

# Leeds Studies in English

## Article:

Paul Kerswill, Carmen Llamas and Clive Upton, 'The First SuRE Moves: Early Steps Towards a Large Dialect Project', *Leeds Studies in English*, n.s. 30 (1999), 257-69

## Permanent URL:

[https://ludos.leeds.ac.uk:443/R/-?func=dbin-jump-full&object\\_id=124878&silo\\_library=GEN01](https://ludos.leeds.ac.uk:443/R/-?func=dbin-jump-full&object_id=124878&silo_library=GEN01)



*Leeds Studies in English*  
School of English  
University of Leeds  
<http://www.leeds.ac.uk/lse>

# **The First SuRE Moves: Early Steps Towards a Large Dialect Project**

Paul Kerswill, Carmen Llamas and Clive Upton

## ***Abstract***

Two factors have led to a sea-change taking place in the field of dialectology during recent times. The first is a considerable broadening of research aims to include models of the diffusion of changes through both geographical space (geographical spread) and social space (permeating different social groups at different times). The second is the very recent use of digital technology.

This paper argues that the time has come for a new survey of British and Irish English to be instituted, taking account of new orientations and methodologies. Firstly we propose continuous recording, to document the simplest facts of language variation over a wide geographical area in a way that will be useful to future linguists. Further, in a more elaborately structured enquiry of more restricted time-span, it should be possible to provide clear and detailed outcomes directly relating to current issues concerning the diffusion of language change.

In order to undertake the proposed large-scale survey of regional English, data must be obtained which are analysable on three levels of variation: phonological, grammatical and lexical. This paper outlines a new method of data elicitation which will prove to be a workable, effective and relatively simple way of obtaining data which allow for a complete picture of regional variation found throughout the British Isles at the turn of the Millennium and onwards.

## **1. *Introduction***

The 20th century has seen unprecedented changes in the regional speech of the British Isles. The industrialisation of Britain in the 18th and 19th centuries led to the

establishment of new, urban varieties of English in all its cities and large towns. It was recognition of the upheavals caused by the rapid pace of change which led Joseph Wright (1905: vii) to assert that by 1925 it would be 'quite impossible to get together sufficient pure dialect material to enable any one to give even a mere outline of the phonology of our dialects as they existed at the close of the nineteenth century'. The varieties created were in some sense a compromise between the dialects of the people who migrated there from the respective hinterlands and from further afield. However, in a time of unprecedented geographical and social mobility, this process of the development of 'compromise' dialects seems to be even stronger at the close of our present century. Not only are there now very few speakers of genuine rural dialects, but the overwhelmingly urban speech heard in the streets today is being 'levelled' in the direction of a set of relatively uniform varieties which have a clearly identifiable regional flavour but which are relatively difficult to pin down to a more specific locality.

It is the argument of the authors of this paper that it is now time to begin a two-pronged assault on a deficiency which exists in our knowledge of the contemporary situation. Firstly, and at bottom, we propose a process of continuous principled recording to document the simplest facts of language variation over a wide geographical area, in a way that will be useful to future generations of linguists. Further, in a more elaborately structured enquiry of more restricted time-span, it should be possible to provide clear and detailed outcomes directly relating to current issues concerning the diffusion of language change.

The previous network surveys of regional speech in Britain are the Survey of English Dialects (Orton and Dieth 1962-71), the Survey of Anglo-Welsh Dialects (Parry 1977, 1979), the Linguistic Survey of Scotland (Mather and Speitel 1975), and the Tape-Recorded Survey of Hiberno-English Speech (Barry 1981). To these can be added recent and ongoing endeavours of restricted theoretical scope or geographical spread, such as the Survey of British Dialect Grammar (Cheshire et al. 1989), Kerswill and Williams' (1997) work on the role of adolescents in dialect levelling, and the Tyneside and Derby study (Docherty et al. 1997). Each of inestimable value in its own way, these widespread or more focussed projects are of different times, have different aims, and employ different methodologies: together they do not amount to a cohesive record of the dialects.

