Predictive Affective Natural Language Generation
PhD Thesis Proposal

Saad Mahamood
Department of Computing Science,
King's College, University of Aberdeen
saad.mahamood@abdn.ac.uk
Supervisors: Ehud Reiter & Chris Mellish

1 Introduction

In a digital age, in which instant access to information is becoming more and more prevalent in industrialised society the Internet revolutionises when and how individuals access information. This growth of information has led to research into finding ways in which such information can be tailored to an individual recipient. Such is the growth of information it has resulted in the need of adapting information so that it is personalised on a individual basis to make such information contextually relevant to the user. This is especially the case in domains in which the end user is a non-expert, but the information given to them is highly technical in nature. In addition, particular domains can give information which causes a negative emotional effect such as stress and anxiety. There is a need for computers to be able to reason about emotions and adjust information that can minimise the distress felt by the recipient. One approach of dealing with this particular phenomenon is to allow computers to infer and react to human emotions in an intelligent way so to allow for more positive human-computer interaction when the recipient is distressed. The use of emotions as a component of computer interfaces was a concept that Rosalind Picard coined as "Affective Computing". For machines, the recognition of emotions and distinguishing them from other communicative or social signal signals is a difficult task [74]. However, recognition of emotional states by computers through the use of physiological sensors can give a high success rate of 80% as well [57]. Additionally, research has show that computers are more capable of conveying the correct emotion 70% of the time compared to 50% of the time for a human actor [56]. Nevertheless, the capturing of emotions by computers still has a long way to go in non-controlled enviroments [53]. Picard's work has led to an increasing interest in utilising human emotions as a means of enriching human-computer interactions in many different computing applications.

Traditional Natural Language Generation (NLG) systems have tended to focus on converting non-linguistic data, such as knowledge bases or numerical data, into textual output with the main focus of acting as an information resource for the recipient. However, a growing focus within the NLG community has been to use of NLG systems to represent less rational aspects, such as emotions, of the recipient in a way that allows for a richer quality of interaction. Such systems have been defined as "Affective" Natural Language Generation (ANLG), since they allow for NLG systems that "relate to, arise from or deliberately influence emotions or other non-strictly rational aspects of the Hearer" [12]. The need to include emotion as an aspect of textual output has come about as NLG systems are required to communicate information to a greater and more varied audience, of varying knowledge levels, in which particular recipients can be adversely emotionally effected by information that has an distressing effect.

Whilst there has a been great interest in the field of ANLG and there are many examples of past ANLG systems [29, 31, 13, 16], these systems have tended to have kept the emotional modelling fairly simple or are still "works in progress". Past ANLG systems have also tended not to clearly define the dynamic relationship between the inference of emotion for the recipient and how that inference is translated into action as ANLG output. Additionally, most ANLG systems have had very little or no form of empirical testing to assess their effectiveness [3].

2 Hypothesis

This thesis aims to try and extend the existing work in the field of ANLG by attempting to define a model for varying natural language output based upon the level of stress/anxiety that the recipient is experiencing at a given point in time. Rather than trying to build a complete emotional model to vary textual output by, this work aims to try and model a particular subset of human emotions. The hypotheses for this work can be stated as the following:

1. It is possible to build an predictive model, in the medical domain, that can make a estimation of the level of stress/anxiety that the intended recipient is experiencing by undertaking a situational evaluation that analyses the variables that are responsible for the recipient's distress.
2. By using the estimated level of distress from the predictive model it is also possible to allow for a fine grain inclusion of textual elements that can induce/mitigate affect in the resultant text output. This will allow for the activation of differing ANLG strategies to respond and generate textual output that meets the emotional and informational needs of the recipient and therefore prevents "information overload" at key emotional junctures [4, 69] and additionally any unnecessary distress as well.

Additionally, this work aims to evaluate the effectiveness of such a system by seeing how well it meets the needs of recipients who are in a highly emotional situation by reducing the level of distress experienced, giving recipients a sense of hope to cope with the stressors that they face [7], and finally increasing the level of information recall by the recipient.

3 Motivation

Our motivation for using ANLG as a way of tailoring information to take account of the recipients emotion comes from the need in the medical domain in which non-experts such as parents and patients have to come to terms with information that they may not be familiar with and have to deal with any emotional implications for any distressing information that may arise. Therefore, just like that produced by medical staff, information generated by computers must be able to adapt to reflect the emotional state of the recipient so that sensitivity is shown to the nature of information being presented. However, whereas medical staff can express affect through voice tone and body language, the affect-limited nature of text [55] means that information given to recipients must carry the appropriate affective tone in the way that words are expressed and presented. Allowing computers to use emotional intelligence [21] to express, understand, and to interpret emotions of human recipients [50] could allow end users to interact in a more intelligent way and provide a richer quality of interaction [56]. The use of empathy to recognise and express emotions to efficiently convey the affective tone of information could allow computers to influence the mood of their users [56] and help users to come to terms with information that may have a distressing effect. In effect, this could allow recipients to engage in acceptance and/or problem-solving based forms of coping as a way of dealing with distressing information, which has been shown to be effective at reducing stress and distress [51].

