

# Knowledge Acquisition for Natural Language Generation

Ehud Reiter and Roma Robertson

Dept of Computing Science  
Univ of Aberdeen, Scotland  
{ereiter,rroberts}@csd.abdn.ac.uk

Liesl Osman

Dept of Medicine and Therapeutics  
Univ of Aberdeen, Scotland  
l.osman@abdn.ac.uk

## Abstract

We describe the knowledge acquisition (KA) techniques used to build the STOP system, especially sorting and think-aloud protocols. That is, we describe the ways in which we interacted with domain experts to determine appropriate user categories, schemas, detailed content rules, and so forth for STOP. Informal evaluations of these techniques suggest that they had some benefit, but perhaps were most successful as a source of insight and hypotheses, and should ideally have been supplemented by other techniques when deciding on the specific rules and knowledge incorporated into STOP.

## 1 Introduction

An important aspect of building natural-language generation (NLG) systems is *knowledge acquisition*. This is the process of acquiring the specific knowledge needed in a particular application about the domain, the language used in the domain genre, the readers of the texts, and so forth. Such knowledge influences, for example, the system's content selection rules (whether represented as schemas, production rules, or plan operators); the system's microplanning choice rules (lexicalisation, referring expression generation, aggregation); and perhaps even the system's grammar (if a genre grammar is needed, as is the case, for example, with weather reports).

To date, knowledge acquisition for NLG systems has largely been based on corpus analysis, informal interactions with experts, and informal feedback from users (Reiter et al., 1997; Reiter and Dale, 2000). For example, the PlanDoc developers (McKeown et al., 1994) interviewed users to get a general understanding of the domain and user requirements; asked a single expert to write some example output texts; and then analysed this corpus in various ways. Other KA techniques used in the past for building NLG systems include letting domain experts specify content rules in pseudo-code (Goldberg et al., 1994) and ethnographic techniques such as observing doctors and patients in real consultations (Forsythe, 1995).

As part of the STOP project (Reiter et al., 1999) to

generate personalised smoking-cessation leaflets, we investigated using some of the structured KA techniques developed by the expert-system community (see, for example, (Scott et al., 1991)) for acquiring the knowledge needed by an NLG system. In this paper we summarise our experiences. Very briefly, our overall conclusion is that in STOP, structured KA was probably useful for getting insight into and formulating hypotheses about the knowledge needed by an NLG system. However, formulating detailed rules purely on the basis of such KA was not ideal, and it would have been preferable to use other information as well during this process, such as statistics about smokers and feedback from smoker evaluations of draft STOP leaflets.

## 2 Background: The STOP System

The STOP system generates personalised smoking-cessation leaflets, based on the recipient's responses to a questionnaire about smoking beliefs, concerns, and experiences. STOP leaflets consist of four A5 pages, of which only the two inside pages are fully generated; an example of the inside pages of a STOP leaflet are shown in Figure 2. Internally, STOP is a fairly conventional shallow NLG system, with its main innovation being the processing used to control the length of leaflets (Reiter, 2000). STOP has been evaluated in a clinical trial, which compared cessation rates among smokers who received STOP leaflets; smokers who received a non-personalised leaflet with similar structure and appearance to a STOP leaflet; and smokers who did not receive any leaflet (but did fill out a questionnaire). Unfortunately, we cannot discuss the results of the clinical trial in this paper<sup>1</sup>.

One of the research goals of the STOP project was to explore the use of expert-system knowledge acquisition techniques in building an NLG system. These knowledge acquisition sessions were primarily carried out with the following experts:

---

<sup>1</sup>Our medical colleagues intend to publish a paper about the clinical trial in a medical journal, and have requested that we not publish anything about the results of the trial in a computing journal or conference until they have published in a medical journal.

- three doctors (two general practitioners, one consultant in Thoracic Medicine)
- one psychologist specialising in health information leaflets
- one nurse

None of these experts were paid for their time. We also did a small amount of KA with a (paid) graphic designer on layout and typography issues.