To the fact of the fragmentary nature of our holdings of information can be added the fact that dialectology has quite recently undergone two sea-changes. Firstly, there has been a considerable broadening of its research aims to include models of the diffusion of changes through both geographical space (geographical spread) and social

space (permeating different social groups at different times) (Butters 1997). Dialectologists have used their data to model types of language change in a way entirely compatible with Labovian sociolinguistics (Bailey, Guy & Wikle 1993); indeed, it has recently been argued that dialectology must be regarded as part of sociolinguistics (Butters 1997: 11). There has also been the use of models of geographical diffusion to account for patterns of spread (Trudgill 1974; Upton 1995; Hernandez-Campoy 1996).

Secondly, there have been enormous technological advances in the recording, storage, sorting, and retrieval of data. Quite early use of such technology is exemplified by the Computer Developed Linguistic Atlas of England (Viereck 1991; Ramisch 1997), a project which has entered some of the Basic Material (Orton et al. 1962-71) of the SED onto a database allowing sophisticated displays of individual linguistic features and, importantly, displays showing overall differences between locations measured in terms of a large number of features. It is now finding further expression in the work of Elmer and Rudin (1997) and Schiltz (1997). Such scholarship as theirs shows what can be done with material not designed with computerisation in mind, and hints at the possibilities for enquiries which presume upon the digitised storage of speech signals, the tagging of speech samples for automatic retrieval, and the facility of remote retrieval of stored data.

Large-scale surveys of regional speech in this country have not before had the benefit of full computerization. The obvious UK point of reference, the British National Corpus (1994), did not intend to achieve any degree of representativeness as a sample of regional speech, and a more directed approach is needed.

## **2. *SuRE: The New Project***

The project being embarked upon, the Survey of Regional English (SuRE), will create from the outset a large computer-held corpus whose form will be guided by the need for it to be the object of analytical work addressing current research questions, those to do with levelling, while at the same time being sufficiently broad to allow to be addressed research questions which may arise in the future.

Having regard to issues of current and continuing theoretical linguistic interest, and to the technologies available to us, active steps are under way to put in place a double-banked project. Funding is being sought by a Leeds/Sheffield/Reading axis intensively to survey the speech of a planned network of British and Irish localities. In the meantime, a doctoral research grant has been secured, part of the brief of the holder

being to advance thinking on that method and, in particular, to generate ideas for rapid data-collection which can be implemented both within and beyond a planned network and limited time-frame.

The 'double-banked' nature of our approach is fundamental (and here it seems reasonable to follow Orton and Wright's Introduction to *A Word Geography of England* by venturing into military metaphor). We see the need in the most detailed way possible to reconnoitre in strength the dialects as they exist at the turn of the Millennium: for this reason a 'heavy brigade' approach is required, targetting a finite set of localities with a comprehensive, tightly structured campaign. Such a foray is logistically demanding, however, and cannot be sustained indefinitely. For this reason, a 'light brigade' action is envisaged: this can be sustained indefinitely, with field linguists super-adding a simple technique of elicitation to any methodology which they are employing, in any locality and at any time, so that a bank of data accumulates. This nugget of comparable data, properly tagged, will be the kernel of all collecting, growing to an open-ended resource capable of analyses known and as yet unforeseen.

At the moment the envisaged characteristics of the 'heavy brigade' project are as follows. Firstly, a network of pre-defined localities in England, Wales, Scotland, Northern Ireland and the Republic of Ireland will be targetted. Locality distribution will be determined essentially by density of population along lines set by the Survey of British Dialect Grammar (Cheshire et al. 1989: 190), a project which made use of a classification of locations in Britain according to, among other variables, the degree of urbanisation (CURDS Functional Regions framework, see Champion et al. 1987). In principle, this will allow every resident to have an equal chance of selection, and will reflect the urban bias of much of the British Isles. Geographical models of diffusion can be applied to the data deriving from the network. Subsets of localities will be subject either to fuller or less detailed sociolinguistic investigations.