For parents of children in Neonatal Intensive Care Unit (NICU) the need for information that is tailored to emotional and informational needs is very much evident. The birth of a child that requires neonatal care is a particular circumstance that has the potential to cause a considerable amount of stress and anxiety for the parents. For parents the sequence of events in NICU can be akin to a roller coaster ride, with many unexpected ups, downs, and turns of events. Parents rarely feel safe from the fear and uncertainty of the problems that can occur whilst the child is in care [46]. In addition the stress and shock of having a sick child in neonatal intensive care might also mean that parents would not be able to process large amounts of information [4].

Past e-health systems [22, 18], in the NICU domain, have attempted to fill in this gap of information but have not adequately filled the need to address the emotional needs of parents by adapting information in a relevant way. The BabyTalk project [59, 64] is one such e-health system which aims to develop several systems, using an

NLG data-to-text architecture to produce textual summaries of clinical data about babies in NICU. One of these systems, BT-Family, aims to provide information for parents on the condition of their child that is in the hospital's NICU. The BT-Family system provides the appropriate test domain setting for the use of an ANLG system to generate textual information reports as it is a domain that is highly stressful for parents due to multiple differing sources of stress. It is also a domain in which the parents are non-experts as they typically lack familiarity with the technical aspects of NICU care, but nevertheless desire to be informed about the status and condition of their child. Therefore, there is a need for an ANLG system to help cater for the parents desire for information, whilst at the same time understanding how the presentation of such information can have an emotional impact on the recipient.

4 Related Research

ANLG systems are a recent extension to NLG that aims to tailor textual output according to Human emotions. One of the earliest ANLG systems was the work done by Eduard Hovy with his PAULINE system [29] that explored how to generate textual information that could be tailored to a recipient's point of view on subject matters that are potentially contentious. To generate text to recipients' emotional state PAULINE was able to be configured to take into account the emotional state of the recipient with a predefined variable setting. This emotional setting could be defined to "happy", "angry", or "calm" and was selected manually by the operator of the system. Other ANLG systems like Haimowitz's SERUM system, have used a numerical measure to calculate the likely impact of a given textual output [24]. By using the emotional setting and 21 other parameters the PAULINE system used a simple "Affect-Rule" in which enhancer or migrator words/phrases were used to either strengthen a favourable concept or weaken an unfavourable one. In addition to the "Affect-Rule" PAULINE had several other affect related strategies that defined the content selection that allowed for only favourable content to be present in the resulting textual output. The OPADE system created by de Rosis et al. [13], used a not too dissimilar approach for generating "emotionally emphatic responses" for Drug Prescriptions, in which a set of five rules defined the affective responses of the system. These rules placed empathy and/or de-empathsised particular parts of the text that would be unfavourable, whilst placing emphasis on those items that are favourable. However, whilst PAULINE and OPADE contained many interesting strategies for an ANLG system to induce affect, there isn't any empirical measure of the effectiveness of the PAULINE and OPADE systems. Additionally, the emotional model used in both systems relies on the presumption that the recipients emotional state is either considerably distressed or not. No variation is given to varying the level of affective response to the recipient's dynamic level of emotional state.

In the field of Affective Computing there has been a growing interest in building computational systems that can react dynamically to the end user's emotions. However, such approaches have typically been done in controlled environments [53] or have relied on the use of invasive physiological sensors [57]. One alternative to using physiological sensors is to use a probabilistic assessment of the recipient emotional state to determine the type of computational reaction to take. Work in Affective Educational Games by Cristina Conati took this particular approach by using Dynamic Decision Networks (DDN) to situationally evaluate the end user's emotional state by utilising indirect evidence [9]. This included whether particular game playing goals had been met, the end user's personality traits, and the end user's behaviour with the interactive interface. However, the uncertainty of predicting the person's emotions for his/her goals and their perception of particular events makes for a challenging task. Hence, in Contai's work the DDN utilises several physiological sensors that collaborate the predictive assessment. However, whilst progress has been made in recent years to make such sensors less intrusive [56], it still remains unfeasible in many sensitive domains such as the BabyTalk domain. Therefore, a similar probabilistic approach for predicting emotions is needed, but without the reliance of physiological sensors.