## 2.1 Unusual Aspects of STOP from a KA Perspective

KA research in the expert-system community has largely focused on applications such as medical diagnosis, where (1) there is a single correct solution, and (2) the task being automated is one currently done by a human expert. STOP is a different type of application in that (1) there are many possible leaflets which can be generated (and the system cannot tell which is best), and (2) no human currently writes personalised smoking-cessation leaflets (because manually writing such leaflets is too expensive). Point (2) in particular was repeatedly emphasised by the experts we worked with. The doctors and the nurse were experts on oral consultations with smokers, and the health psychologist was an expert on writing non-personalised health information leaflets, but none of them had experience writing personalised smoking-cessation leaflets.

Many NLG systems have similar characteristics. The flexibility of language means that there are almost always many ways of communicating information and fulfilling communicative goals in a generated text; in other words, there are many possible texts that can be generated. Furthermore, while some synthesis tasks, such as configuration and scheduling, can be formalised as finding an optimal solution under a well-defined numerical evaluation function, this is difficult in NLG because of our poor understanding of how to computationally evaluate texts for effectiveness.

With regard to human expertise, some NLG systems do indeed generate documents, such as weather reports and customer-service letters, which are currently written by humans. But many systems are similar to STOP in that they generate texts — such as descriptions of software models (Lavoie et al., 1997) and customised descriptions of museum items (Oberlander et al., 1998) — which are useful in principle but are not currently written by humans, perhaps because of cost or response-time issues.

## 3 KA Techniques Used in STOP

### 3.1 Sorting

Sorting is a standard KA technique for building up taxonomies. Experts in a sorting exercise are given a set of entities, and asked to divide the set into

subsets, and ‘think aloud’ into a tape recorder as they do so.

In STOP, we used sorting to build a classification of smokers. We started off with an initial classification which was motivated by the Stages of Change psychological theory (Prochaska and diClemente, 1992): this divided smokers into the three categories of Precontemplators (not intending to quit anytime soon), Contemplators (seriously considering quitting), and Preparers (definitely decided to quit, and soon). We wished to refine these categories, especially Precontemplator (which includes 67% of smokers in the Aberdeen area), and used sorting to do so. The basic exercise consisted of giving a doctor three sets of questionnaires (a set from Precontemplators; a set from Contemplators; and a set from Preparers), and asking him or her to subdivide each set into subsets. We repeated this exercise with three different doctors.

The results of this exercise were complex, and the doctors were not in full agreement. After some analysis, we proposed to them that we subdivide all three categories on the basis of desire to quit. Precontemplators in particular would be divided up into people who neither want nor intend to quit (Committed Smokers); people who have mixed feelings about smoking but don’t yet intend to quit (Classic Precontemplators); and people who would like to quit but aren’t intending to quit, typically because they don’t think they’ll succeed (Lacks Confidence). The three doctors agreed that this was a reasonable subcategorisation, and we proceeded on this basis. In particular, we operationalised this categorisation as follows:

- We added the question *Would you like to stop if it was easy* to the questionnaire. People who answered *No* were put into the ‘Committed Smoker’ category. For people who answered *Not Sure* or *Yes*, we looked at their decisional balance, that is the number of likes and dislikes they had about smoking, and placed them into Lacks Confidence if their dislikes clearly outnumbered their likes, and Classic Precontemplator otherwise.
- We defined different high-level schemas for each of these categories; these schemas essentially specified which sections and (in some cases) paragraphs should be included in the leaflet, but not the detailed content of individual paragraphs. Under these schemas, Committed Smokers got short non-argumentative letters which gently reminded smokers of some of the drawbacks of smoking, and suggested some sources of information if the smoker ever changed his/her mind; Classic Precontemplators got letters which focused on the drawbacks of smoking; and Lacks Confidence smok-

ers got letters which focused on confidence-building and dealing with barriers to quitting (such as addiction or fear of weight gain). The example leaflet shown in Figure 2, incidentally, is for a Lacks Confidence smoker.

### 3.1.1 Evaluation

After the clinical trial was underway, we attempted to partially evaluate the sorting-derived categories by doing a statistical analysis of the differences between smokers in the groups. In other words, we hypothesised that if our categories were correct in distinguishing different types of smokers, then we should observe differences in characteristics such as addiction and confidence between the groups. Of course, this is not an ideal evaluation because it does not test the hypothesis that the different classes of smokers we proposed should receive different types of leaflets; but this is a difficult hypothesis to test directly.