Secondly, a standard set of items will be elicited, with some variation allowed to enable known regionally-significant variables to be investigated. The methodology for this aspect of the collection, which is currently being designed and trialled, is now outlined.

### ***3. SuRE: towards a new methodology***

In order to undertake a large-scale survey of regional variation in contemporary spoken British English, data must be obtained which can be analysed on three levels

of possible variation; phonological, grammatical and lexical. Although it is difficult to combine the three levels, to discount any would be to obtain an incomplete picture of the regional linguistic variation found in the British Isles at the turn of the Millennium. The phonological, grammatical and lexical data must be comparable across the localities to be studied, permitting quantitative analyses of the different levels of regional and social variation.

As the SuRE project is collaborative in nature, the problem of combining the three levels of analysis into a single data elicitation method which will be satisfactory to all interested parties is considerable. The problem is further compounded by the necessity of any data elicitation technique to be relatively quick and easy to administer. Researchers must be able to apply the methodology to their fieldwork with the minimum of prior preparation or administration superfluous to their particular fieldwork needs. With these considerations and underlying difficulties in mind, a new method of data elicitation and collection is proposed. After refinement, it is anticipated that this method will prove to be an effective and relatively simple way of gaining data which are analysable on a number of different levels, and which will be usable by researchers whatever their particular research interest.

The primary aim of the SuRE interview is to obtain informal speech from the informant from which an analysis can be made on one, or more than one, level. This being the case, the somewhat formal context of the fieldworker asking set questions to elicit lexis or grammar in an extremely lengthy interview, as in the SED (Orton et al. 1962-71), would be entirely inappropriate. Similarly, methods involving the fieldworker asking questions to elicit involving personal narratives (cf. Labov 1972), or allowing the informants to converse in pairs on topics of their choosing (cf. Docherty et al. 1997; Llamas 1998) would also be unsuitable, as the possibility of obtaining any comparable data on lexical variation would be almost completely removed.

It is necessary, then, to find a way of combining informal conversation, from which data for phonological and, to some extent, grammatical analyses are obtained, with information on lexical variation. With this in mind it is intended that the fieldworker will 'lead' a conversation around linguistic domains, ideally with socially paired informants, permitting interaction to be more like a conversation and less like an interview. The fieldworker prompts discussion about lexical items used in the given area, encouraging the informants to discuss their 'dialect' words; how they are used and what connotations they have. Whilst producing relatively casual conversation, this means of eliciting data yields valuable information on knowledge and use of lexical items, as well as revealing possible age or gender variation in lexis

within a given dialect. In the course of the conversation, how much the speakers are actually aware of variation, as well as attitudinal information on lexis and dialect, are also revealed.

#### 4. *Sense Relation Network Sheets*

##### 4.1 *SRNs: Visual Design and Content Design*

However, although it is successful in obtaining informal speech, simply *talking* about lexical variation does not yield comparable or quantifiable lexical data. In order to allow the information on lexical items to be comparable across the network of localities to be studied, and to give a somewhat flexible structure to the interview, Sense Relation Network sheets (SRNs) have been designed. An example of one of the SRNs, three of which form the core of the SuRE interview, is given in Figure 1 (see Appendix). Both the visual design and the content design of the SRNs are inspired by the idea that there exists a 'web of words', or a series of interconnected networks which define, delimit and store linguistic expressions in the mind (Aitchison 1994, 1996).

Visually, a network is designed as shown in the example given in Figure 1. The language domain of *Feelings, Actions & States* is broken down into subdivisions which form the network. Standard notion words are listed, and space is provided for the informant to insert a dialectal partial synonym for the standard notion word. The visual design of the SRN is aimed at making it as visually pleasing and unthreatening as possible, so the informant will actually *want* to complete it.