This need for predicting the level of emotional distress is particularly important as it can vary significantly for parents of NICU children. After birth to a premature child, parents are thrust into a state of crisis [62]. This in itself can be a frightening experience [40] and can lead to parents feeling overwhelmed with the amount of new information they must assimilate [4, 40, 68]. However, over time parents try and gain an intellectual understanding of their infants condition and treatment [2]. This gradual change suggests that parents information needs aren't static, but instead it is quite dynamic, as the parent's concern moves from infant survival to infant care and personal

coping [4]. However, after childbirth parents can be significantly distressed if they perceive that their child is severely ill [72]. A child with a debilitating condition, such as Down's syndrome, may be physiologically fine, but because of the presence of such a condition the parents could possibly be quite distressed [72]. Parents can also become distressed by seeing their child surrounded by invasive medical equipment [30], noticing colour changes such as jaundice, and also witnessing episodes of apnea or respiratory distress [49, 2]. Past work in clinical psychology by DeMier et al. [14] has shown that a predictive model approach is feasible for identifying mothers that are risk of postnatal emotional distress after giving birth to high-risk infants. However, DeMier's predictive three and two factor models only cover the variables associated with stress at the time after birth and do not cover any variables associated with stress that could occur during the course of the child's treatment in NICU. One other source that could be used in helping to define a predictive model is the Parental Stress Scale: NICU (PSS: NICU) self-assement instrument developed by Miles et al. [48]. This particular instrument has been validated for both US and UK populations [17] and contains many parental NICU stress variables and a rating scale that could be utilised to help define a predictive stress model. Whilst some questions have been raised about PSS: NICU's comprehensiveness [63], it remains one of the few instruments for NICU parents that is widely used. By understanding and appreciating the differing variables that are responsible for the level of distress experienced by parents in NICU, it is possible to build a model that can make a predicative assessment of the level of stress experienced by parents and couple this with an ANLG engine that is capable of varying the level of affective response to the recipient.

5 Research Approach

The diagram below (Figure 1) gives an general overview of the architecture to be used in building an ANLG system that converts data into text:

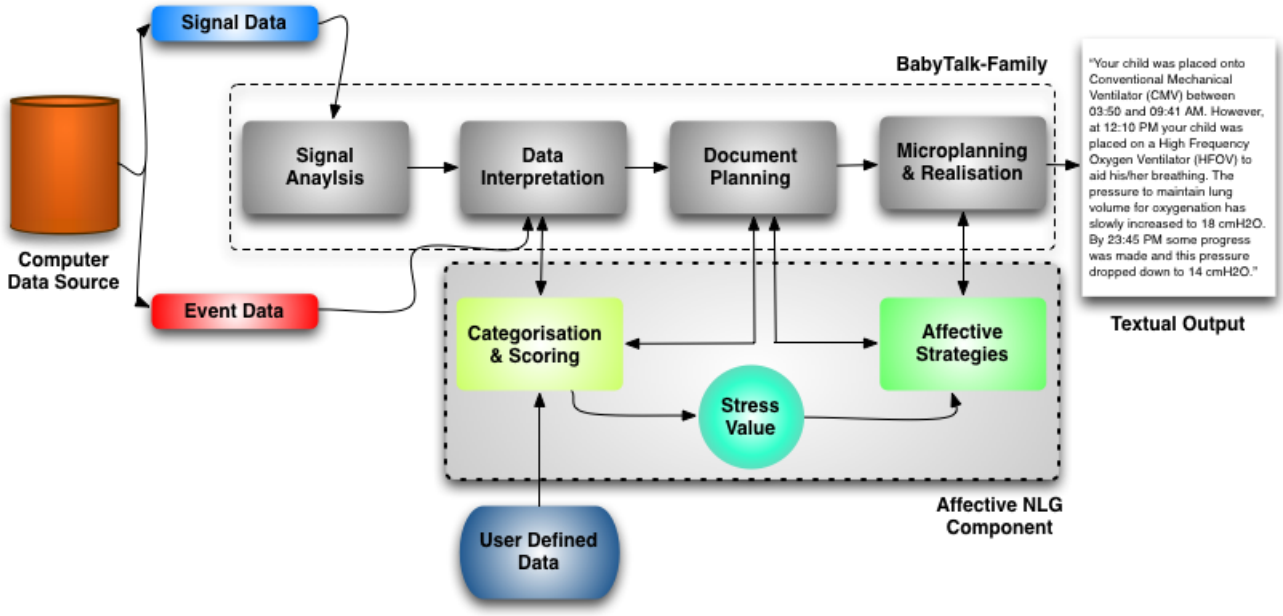

Figure 1: Diagram illustrating the overall ANLG architecture

The ANLG architecture is an extension of the NLG data-to-text architecture that was proposed by Reiter [64], in which natural language text is generated from a non-lingusitic data resource. The primary aim of the Affective extension is to produce textual output that accounts for the emotional status of the recipient and influences the generation of text to meet their emotional needs. The data source for this research comes from the BabyTalk domain, in which continuous multi-channel time series data from physiological monitors and discrete event data such as

the type of ventilation equipment used, staff comments, etc. [59], are used to make a probabilistic assessment of the emotional status of the recipient. Only once this "Emotional Assessment" has been made can "Affective Strategies" be used to add the appropriate affect into the generated textual output.

