide: Theory and Clinical Aspects. Ch.12, p.143-163. Massachusetts: P.S.G. Publishing Company Inc.

Shneidman, E.S. (1999). "Letters of Enforced Death Versus Suicide Notes". In Leenaars, A.A. (Ed.) (1999). Lives and Deaths: Selections from the Works of Edwin S. Shneidman. Ch.22, p.303-310. London UK: Brunner/Mazel.

Shneidman, E.S. & Farberow, N.L. (1957a). "Some comparisons between genuine and simulated suicide notes in terms of Mowrer's concepts of Discomfort and Relief". Journal of General Psychology, 56, p.251-256.

Shneidman, E.S. & Farberow, N.L. (Eds.) (1957b). Clues to Suicide. New York: McGraw-Hill Book Company.

Shneidman, E.S. & Farberow, N.L. (1957c). "Clues to Suicide." In Shneidman, E.S. & Farberow, N.L. (Eds.). p.3-10.

Shneidman, E.S. & Farberow, N.L. (1957d). "The logic of suicide". In Shneidman, E.S. & Farberow, N.L. (Eds.). p.31-40.

Shneidman, E.S. & Farberow, N.L. (1957e). "Genuine and Simulated Suicide Notes". In Shneidman, E.S. & Farberow, N.L. (Eds.). p.197-215.

Sinclair, John M. (2001). "Preface". In Ghadessy, Mohsen; Henry, Alex; & Roseberry, L. Robert. (Eds.) Small Corpus Studies and ELT: Theory and Practice. (Studies in Corpus Linguistics, V.5.) John Benjamins.

Skinner, B.F. (1957). Verbal behaviour. New York: Appleton-Century-Crofts.

Smith, Janet. (2002). "The Numbers". [In "Reports: First Report 'Death Disguised', Vol.1, Ch.14 - 'Conclusions'"]. The Shipman Inquiry. p.1-1.
http://www.the-shipman-inquiry.org.uk/fr_page.asp?p=1&id=187
[Accessed September 27, 2006.]

Smith, Janet. (2003). "Death Registration and Cremation Certification". [In "Reports: Second Report: 'The Police Investigation of March 1998', Ch.1 - 'Concerns Are Raised'"]. The Shipman Inquiry. p.1-3.
http://www.the-shipman-inquiry.org.uk/tr_page.asp?p=1&id=40
[Accessed September 27, 2006.]

Smith, Janet. (2005). "Chronology". [In "Background"]. The Shipman Inquiry. p.1-4.
http://www.the-shipman-inquiry.org.uk/bg_chronology.asp
[Accessed September 1, 2006.]

Solan, Lawrence M. (1998). "Linguistic experts as semantic tour guides". Forensic Linguistics: The International Journal of Speech Language and the Law. Vol.5, No.2, p.87-106.

Stincelli, Rebecca. (2001). "My 'Official' Definition and Classification" and "Chapter 1: What is Suicide by Cop?" In "Home" link. SuicideByCop.com
http://www.suicidebycop.com/page2.html [Accessed 14 October, 2003.] [Mapped as p.2-6.]

Stone, Philip J.; Dunphy, Dexter C.; Smith, Marshall S.; Ogilvie, Daniel M.; and Associates. (1966). The General Inquirer: A Computer Approach to Content Analysis. Cambridge, Massachusetts: MIT Press.

Suicide Act 1961. (9 & 10, Elizabeth 2, Ch.60). London: HMSO.
[Revised Statute from The UK Statute Law Database, Office of Public Sector Information.] http://www.opsi.gov.uk/RevisedStatutes/Acts/ukpga/1961/cukpga _19610060_en_1 [Accessed 28 August, 2010.]

Swales, John. (1996). "Occluded genres in the academy: the case of the submission letter". In Ventola, E. & Mauranen, A. (eds.) Academic writing: Intercultural and textural issues. Amsterdam: John Benjamins.

Swales, John. (2007). "Worlds of Genre – Metaphors of Genre". Paper given at SIGET: The 4th International Symposium on Genre Studies in University of Southern Santa Catarina, Tubarão, Santa Catarina, Brazil on August 15th-18th, 2007. p.147-157.
http://www3.unisul.br/paginas/ensino/pos/linguagem/cd/English/15i.pdf
[Accessed 12 August, 2009.]