In any case, our analysis suggested that the smokers in each group did indeed have different characteristics. However, it also suggested that we might have done as well (in terms of creating subgroups with different characteristics) by subcategorising purely on the *Would you like to stop if it was easy* question, and ignoring likes and dislikes about smoking. The analysis also suggested that it might have been useful to subcategorise on the basis of addiction, which we did not do. In fact during the sorting exercises the doctors did mention dividing into groups partially on the basis of the difficulty that individuals would have in quitting, but we did not implement this.

The statistical analysis also suggested some ways of possibly improving the content schemas. For example, the analysis showed that the Committed Smoker category included many light smokers who probably smoked for social reasons; it might have been useful to specifically address this in the STOP leaflets ('quit now, before you become addicted').

In retrospect, then, the sorting exercise was useful in proposing ideas about how to divide Stages of Change categories, and in new questions to ask smokers. However, the process of defining detailed category classification rules and content schemas would have benefited greatly from statistical data about smokers in our target region. In STOP we did not have such data until after the clinical trial had started (and smokers had returned their questionnaires), by which time the system could not be changed. So it would have been difficult to base smoker classification on statistical smoker data in STOP; but certainly we would recommend such an approach in projects where good data is available at the outset.

## 3.2 Think-aloud Protocols

The detailed content and phrasing of STOP letters was largely based on think-aloud example sessions with experts. In these sessions, health professionals would be given a questionnaire and asked to write a letter or leaflet for this person. They were also asked to 'think aloud' into a tape recorder while they did this, explaining their reasoning. Again this is a standard expert-system technique for KA.

### 3.2.1 Example

A simple example of the think-aloud process is as follows. One of the doctors wrote a letter for a smoker who had tried to quit before, and managed to stop for several weeks before starting again. The doctor made the following comments in the think-aloud transcript:

Has he tried to stop smoking before? Yes, and the longest he has managed to stop — he has ticked the one week right up to three months and that's encouraging in that he has managed to stop at least once before, because it is always said that the people who have had one or two goes are more likely to succeed in the future.

He also included the following paragraph in the letter that he wrote for this smoker:

I see that you managed to stop smoking on one or two occasions before but have gone back to smoking, but you will be glad to know that this is very common and most people who finally stop smoking have had one or two attempts in the past before they finally succeed. What it does show is that you are capable of stopping even for a short period, and that means you are much more likely to be able to stop permanently than somebody who has never ever stopped smoking at all.

After analysing this session, we proposed two rules:

- IF (previous attempt to quit) THEN (message: more likely to succeed)
- IF (previous attempt to quit) THEN (message: most people who quit have a few unsuccessful attempts first)

The final system incorporated a rule (based on several KA sessions, not just the above one) that stated that if the smoker had tried to quit before, then the confidence-building section of the leaflet (which is only included for some smoker categories, see Section 3.1) should include a short message about previous attempts to quit. This message should mention length of previous cessation if this

was greater than one week; otherwise, it should mention recency of previous attempt if this was within the past 6 months. The actual text generated from this rule in the example leaflet of Figure 2 is

Although you don't feel confident that you would be able to stop if you were to try, you have several things in your favour.

- You have stopped before for more than a month.

Note that the message (text) produced by the actual STOP code is considerably simpler than the text originally written by the expert. This is fairly common, as is simplifications in the logic used to decide whether to include a message in a leaflet or not. In some cases this is due to the expert having much more knowledge and expertise than the computer system (Reiter and Dale, 2000, pp 30–36). Consider, for example, the following extract from the same think-aloud session

The other thing I notice is that he lives in [Address] which I would suspect is quite a few floors up and that he is probably getting quite puffy on the stairs ...and if he gets more breathless he'll end up being a prisoner in his own house because he'll be able to get down, but he won't be able to get up again

This type of reasoning perhaps requires too much general 'world knowledge' about addresses, stairs, and breathlessness to be implementable in a computer system.