The subsequent two sections will describe how the "Emotional Assessment" will be performed and the manner in which the "Affective Strategies" will be used.

5.1 Emotional Assessment

To induce relevant affect through textual output the first part of the ANLG system must be able to make a predictive emotional assessment of the recipient. In the BabyTalk-Family domain, this is done by making an assessment of the levels of stress/anxiety that parents may have by interpreting the medical records of the baby, since one of the main factors of parental stress is the physiological heath of the child [72]. For the BabyTalk-Family domain these records are contained within an database (called Badger), which not only contains the signal and event medical data for the child, but also supporting data that describes details about the mother, birth details, parent-staff interaction notes, and staff observation notes. It is possible by combining these details together, in a coherent way, to gain a detailed insight of all the potential stressors faced by parents at given point during the child's neonatal care. The diagram below (Figure 2) illustrates the architecture for the stress predication subsystem:

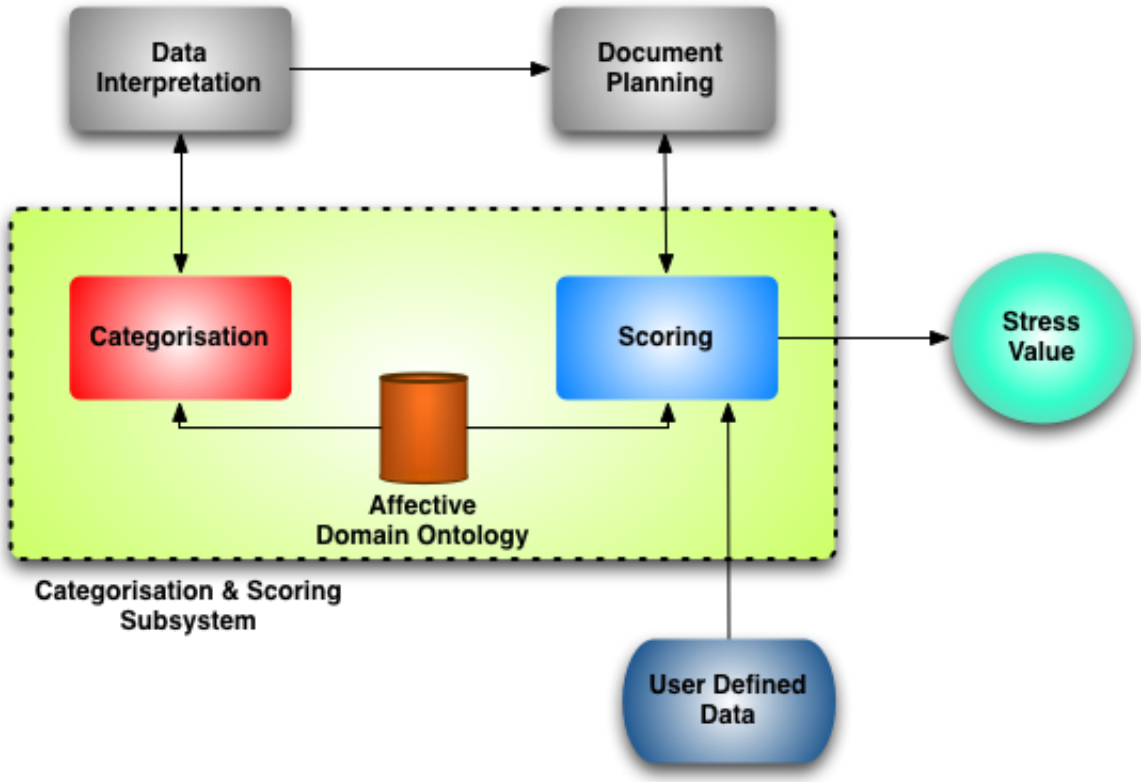

Figure 2: Diagram illustrating the stress prediction subsystem

The main feature of this subsystem is the "Categorisation" and "Scoring" components that categorises medical records and also performs an additive calculation on the level of stress that the parents are experiencing. There are two main categories that the Badger medical records can be organised into:

1. Data pertaining to the baby's physiological condition:
    - Baby's heart rate
    - O_2 consumption rate
    - Ventilation type
    - etc.
2. Data that relate to the parents emotional state:
    - Free text comments by clinicians
    - The type of delivery method used
    - Type and effect of drugs administered to the mother
    - etc.