In terms of content, the SRNs are built around domains of language, much like the grouping of questions by subject matter in the SED questionnaire. Standard notion words are given as pointers, since interviews which use indirect elicitation techniques are much more time-consuming than those which use direct ones. Also, with an indirect question, the interaction may feel more like an interview or a test of some sort, rather than a conversation, and this may have the affect of increasing the formality of the speech of the informant. In the SRNs, the presence of an *indirect* prompt would considerably increase the density of the written input, which may result in a negative reaction from the informants. Through trialling and revision, although none of the original domains have been entirely lost, the number of SRNs has been drastically cut from eight to three, reducing both the time needed by informants to complete them and the time necessary to conduct the interview.

The SRNs then, as well as being a visual network, rather than a list of

questions, represent the interrelated network of paradigmatic and syntagmatic sense relations in which linguistic expressions from similar semantic fields define and delimit each other's meaning. They also represent the sense relation of partial synonymy, which the dialectal variant holds with the standard notion. Additionally, in time they will represent a geographical sense relation network of dialectal variation of partial synonyms found throughout the British Isles.

#### *4.2 SRNs: Technique of Administration, and Data Yielded*

A crucial part of the new methodology used for the SuRE interview involves the actual administration of the SRNs and the conduct of the interview. Informants are given the SRNs some five days in advance of the interview. This allows them time to consider the words they use, and lessens drastically the possibility of the mind going blank if an informant is called upon to give an immediate response to something which is not often consciously thought about. Also, if the informants are aware of what is going to be discussed, the feeling of somehow being tested may be lessened considerably. It is important that the interview is enjoyable and unthreatening, in order to ensure the ready recruiting of informants and to maximise the possibility of gaining access to their least overtly careful or monitored speech style.

The method of basing an informal recorded conversation on the SRNs allows the fieldworker to secure the written record of the informant's responses on the SRNs, which are retained by the fieldworker, and also to secure the backup of the recorded spoken version of the responses for pronunciation purposes. The spoken recording of the responses also acts as a safeguard against mis-spellings, which may indeed prove to be an interesting research exercise in itself. In the course of the interview, other lexical items not given on the SRNs may also be revealed, with informants becoming aware that they use a particular word when they hear someone else use it, or with informants using dialectal variants without necessarily being aware they are doing so. The amount of lexical data obtained through the SRNs is considerable. In terms of obtaining phonological and grammatical data, informal speech is produced through the use of the SRNs. Informants are willing to talk at length about lexis, and about attitudes towards lexical items and awareness of variation which also yields a mass of attitudinal data.

### *5. Identification Questionnaire*

Combined with the SRNs an Identification Questionnaire (IdQ) is included in the



SuRE interview. An example of the questions posed in the IdQ is given in Figure 2 (see Appendix). The complete IdQ comprises 15 questions, whose primary aim is to act as a safety net: the questions posed elicit a relatively extended response, should the informant respond minimally to the SRNs. The questions on the IdQ are designed to combine informal speech with valuable information on people's attitudes towards language and identity (Le Page & Tabouret-Keller 1985), existence and awareness of age and gender differences (Kerswill 1996; Kerswill and Williams 1997; J. Milroy, L. Milroy and S. Hartley 1994; Trudgill 1974), and rudimentary ideas on density of networks (Milroy 1987). They may also elicit short personal narratives, and information on people's perception of language areas and boundaries (Preston 1988). As communities and boundaries are often symbolic, it is difficult to impose a definition of speech community onto a geographical area and a group of people, even when an investigator is a native of the geographical area to be studied. The similarities and differences which define and delimit communities are often not a matter for objective assessment, but are largely subjective, involving feelings and existing in the minds of the members of the community (Cohen 1985: 21). There is no reason why the topical content of the interview should not be of use in this regard, with the fieldworker tapping the natural resource of the informant for information on language, area, boundaries and attitudes found in Britain at the turn of the Millennium. Attitudinal information from the individual informant is yielded through the use of the IdQ. This attitudinal information will give comparable data across regions of Britain and may reveal differing regional attitudes towards areas and dialects.