Tadros, George and McGrath, Melissa. (2004). "A Comparison of Suicide Notes Written by Younger and Older Adults". Poster presentation at the Annual General Meeting of the Royal College of Psychiatrists, 6-9 July 2004, Harrogate, U.K. [Mapped as p.1-4.]

http://www.rcpsych.ac.uk/pdf/PRELIMINARYPROGRAMME.pdf
[Accessed circa October 2004.]

Tochon, F.V. (1999). "Semiotic foundations for building the New Didactics: An introduction to the prototype features of the discipline". Instructional Science. 27: 9-32. www.springerlink.com/content/g2767g144633x775/fulltext.pdf [Accessed 26 October, 2010.]

Toshiki, Shioiri; Akiyoshi, Nishimura; Kohei, Akazawa; Ryo, Abe; Hideyuki, Nushida; Yasuhiro, Ueno; Maki, Kojika-Maruyama; Toshiyuki, Someya. (2005). 'Abstract' from "Incidence of note-leaving remains constant despite increasing suicide rates". Psychiatry and Clinical Neurosciences. April 2005. Vol.59, No.2, p.226-228. http://www.ingentaconnect.com/content/bsc/pcn/2005/00000059/00000002/ art00021;j... [Accessed May 2, 2005.]

Towl, Graham. (1999). "Self-Inflicted Deaths in Prisons in England and Wales from 1988 to 1996". In Towel, Graham; McHugh, Martin; & Jones, Dilys (Eds.) Suicide in Prisons: Research, Policy and Practice. p.13-20. Brighton: Pavilion Publishing Ltd.

Tuckman J., Kleiner, R.J. & Lavell, M. (1959). "Emotional content of suicide notes". American Journal of Psychiatry. 116, p.59-63.

Van Dijk, T.V. (1977). Text and Context: Explorations in the Semantics and Pragmatics of Discourse. London: Longman.

Van Wormer, Katherine and Odiah, Chuk. (1999). "The Psychology of Suicide-Murder and The Death Penalty". Journal of Criminal Justice. Vol.27, No.4, p.361-370. [Online journal. Accessed May 2005.]

Wellisch, Hans H. (1994). "Some comments on 'Subject analysis and indexing'". The Indexer. Vol.19, No.2. http://www.theindexer.org/files/19-2/19-2_131.pdf [Accessed October 2, 2010.]

Wilson, Andrew & Thomas, Jenny. (1997). "Semantic Annotation". In Garside, R., Leech, G. & McEnery, A. (Eds.) Corpus Annotation: Linguistic Information from Computer Text Corpora. Ch.4, p.53-65. London: Longman.

Winter, Eugene. (1996). "The Statistics of Analysing Very Short Texts in a Criminal Context". In Kniffka, H., Blackwell, S. & Coulthard, M. (Eds.) Recent Developments in Forensic Linguistics. Ch.5, p.141-179. Frankfurt am Main: Peter Lang.

2. Legal Cases
[All cases are from "LexisNexis" Database, http://web.lexis-nexis.com/professional/...]
[U.K. cases are from: UK Cases, Combined Courts. Accessed May 21, 2003.]
[Non-U.K. Cases are from: ACT Supreme Court Unreported. Access dates: see below.]

Attorney-General's Reference (No 24 of 1999). (1999). Court of Appeal (Criminal Division), (Transcript: Smith Bernal), 29 JUNE 1999.

Attorney-General's Reference (No 40 of 2002). (2002). Court of Appeal (Criminal Division), [2002] EWCA Crim 2219, (Transcript: Smith Bernal), 3 October 2002.

Cooke v. Cooke. (1960). Probate, Divorce and Admiralty Division, [1960] 3 All ER 39, [1960] 3 WLR 807, 16 June 1960.

Cross and another v. Highlands and Islands Enterprise and another. (2000). Court of Session: Outer House, 2001 SCLR 547, 5 December 2000.