The baby's physiological performance is one of the major indicators of how well a child is doing in care over a given time-frame. For most parents there is direct link between the level of stress that they are experiencing and the the baby's physiological state [72]. However, in some cases the physiological well-being of the child can differ from the parents perception of how the child is doing [72]. A child with a disease or disability could potentially mean that a parent would be more stressed regardless of how well the child might be doing physiologically. This can occur when less serious ailment, such as Jaundice, have a dramatic effect on the physical appearance of the child. Whilst the data describing the child's physiological performance shows that the child is stable, there is a divergence between this description and the distress experienced by parents who witness significant change in the appearance/behaviour of their child [49].

One way of dealing with such a situation is to take into account data that relate to the parents emotional state. Past Affective Textual Interaction systems have used keyword spotting as a way of determining the emotional state of those that use the system [41]. In a similar fashion, free-text comments by medical staff on the parents behaviour when visiting the neonatal unit can give key indications of how distressed the parents might be, as the example below illustrates:

(1) "Discussed with parents baby's current medical condition. Condition is thought to be very poor and getting worse. Parents very distressed by this news."

In addition to the parent's emotional state being influenced by the physiological condition of their child, it is also influenced by how stressful the birth of the child was, especially if the birth of the child diverges greatly from a normal child birth and is quite invasive [49]. This distressful event is also a time in which the parent's ability to absorb information is limited [4]. Therefore, each event must be ranked to give it a value that relevantly expresses the level of distress experienced by parents. The stress value scale used by the system is based on the PSS: NICU self-assement instrument developed by Miles et al. [48], which uses a five point scale to indicate the severity of the stress being experienced by the parents:

- 5 – Extremely Stressful
- 4 – Very Stressful
- 3 – Moderately Stressful
- 2 – A little stressful
- 1 – Not at all stressful

The perception of stress by parents is subjective [71] and changeable over time. Therefore, the rating of individual NICU elements should not be fixed to a stress scale value and should go up and down depending on the circumstances involved. This is particularly the case as physiological development milestones for the child in NICU are met or when setbacks occur [43]. However, whilst past literature has indicated general areas of NICU care that can cause stress for parents, at present no current literature has managed to give a numerical score for specific individual elements of NICU care and environment. Therefore, it is envisioned that an experiment will be carried out to allow former NICU parents to subjectively rate the stress caused by individual elements of NICU care in much greater detail using the PSS: NICU scale. The example below illustrates what this might look like for the types of ventilation equipment used when a child is in NICU:

- High Frequency Oxygen Ventilation (HFOV) - (Stress Score – **3-5**)
- Synchronised Intermittent Mandatory Ventilation (SIMV) - (Stress Score – **3-5**)
- Conventional Mechanical Ventilation (CMV) - (Stress Score – **3-5**)
- Continuous Positive Airway Pressure (CPAP) - (Stress Score – **2-3**)
- Nasal Prongs - (Stress Score – **2-3**)
- Self-breathing - (Stress Score – **1**)

The variability in the stress score for individual NICU elements, like ventilation equipment, indicates that for some elements a static score is in itself not representative of the level of distress that is experienced by parents. Additional related factors must be taken into account such as the level of pressure, oxygen concentration, the perceived invasiveness of the equipment, the use of nitrates, etc. when calculating a potential stress score.

After each element has been ranked by the system are then combined together and a mean value is taken to indicate the level of stress in each of the two categories. An example of what this might look like is shown below:

- Baby's Physiological Condition – **4**
- Parent's Emotional Condition – **3**

From the two category scores a final stress score is derived by taking the mean of the two category scores. It is this final score that will indicate the level of probable stress experienced by the parent at a given point in time of the child's NICU care. However, environmental, background, and sociological factors can also have an impact on the final stress score and can influence the score up or down. For mothers, higher levels of education can correlate with higher anxiety scores [72] and is a background factor that influences the final stress score up higher. Additionally. the other possible factors that could have an influence on the final score are:

- The distance that parents are from the hospital
- Level of medical expertise
- Prior experience of NICU care
- Number of children in NICU care
- Linguistic Preferences
- Parent's Personality
- etc.

Some data items like the parent's personality and their previous experience with NICU care can only be obtained by asking the intended recipients themselves directly. It is hoped that using the above optional factors allows for a more accurate stress value calculation. However, access to such additional information might be difficult to achieve in practice.