There also exists the possibility of scoring the responses given in the IdQ in terms of positive, negative and neutral answers, as in Labov's Martha's Vineyard study (1972). Such quantification can be used to discover whether any correlation exists between attitudes to the area and linguistic or non-linguistic variables, again considered both in terms of the individual or social groups of informants within a given region and in terms of differing attitudes towards regional areas as a whole, on a national scale. The questions on the IdQ are also given to the informants prior to the interview, and are included with the SRNs in a small interview pack.

Ultimately included in the interview may be a word or sentence list to be read by the informant to facilitate observation of stylistic variation, and for control of phonological features. Also, a more formal grammatical element may be included along the lines of an adapted questionnaire similar to that used by Cheshire et al. (1989) for the Survey of British Dialect Grammar.

## 6. Conclusion

Two elements of SuRE are envisaged: firstly, a detailed survey designed to provide up-to-date information over a wide area, informing the linguistic community on current issues; secondly, accumulation of an uncomplicated bank of consistently collected data, rolling forward, built by and available to anyone committed to its creation.

At present the methodology envisaged for the data elicitation involves an informal interview with socially paired informants, providing the data for phonological and, to some extent, grammatical analyses. This interview is centred on the three SRNs, which will also yield data on lexical variation, and the IdQ, which will provide attitudinal and ethnographic information. This, then, can be seen as the core element of the SuRE interview, which can be expanded for the purposes of the 'heavy brigade' project, i.e. the detailed survey. A reduced version of this core will act as the 'light brigade', i.e. the bank of consistently collected data, if the whole proves to be too lengthy or complex to administer.

Knowledge of the early periods of each stage of SED shows that a final viable project is different in detail from that which is initially envisaged, and this may prove the case for the SuRE project. But what we are decided upon is that we should make every effort to follow the example of our predecessors in principled data gathering. It would be a sad reflection on us if we were now to continue to rely on the ever more dated findings of those who preceded us, without setting in place the coherent collection of data which can in some measure at least be compared with earlier findings, and which will just as importantly provide data for the instruction of those who will follow.

## REFERENCES

- Aitchison, Jean. 1994. *Words in the Mind: An Introduction to the Mental Lexicon*. 2nd edn. Oxford: Blackwell.
- Aitchison, Jean. 1996. *Reith Lectures*. B.B.C. Radio 4.
- Bailey, Guy, Tom Wike, Jan Tillery & Lori Sand. 1993. Some patterns of linguistic diffusion. *Language Variation and Change* 5: 359-90.
- Barry, Michael V. 1984. The English Language in Ireland. In R.W. Bailey and M. Gorlach (eds) *English as a World Language* 84-133. Cambridge: Cambridge University Press.