Devine v. Attorney General for Northern Ireland and another, Breslin v. Attorney General for Northern Ireland and another. (1992). House of Lords, [1992] 1 All ER 609, [1992] 1 WLR 262, 6 February 1992.

R v. Bamber. (2002). Court of Appeal (Criminal Division), [2002] EWCA Crim 2912, [2002] All ER (D) 165 (Dec), (Approved judgment), 12 December 2002.

R v. Campbell. (1984). Court of Appeal (Criminal Division), (Transcript: Marten Walsh Cherer), 20 December 1984.

R v. Dear. (1996). Court of Appeal (Criminal Division), [1996] Crim LR 595, (Transcript: John Larking), 14 March 1996.

R v. Gilfoyle. (2000). Court of Appeal (Criminal Division), Lexis UK CD M428, [2001] 2 Cr App Rep 57; The Times 13 February 2001, (Transcript: Smith Bernal), 20 December 2000.

R v. HM Coroner for Ceredigion ex parte Wigley. (1993). Queen's Bench Division (Crown Office List), CO/290/92R, (Transcript: Marten Walsh Cherer), 8 February 1993.

R v. HM Coroner for Newbury, ex parte John. (1991). Queen's Bench Division, 156 JP 456, 15 March 1991.

R v. HM Coroner for Northamptonshire, ex parte Anne Walker. (1988). Queen's Bench Division, 153 JP 289, 20 December 1988.

R v. HM Coroner for Portsmouth, ex parte John Keane. (1989). Queen's Bench Division, 153 JP 658, 23 June 1989.

R v. HM Coroner for Solihull ex parte Nutt. (1993). Queen's Bench Division (Crown Office List), CO/1318/92, (Transcript: John Larking), 21 April 1993.

R v. James. (1998). Court of Appeal (Criminal Division), (Transcript: Smith Bernal), 31 July 1998.

R v. Kavanagh. (2002). Court of Appeal (Criminal Division), [2002] EWCA Crim 904, (Transcript: Smith Bernal), 26 March 2002.

R v. Michael Anthony Taylor. (1996). Court of Appeal (Criminal Division), CARF 2087, (Transcript), 31 May 1996.

R v. Tonge. (1993). Court of Appeal (Criminal Division), 157 JP 1137, [1993] Crim LR 876, 2 April 1993.

R v. Torney. (1997). Court of Appeal (Criminal Division), MACC2236M (Transcript), 18 April 1997.

R v. Wood. (1989). Court of Appeal (Criminal Division), [1990] Crim LR 264, (Transcript: Marten Walsh Cherer), 24 July 1989.

R (on the application of Middleton) v. Western Somersetshire Coroner. (2001). Queen's Bench Division (Administrative Court), [2001] EWHC Admin 1043, CO/563/2001, (Transcript: Smith Bernal), 14 December 2001.

R (on the application of Pretty) v. Director of Public Prosecutions and another. (2001). Queen's Bench Division, [2001] EWHC Admin 788, 63 BMLR 1, 18 October 2001.

Re J (a minor). (1994). Court of Appeal (Civil Division), (Transcript: John Larking), 11 June 1994.

Re N (Minors). (1993). Court of Appeal (Civil Division), (Transcript: John Larking), 27 October 1993.

Reeves Respondent and Commissioner of Police of the Metropolis Appellant. (2000). House of Lords, [2000] 1 AC 360, 15 July 1999, (c)2001 The Incorporated Council of Law Reporting for England & Wales.

Ryan v. Kazacos. (2001). Estate of Michael Harvey Kazacos, NSWSC 140, 104947/97, Supreme Court of New South Wales Equity Division Probate List, BC200100812, 183 A.L.R. 506. 20 and 21 February 2001 heard, 13 March 2001 delivered, Young J. [Accessed May 27, 2003.]

Stead v. Stead. (1985). Court of Appeal (Civil Division), (Transcript: Association), 17 January 1985.

U.S.A. v. Frank Feichtinger. (1997). United States Court of Appeals for the Seventh Circuit. No. 96-2375. 105 F.3d 1188. LEXIS 1835. [Accessed September 15, 2003.]