5.2 Affective Strategies

Once the stress value has been calculated the question remains how to generate textual output that invokes or mitigates emotional affect on the recipient. Past work in ANLG has shown that various strategies can be used in the generated textual output to achieve this aim [29, 13, 12]. Based upon this prior work it is envisioned that the BabyTalk-Family ANLG system will use the following affective strategies:

- Re-ordering Material
- Explanatory Justification
- Additional Information
- Level of Detail Adjustment
- Additional Reassurance Adjustment
- Typographic Adjustment

The first strategy, "Re-ordering Material", is inspired by Hovy's second evasion rule in which a good effect is obtained by the juxtaposition of neutral topics with good ones in enhancer phrases [29]. Additionally, those topics that could have an negative effect on the recipient would be placed among any supporting positive topics. For example, if a child has a minor respiratory setback such as the temporary increase of oxygen, this negative information could be offset against positive information that states the overall long-term good progress that the child has made since being admitted to NICU.

The "Explanatory Justification" strategy is a strategy that aims to provide additional explanatory textual information to parents for why particular medical actions have occurred. For example, if a child is moved from one ventilation equipment to another that could be viewed by the parents as a negative reflection on the child's well being, then reassurance could be provided by the system in the form of a explanatory justification such as: *"However, in the best interests of your child [...] to give additional support to your child's breathing."* This strategy is quite similar to first rule of the "emotionally empathetic responses" employed by de Rosis et al. [13] in which empathy is added to any relevant item that is unfavourable to the recipient.

The third strategy aims to provide additional information to the recipient for particular medical concepts that the parent may not be familiar with or would not likely be able to recall from their consultations with medical practitioners. It is hoped that the additional provision of information about the exact nature of the care given to their child will enable parents to have a better understanding of the medical situation that their child faces. Additionally, the provision of such extra information could prevent parents from feeling that they are not being told all they can and thus would lose hope or assume the worst [7]. However, during very stressful situations such a strategy cannot be deployed as high levels of distress could limit the information that parents can absorb [4]. Future qualitative work with NICU parents will hopefully identify areas of information that parents have difficulty understanding or recalling during their consultations with medical consultants and their relationship to the level of emotional distress experienced by the parents.

The "Level of Detail" strategy aims to reduce the complexity of the information that parents are presented with when they are experiencing very high levels of distress. This is particularly important as parents can feel during such circumstances that they are overloaded with more information than they can assimilate [4]. In particular this strategy has four particular effects when used:

1. Removal of acronyms from the textual description.
2. Removal of specific time values from the the textual description – e.g. 1:30 PM turns into "In the afternoon..."

3. Removal non-salient information from the textual description – Any information that is not directly related to the domain object will be removed.

4. Shorter sentences, shorter paragraphs, and a simpler lexicon.

Each of these four effects can be used independently of each other, which allows for the more drastic effects such as the third and fourth ones to be only used when the parent is potentially extremely distressed.

The fifth affective strategy used by the system aims to provide additional reassurance phrases in the resultant textual output. Phrases like "Your baby has made good progress today" could be used by the system to reassure parents that physiological progress is being made by the child in which positive outcomes that have occurred are communicated. Such a strategy can be also used to help parents cope with the distress of seeing their child having a temporary downturn by helping to reassure parents that the baby has made significant progress over the long term: "Over a period of time you can see that your child has made quite definite progress in the right direction." However, if the child has a longer downturn in medical progress and is not a temporary phenomenon then a third type of reassurance phrases could be employed by the system to reassurance parents that the medical practitioners are doing all they can to help the baby: "We will continue to support your baby and hope that your chid will improve in the near future."

The last affect strategy aims to change typographic parameters for the text that is being presented to the recipient. Previous work by Sánchez et al. has demonstrated the use of differing font type and colour selection for a range of affective states in a instant messaging application [70]. A similar approach could be undertaken here as well, in which different font weights and colour could be used to emphasise key pieces of information, such as emphasising positive information and any reassurance phrases that occur in the textual output. Additionally, an overall increased font weight and increased line widths and character spacing could help increase the readability of text [1] for parents that are considerably distressed.