- British National Corpus. 1994. Longman, Chamber-Harrap, OUP, British Library, Unit for Computer Research, Lancaster University.
- Butters, Ronald. 1997. Dialectology and sociolinguistic theory. In Alan Thomas ed. *Issues and Methods in Dialectology* 1-13. Bangor: Department of Linguistics, University of Wales, Bangor.
- Champion, A.G., A.E. Green, D.W. Owen, D.J. Ellin & M.G. Coombes. 1987. *Changing Places: Britain's Demographic, Economic and Social Complexion*. London: Edward Arnold.
- Cheshire, Jenny, Viv Edwards & Pam Whittle. 1989. Urban British Dialect Grammar: the question of dialect-levelling. *English World Wide* 10: 185-225.
- Cohen, Antony, P. 1985. *The Symbolic Construction of Community*. Chichester: Ellis Horwood.
- Docherty, Gerard J., Paul Foulkes, James Milroy, Lesley Milroy & David Walshaw. 1997. Descriptive adequacy in phonology: variationist perspective. *Journal of Linguistics* 33: 275-310.
- Elmer, Willy & Ernst Rudin. 1997. The Survey of English Dialects as an electronic database for research in areal and variationist linguistics. In Alan Thomas (ed) *Issues and Methods in Dialectology* 234-46. Bangor: Department of Linguistics, University of Wales, Bangor.
- Hernandez-Campoy, Juan Manuel. 1996. *Modelos de difusión geográfica de las innovaciones sociolingüísticas en los acentos del Reino Unido*. Unpublished Ph.D. thesis. Universidad de Murcia.
- Kerswill, Paul. 1996. Children, adolescents and language change. *Language Variation and Change* 8: 177-202.
- Kerswill, Paul & Ann Williams. 1997. Creating a new town koine: children and language change in Milton Keynes. In *Reading Working Papers in Linguistics* 3: 205-58. Department of Linguistic Science, University of Reading.
- Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Le Page, Robert B., & Andrée Tabouret-Keller. 1985. *Acts of Identity: Creole-based Approaches to Language and Ethnicity*. Cambridge: Cambridge University Press.
- Llamas, Carmen. 1998. Language variation and innovation in Middlesbrough: a pilot study. In *Working Papers in Linguistics* vol. 6. Department of Linguistics and Phonetics, University of Leeds.
- Mather, J.Y., H.H. Speitel & G.W. Leslie. 1975. *The Linguistic Atlas of Scotland*. 3 vols. London: Croom Helm.
- Milroy, James, Lesley Milroy & Sue Hartley. 1994. Local and supra-local change in British

## *The First SuRE Moves: Early Steps Towards a Large Dialect Project*

- English: The case of glottalisation. *English World Wide* 15:1 1-33.
- Milroy, Lesley. 1987. *Language and Social Networks*. Oxford: Blackwell.
- Orton Harold et al. 1962-71. *Survey of English Dialects*. Introduction and 4 vols. in 3 parts. Leeds: E.J. Arnold.
- Orton, Harold & Nathalia Wright. 1974. *A Word Geography of England*. London: Seminar Press.
- Parry, David (ed). 1977, 1979. *The Survey of Anglo-Welsh Dialects*. 2 vols. Swansea: University College of Swansea.
- Preston, Dennis R. 1988. Methods in the Study of Dialect Perceptions. In Alan R. Thomas (ed) *Methods in Dialectology* 373-95. Clevedon: Multilingual Matters.
- Ramisch, Heinrich. 1997. Dialectological and cartological features of the *Computer Developed Linguistic Atlas of England*. In Alan Thomas (ed) *Issues and Methods in Dialectology* 224-33. Bangor: Department of Linguistics, University of Wales, Bangor.
- Schiltz, Guillaume. 1997. A dialectometric analysis of the dialects in the Grand-Duchy of Luxembourg. in Alan Thomas (ed) *Issues and Methods in Dialectology* 94-108. Bangor: Department of Linguistics, University of Wales, Bangor.
- Trudgill, Peter. 1974. *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.
- Upton, Clive. 1995. Mixing and fudging in Midland and southern dialects of England: the *cup* and *foot* vowels. In Jack Windsor Lewis (ed) *Studies in General English Phonetics: Essays in Honour of Professor J.D. O'Connor*. London: Routledge.
- Viereck, Wolfgang. 1991. *Computer Developed Linguistic Atlas of England 1*. Tübingen: Niemeyer.
- Wright, Joseph. 1905. *The English Dialect Grammar*. Oxford: Clarendon Press.