### 3.2.2 Evaluation

Afterwards, we attempted to partially evaluate the rules derived from think-aloud sessions by showing STOP leaflets to smokers and other smoking professionals, and asking for comments. The results were mixed. In terms of content, some smokers found the content of the leaflets to be useful and appropriate for them, but others said they would have liked to see different types of information. For example, STOP leaflets did not go into the medical details of smoking (as none of the think-aloud expert-written letters contained such information), and while this seemed like the right choice for many smokers, a few smokers did say that they would have liked to see more medical information about smoking. Reactions to style were also mixed. For example, based on KA sessions we adopted a positive tone and did not try to scare smokers; and again this seemed right for most smokers, but some smokers said that a more 'brutal' approach would be more effective for them.

An issue which our experts (and other project members) disagreed on was whether leaflets should always use short and simple sentences, or whether

sentence length and complexity should be varied depending on the characteristics of the smoker. In the STOP implementation we decided to always use moderately simple sentences, and not vary sentence complexity for different users. After the clinical trial started, we performed a small experiment to test this hypothesis. In this experiment, we took a computer-generated leaflet and asked one expert (who believed that short sentences with simple words should always be used) to revise the computer-generated leaflet to make it as easy to read as possible, and another expert (who believed that more complex sentences were sometimes appropriate, and such sentences could in some cases make letters seem friendlier and more understanding) to revise the computer-generated leaflet to make it friendlier and more understanding. The revisions made by the experts were primarily microplanning ones (using NLG terminology) — that is, aggregation, ellipsis, lexical choice, and syntactic choice. We then showed the two expert-revised leaflets to 20 smokers and asked them which they preferred. The smokers essentially split 50-50 on this question (8 preferred the easy-to-read leaflet, 9 preferred the friendly-understanding leaflet, 3 thought both were the same). This suggests that in principle it indeed may be useful to vary microplanning choices for different leaflet recipients. We hope to further investigate this issue in future research.

Overall, a general finding of the evaluation was that there were many kinds of variations (including whether to include detailed medical information, whether to adopt a 'positive' or 'brutal' tone, and how complex sentences should be) which were not performed by STOP but might have increased leaflet effectiveness if they had been performed. These types of variations were either not observed at all in the think-aloud sessions, or were observed in sessions with some experts but not others.

In terms of KA methodology, perhaps the key lesson is similar to the one from the sorting sessions; the think-aloud KA sessions were very useful in suggesting ideas and hypotheses about STOP content and phrasing rules, but we should have used other information sources, such as smoker evaluations and small comparison experiments, to help refine and test these rules.

### 3.3 Other techniques

Some of the other KA techniques we tried are briefly described below. These had less influence on the system than the sorting and think-aloud exercises described above.

#### 3.3.1 Expert Revision

We gave experts leaflets produced by the STOP system and asked them to critique and revise them. This was especially useful in suggesting local

changes, such as what phrases or sentences should be used to communicate a particular message. For example, an early version of the STOP system used the phrase *there are lots of good reasons for stopping*. One of the experts commented during a revision session that the phrasing should be changed to emphasise that the reasons listed (in this particular section of the STOP leaflet) were ones the smoker himself had selected in the questionnaire he filled out. This eventually led to the revised wording *It is encouraging that you have many good reasons for stopping*, which is in the first paragraph of the example leaflet in Figure 2.

Revision was less useful in suggesting larger changes to the system, and after the clinical trial was underway, one of our experts commented that he might have been able to suggest larger changes if we had explained the system's reasoning to him, instead of just giving him a leaflet to revise. In other words, just as we asked experts to 'think-aloud' as they wrote leaflets, in order to understand their reasoning, it would be useful if we could give the experts something like the computer-system 'thinking aloud' as it produced a leaflet, so they could understand its reasoning.

### 3.3.2 Group activities

Because experts often disagreed, we tried a variety of activities where a group of experts either discussed or collaboratively authored a leaflet, in the hopes that this would help resolve or at least clarify conflicting opinions. This seemed to work best when we asked two experts to collaborate, and was less satisfactory with larger groups. Several experts commented that the larger (that is, more than 2-3 people) group sessions would have benefited from more structure and perhaps a professional facilitator.