5.2.1 Scaling Affect for Differing Stress Levels

Differing levels of stress can also determine which affective strategies are used and their intensity. This is particularly important as it is possible that excessive use of affective strategies could lead to the recipient feeling patronised and thus it must be scaled to the level of distress experienced by the parents. Previous evaluation work on inserting hedge phrases into NLG text has found that overuse can lead some recipients to feel that the text is "condescending" and "wishy-washy" [42]. Therefore, to prevent the same occurrence of sentiment with parents the ANLG strategies described in the previous section must be activated and deactivated at particular stress values. Table 1 below shows the minimum stress value need to activate a particular affective strategy and the value in which that an affective strategy is deactivated and no longer used:

Affective Strategy	**Stress Value Activation**	**Stress Value Deactivation**
Re-ordering Material	2 *and greater*	—
Explantory Justification	2 *and greater*	—
Additional Information Adjustment	1 *and greater*	4 *and greater*
Level of Detail Adjustment	3 *and greater*	2 *and lower*
Additional Reassuarnce Adjustment	3 *and greater*	2 *and lower*
Typographic Adjustment	4 *and greater*	3 *and lower*

Table 1: Table of Activation Strategies with their Activation and Deactivation Stress Values

Additionally, for some of the affective strategies the intensity of the effect employed can increase as the stress levels go up. For example with the Level of Detail Adjustment strategy this could look like the following:

- Stress Value 3 – Minimal detail removal – Removal of all acronyms and specific time values.

- Stress Value 4 – Medium detail removal – Same as above and less detail provided on non-sailent information.
- Stress Value 5 – Maximum detail removal – Same as minimal detail removal, no detail provided on non-sailent, shorter sentences, shorter paragraphs, and simplified lexicon.

By scaling some of the affect strategies in such a fine grained way it is hoped that we can prevent the excessive use of affect that is present in the textual descriptions, and thus prevent any misalignment between the recipients emotional state and the disproportionate use of affective strategies in the resultant textual output. However, for some parents the use of any affect strategies regardless of how it is scaled may be found unnecessary. For such parents it is envisioned that an override will be provided to turn off some or all of the affect strategies.

6 Deliverables

At the end of the completed project it is hoped that a computational prototype system will be developed based upon the principals discussed in the previous section. It is also hoped that this prototype system will provide a general model of how to build ANLG system for various domains in which information that could have an emotional impact needs to be communicated. Additionally, empirical evaluations of the complete prototype will present a complete picture on the effectiveness of the system.

7 Evaluation

The evaluation of the two subsystems described in Section 5 are described in the succeeding sections. The importance of an effective evaluation in this thesis stems from the fact that past ANLG systems have tended to have little or no evaluation of their effectiveness. Conducting an evaluation on the constituent parts that make up the proposed ANLG system, will provide an important and valuable scientific contribution to the ANLG field.

7.1 Emotional Assessment Evaluation

For the evaluation of the emotional assessment it is hoped that two separate evaluations will be conducted:

1. An expert evaluation comparing the stress value generated by the system for specific scenarios and an expert's individual assessment of the stress level.
2. User evaluation with parents that have a child in NICU.

Evaluating the effectiveness of the Categorisation & Scoring subsystem will be done through an expert evaluation that would compare the predictive stress score generated by the computer against the stress scores calculated by human authors. Additionally, it is hoped that a second user evaluation will be undertaken with parents that have a child in NICU to compare the stress value generated by the system from their individual child's medical records against the results obtained from self-assessment questionnaires such the Spielberger State-Trait Anxiety Inventory (STAI), Edinburgh Postnatal Depression Scale (EPDS), and the Parental Stressor Scale: NICU (PSS: NICU). Conducting these two evaluations it would allow for not only for the emotional assessment model to be tested for effectiveness but also for the model to be calibrated as well.

7.2 Affective Strategy Evaluation

It is envisioned that a prototype of the affective strategy subsystem will be evaluated in a single evaluation to assess it's effectiveness. This evaluation aims to compare the generated affective text against human and non-affective computer texts. A range of scenarios would be used in the evaluation by selecting a range of differing baby NICU outcomes that range from the best case to the worst. This will allow of the affective strategy subsystem to utilise all of the strategies that are defined in the range of texts that it produces. Additionally, it is hoped that this evaluation will be conducted on the three separate populations:

- Parents without any NICU experience
- Former NICU Parents
- Parents that are going through NICU

For each of these populations questionnaires would be used to measure the levels of satisfaction and the quality of the information being presented for each of the text types. Also, each of the evaluation participants would be asked to complete self-assement questionnaires that measure the level of distress when reading an outline summary of the scenario and afterwards when the participants select their favoured textual representation (ANLG text, human text, or non-ANLG text). It is hoped that by using such an evaluation the relationships between affect strategies and the type of scenarios that had a beneficial effect on end recipient satisfaction and level of distress will be discovered.