### **Addresses**

Paul Kerswill  
Department of Linguistic Science  
University of Reading  
Whiteknights  
PO Box 218  
Reading RG6 6AA  
England  
<p.e.kerswill@reading.ac.uk>

Carmen Llamas and Clive Upton  
School of English  
University of Leeds  
Leeds LS2 9JT  
England  
<engml@leeds.ac.uk>  
<c.s.upton@leeds.ac.uk>

# APPENDIX

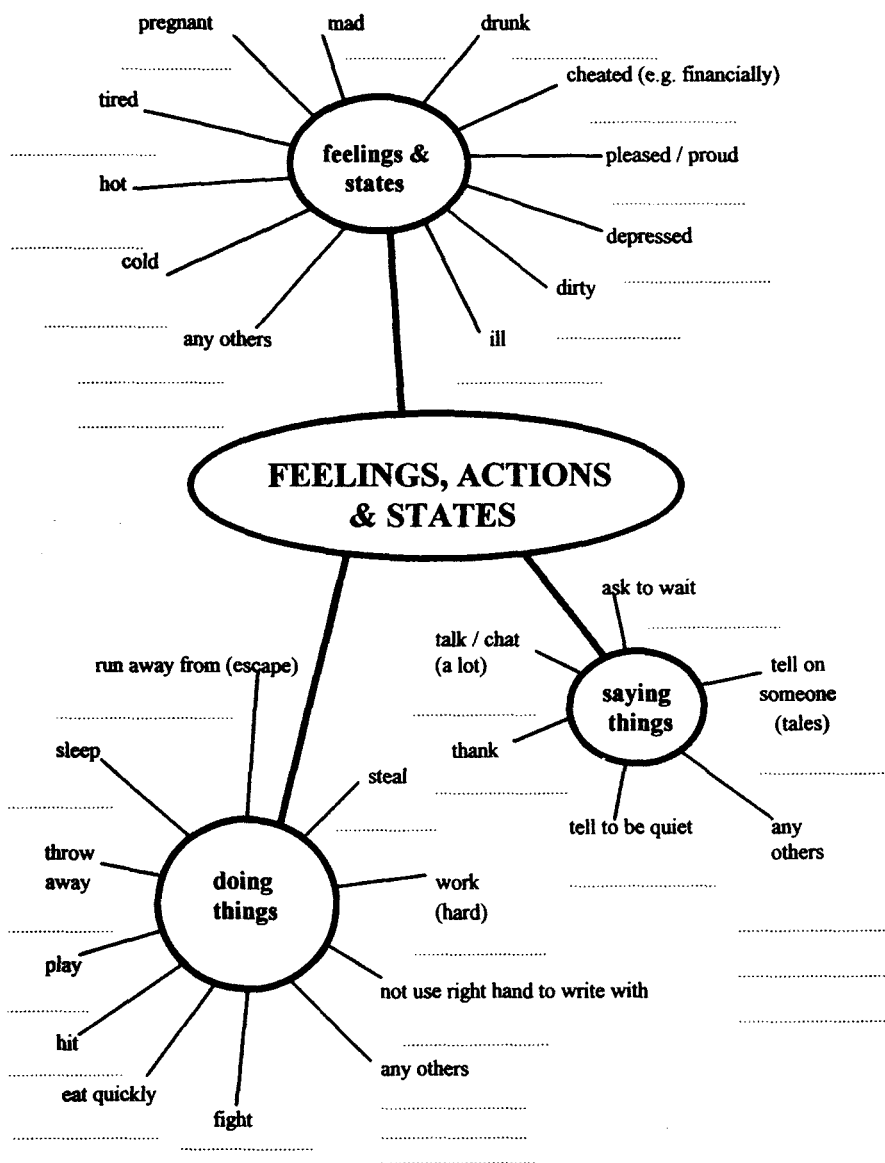


Figure 1: Sense Relation Network sheet

### **Your Language**

- What accent would you say you had, and do you like it?
- Have you ever been in a situation where you've deliberately changed the way you talk? If so, why?
- Do you think there's a difference between how males and females speak here?
- Where, geographically, would you say people stop talking the same as you and start sounding different?

### **Your Area**

- If you were watching a regional news programme, what places would you expect to hear news from?
- What image or description of your home town would you give to someone who didn't know it?
- If you could, would you change where you came from? Why/why not?
- What do you think the best and worst things are about growing up and living in your home town?

Figure 2: Identification Questionnaire (example questions)