### 3.3.3 Smoker feedback

As mentioned in Section 3.2, we showed several smokers the leaflet STOP produced for them, and asked them to comment on the leaflet. In addition to its role as an evaluation exercise for other KA techniques, we hoped that these sessions would in themselves give us ideas for leaflet content and phrasing rules. This was again less successful than we had hoped. Part of the problem was the smokers knew very little about STOP (unlike our experts, who were all familiar with the project), and often made comments which were not useful for improving the system, such as *I did stop for 10 days til my daughter threw a wobbly and then I wanted a cigarette and bought some and after smoking for over 30 years I've tried acupuncture and hypnosis all to no avail*.

We were also concerned that most of our comments came from well-educated and articulate smokers (for example, university students). It was harder to get feedback from less well-educated smokers (for

---

#### *Paragraph from Nov 97 KA exercise:*

Finally, if you do make an attempt to stop, you could consider using nicotine patches. For people like yourself who smoke 10-20 cigarettes per day, patches double your chances of success if you are determined to stop. You can get more information on patches from your local pharmacist or GP.

---

#### *Paragraph from Feb 99 KA exercise:*

You smoke 11-20 cigarettes a day, and smoke your first cigarette within 30 minutes of waking. These facts suggest you are moderately addicted to nicotine, and so you might get withdrawal symptoms for a time on stopping. It would be worth considering using nicotine patches when you stop; these double the chances of success for moderately heavy smokers such as yourself who make a determined attempt to stop smoking. Your pharmacist or GP can give you more information about this.

Figure 1: Paragraphs written by the same doctor for the same smoker in different KA exercises

---

example, single mothers living in public housing estates). This led to the worry that the feedback we were getting was not representative of the population of smokers as a whole.

### 3.4 KA and the Smoker Questionnaire

KA sessions also effected the smoker questionnaire (STOP's input) as well as the text-generation component of the system. We started with an initial questionnaire which was largely based on a literature review of previous projects, and then modified it based on the information that experts used in KA sessions. For example, the original questionnaire asked people who had tried to quit before what quitting techniques they had used in their previous attempts. However, in KA sessions the experts seemed primarily interested in previous experiences with one particular technique, nicotine replacement (nicotine patches or gum); so we replaced the general question about previous quitting techniques with two questions which focused on experiences with nicotine replacement.

## 4 Stability of Knowledge

In order to determine how stable the results of KA sessions were, we asked one of our doctors to repeat in February 1999 a think-aloud exercise which he had originally done in November 1997. This exercise required examining and writing letters for two smokers. The letters and accompanying think-aloud from the 1999 exercise were somewhat different from

the letters and think-aloud from the 1997 exercise; in very general terms, the 1999 letters had similar core content, but expressed the information differently, in a perhaps (this is very difficult to objectively measure) more ‘empathetic’ manner. An extract from this experiment is shown in Figure 1.

We asked a group of seven smokers to compare one of the 1999 letters with the corresponding letter from the 1997 exercise. Five preferred the 1999 letter; one preferred the 1997 letter; one thought both were similar. Written comments from the smokers suggested that they found the 1999 letter friendlier and more understanding than the 1997 letter.

In summary, it appears that our experts may themselves have been learning how to write effective smoking-cessation leaflets during the course of the STOP project. In retrospect this is perhaps not surprising given that none of them had written such leaflets before the project started.

## 5 Evaluation of KA

An issue that arose several times during the project was whether we could formally evaluate the effectiveness of KA techniques, in an analogous way to the manner in which we formally evaluated the effectiveness of STOP in a clinical trial which compared smoking-cessation rates in STOP and non-STOP groups. Unfortunately, it was not clear to us how to do this; how can one evaluate a development methodology such as KA? In principle, perhaps it might be possible to ask two groups to develop the same system, one using KA and one not, and then compare the effectiveness of the resultant systems (perhaps using a STOP-like clinical trial), and also engineering issues such as development cost and time-to-completion. This would be an expensive endeavour, however, as it would be necessary to pay for two development efforts. Also, the size of a clinical trial depends on the size of the effect it is trying to validate, and a clinical trial which compared (for example) the effectiveness of two kinds of computer-generated smoking-cessation leaflet might need to be substantially larger (and hence more expensive) than a clinical trial that tested the effectiveness of a computer-generated leaflet against a no-leaflet control group.

An even more fundamental problem is that in order for such an experiment to produce meaningful results, it would be necessary to control for differences in skill, expertise, enthusiasm, and luck between the development teams. It might be necessary to repeat this exercise several times, perhaps randomly choosing which development team will use KA and which will not. Of course, repeating the experiment  $N$  times will increase the total cost by a factor of  $N$ .

As we did not have the resources to do the above, we elected instead to focus on the smaller ‘informal’

evaluations described above. We also conducted a small experiment where we asked a group of five smoking-cessation counsellors to compare leaflets produced by an early prototype of STOP with leaflets produced by the system used in the clinical trial. 60% of the counsellors thought the clinical trial system’s leaflets were more likely to be effective, with the other 40% thinking the two systems produced letters of equal effectiveness. This suggests (although does not prove) that the development effort behind the clinical trial system improved leaflet effectiveness. However, we cannot determine how much of the improvement was due to KA and how much was due to other development activities.

## 6 Current Work

We are currently in the process of analysing the results of the clinical trial (which we cannot discuss in this paper), to see if it sheds any light on the effectiveness of KA. This is not straightforward because the clinical trial was not designed to give feedback about KA, but there nevertheless seem to be some useful lessons here, which we hope to report in subsequent publications.

We also are applying the KA techniques used in STOP to a project in a different domain, to see how domain-dependent our findings are. A first attempt to do this, in a domain which involved giving advice to university students, failed because the relevant expert, who initially seemed very enthusiastic, did not give us enough time for KA. This highlights the practical observation that KA requires a substantial amount of time from the expert(s), who must either be paid or otherwise motivated to participate in the sessions. In this case we could not pay the expert, but instead tried to motivate him by pointing out that a successful system would be useful to him in his job; this was not in the end sufficient motivation to get the expert to make time for KA in his (busy) schedule.

After the above failure we switched to another domain, giving feedback to adults who are taking basic-literacy courses. In this domain, we are working with a company, Cambridge Training and Development, which is paying experts for their time when appropriate. This work is currently in progress. One interesting KA idea which has already emerged from this work is observing tutors working with students (we did not in STOP observe doctors discussing smoking with their patients); this is similar to the ethnographic techniques suggested by Forsythe (1995).

## 7 Conclusion

The expert system community believes that it is worth interacting with experts using structured KA techniques, instead of just informally chatting to them or non-interactively studying what they do (as

happens in a traditional corpus analysis). We believe structured KA techniques can also be useful in developing NLG systems, but they are not a panacea and need to be used with some caution.

In retrospect, KA was probably most effective in STOP when used as a source of hypotheses about smoker categories, detailed content rules, the phrasing of messages, and so forth. But ideally these hypotheses should have been tested and refined using statistical data about smokers and small-scale evaluation exercises.

Of course, a key problem in STOP was that we were trying to produce texts (personalised smoking-cessation leaflets) which were not currently produced by humans; and hence there were perhaps no real human experts on producing STOP texts. It would be interesting to see if structured KA techniques were more effective for developing systems which produced texts that humans do currently write, such as weather forecasts and instructional texts.

## Acknowledgements

Many thanks to James Friend, Scott Lennox, Martin Pucchi, Margaret Taylor, and all of the other experts who worked with us. Thanks also to Yaji Sripada and the anonymous reviewers for their very helpful comments. This research was supported by the Scottish Office Department of Health under grant K/OPR/2/2/D318, and the Engineering and Physical Sciences Research Council under grant GR/L48812.

## References

- Diana Forsythe. 1995. Using ethnography in the design of an explanation system. *Expert Systems with Applications*, 8(4):403–417.
- Eli Goldberg, Norbert Driedger, and Richard Kitredge. 1994. Using natural-language processing to produce weather forecasts. *IEEE Expert*, 9(2):45–53.
- Benoit Lavoie, Owen Rambow, and Ehud Reiter. 1997. Customizable descriptions of object-oriented models. In *Proceedings of the Fifth Conference on Applied Natural-Language Processing (ANLP-1997)*, pages 253–256.
- Kathleen McKeown, Karen Kukich, and James Shaw. 1994. Practical issues in automatic document generation. In *Proceedings of the Fourth Conference on Applied Natural-Language Processing (ANLP-1994)*, pages 7–14.
- Jon Oberlander, Mick O'Donnell, Alistair Knott, and Chris Mellish. 1998. Conversation in the museum: experiments in dynamic hypermedia with the intelligent labelling explorer. *New Review of Hypermedia and Multimedia*, 4:11–32.
- James Prochaska and Carlo diClemente. 1992.

*Stages of Change in the Modification of Problem Behaviors*. Sage.

Ehud Reiter and Robert Dale. 2000. *Building Natural Language Generation Systems*. Cambridge University Press.

Ehud Reiter, Alison Cawsey, Liesl Osman, and Yvonne Roff. 1997. Knowledge acquisition for content selection. In *Proceedings of the Sixth European Workshop on Natural Language Generation*, pages 117–126, Duisberg, Germany.

Ehud Reiter, Roma Robertson, and Liesl Osman. 1999. Types of knowledge required to personalise smoking cessation letters. In Werner Horn et al., editors, *Artificial Intelligence and Medicine: Proceedings of AIMDM-1999*, pages 389–399. Springer-Verlag.

Ehud Reiter. 2000. Pipelines and size constraints. *Computational Linguistics*. Forthcoming.

A. Carlisle Scott, Jan Clayton, and Elizabeth Gibson. 1991. *A Practical Guide to Knowledge Acquisition*. Addison-Wesley.

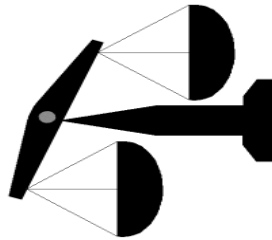
# Smoking Information for Heather Stewart

## You have good reasons to stop...

People stop smoking when they really want to stop. It is encouraging that you have many good reasons for stopping. The scales show the good and bad things about smoking for you. They are tipped in your favour.

### THINGS YOU LIKE

it's relaxing  
it stops stress  
you enjoy it  
it relieves boredom  
it stops weight gain  
it stops you craving



### THINGS YOU DISLIKE

it makes you less fit  
it's a bad example for kids  
you're addicted  
it's unpleasant for others  
other people disapprove  
it's a smelly habit  
it's bad for you  
it's expensive  
it's bad for others' health

## You could do it...

Most people who really want to stop eventually succeed. In fact, 10 million people in Britain have stopped smoking - and stayed stopped - in the last 15 years. Many of them found it much easier than they expected.

Although you don't feel confident that you would be able to stop if you were to try, you have several things in your favour.

- You have stopped before for more than a month.
- You have good reasons for stopping smoking.
- You expect support from your family, your friends, and your workmates.

We know that all of these make it more likely that you will be able to stop. Most people who stop smoking for good have more than one attempt.

## Overcoming your barriers to stopping...

You said in your questionnaire that you might find it difficult to stop because smoking helps you cope with stress. Many people think that cigarettes help them cope with stress. However, taking a cigarette only makes you feel better for a short while. Most ex-smokers feel calmer and more in control than they did when they were smoking. There are some ideas about coping with stress on the back page of this leaflet.

You also said that you might find it difficult to stop because you would *put on weight*. A few people do put on some weight. If you did stop smoking, your appetite would improve and you would taste your food much better. Because of this it would be wise to plan in advance so that you're not reaching for the biscuit tin all the time. Remember that putting on weight is an overeating problem, not a no-smoking one. You can tackle it later with diet and exercise.

## And finally...

We hope this letter will help you feel more confident about giving up cigarettes. If you have a go, you have a real chance of succeeding.

With best wishes,

The Health Centre.



Figure 2: Inside pages of an example STOP leaflet