8 Work Done So Far

An experiment was conducted that compared hedged text and non-texts in the domain of producing a computer-generated summary of exam results [42]. The hedged texts contained evaluative adverbs as the hedge phrases, which were inserted by the NLG application in the front, or back, or in both locations in a given proposition on a stochastic basis. An evaluation was undertaken to compare the preferences for hedged and non-hedged text. A total of 37 Masters students were given questionnaires in which they were asked to assess their preferences between the hedged and non-hedged text. The results showed that on the whole most people preferred the non-hedged text. However, further analysis showed that if the participants were grouped by whether they were native to English the results were different. Those speakers that were not native to the English language were more likely to prefer the hedged text compared to the non-hedged text, whereas native English speakers were more likely to prefer the non-hedged text. One of the typical comments of those speakers who were native speakers was that they preferred the non-hedged text compared to the hedged text since the hedges were perceived to be adding opinion to the text or were just "wishy-washy". In contrast those speakers who were non-native speakers felt that the hedged text was more "humanised" than the non-hedged text. Whilst these early results show some significant effect of hedges, more work is required to find the appropriate use of hedging for those who are native speakers. This experimental work was written up as my first short paper and was accepted for publication at the European Natural Language Generation (ENLG) 2007 workshop.

In addition to the hedging experiment, a qualitative piece of work was undertaken in late 2007 that explored the informational and emotional experiences that former NICU parents had when their child was being looked after [43]. A total of 9 parents, all from a middle-class background were interviewed for this study. The interview recordings were transcribed and analysed using the sociological principle of Grounded Theory [20]. The results obtained from this study revealed that the parents emotional and informational needs during the time their child was in NICU care was dependant on the baby's physiological condition, factors that affect the parent's emotional state, and the parent's goals for the physiological development of their child. Based upon these results, an initial affective model was proposed that took into account the factors found in the study when generating information for NICU parents. The results from this preliminary study founded the basis for many of the ideas that are present in Section 5. However, more work is needed to see whether the factors that were identified in the study by former

NICU parents are the same ones as those identified by parents whilst their child is in NICU care. Additionally, more work is needed to see if the factors identified in the study differ from those parents who have a lower socio-economical background. It is hoped that future qualitative work that is planned with NICU parents at Edinburgh Royal Infirmary Hospital will try and explore these two open questions indirectly. This study was written up a short paper and was accepted for publication at the Computer-Based Medical Systems (CBMS) 2008 symposium.

9 Workplan & Deliverables

Tables 2 & 3 shows the thesis workplan and deliverable milestones for the next 16 months, which starts from June 2008:

Months	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
Activities																
BabyTalk-Family Core Implementation																
Categorisation & Scoring Implementation																
Affective Strategies Implementation																
Edinburgh Research Study																
Evaluations																
Thesis Write up																

Table 2: Gantt chart illustrating the workplan

Deliverable	**Status**
1st Year Report	Completed
Publication	Poster Presentation at ENLG 2007
Thesis Proposal	Completed
Publication	Short Presentation at CBMS 2008
Publication	Conference Paper – End of 2008
Publication	Journal Paper – Mid to End of 2009
Thesis	Due October 2009

Table 3: Research Deliverables

Before any of the ANLG system can be implemented the core BabyTalk-Family system (as illustrated in Figure 1) must first be built. This will involve adapting an existing data-to-text framework called BT-CORE, which is the foundation for all of the data-to-text BabyTalk systems in the NICU domain [58]. Once this work has been completed then the two core ANLG components will be subsequently implemented.

Besides implementing the prototype ANLG system it is also expected that at least four months will be spent collecting qualitative and quantitative data on parents that have a child in neonatal care. The Edinburgh Research Study aims to collect data on four, different, but related aspects of NICU care from the parents perspective:

- The nature of communication between NICU consultants and parents (Qualitative)
- The parents linguistic preferences for communication (Quantitative)
- The level of emotional distress experienced by parents in NICU (Quantitative)
- An understanding of the parents socio-demographic background (Quantitative)

Collecting information on these four aspects will hopefully give a better understanding on the types of information given by consultants to parents during consultations in which parents find they have difficultly understanding

particular medical concepts and/or misinterpret the significance of the information given to them. For example the word "prognosis" may have a positive or negative connotation depending on the recipients degree of understanding. Past research by Perlman et al. [54] has shown that there is a significant difference in the type of information that is emphasised by consultants and parents. Additionally, the parents level of understanding of the information given in consultations could be influenced by the level of emotional distress that they are facing. Therefore, for the ANLG system it could be quite beneficial to have a better understanding of the types of information that parents may need additional help in recalling by placing emphasis on such types of information. This could enable parents to have a better understanding and remind them of what was spoken in the consultations [32]. In addition, the Edinburgh study will explore the relationship between the level of medical information recall, the level of emotional distress, linguistic preferences, and the parents socio-demographic background to see if any significant relationship exists between these elements. By collecting the socio-demographic details and linguistic preferences of the parents it is hoped that valuable insights will be collected that outlines the type of language presentation that parents favour and an understanding of the socio-demographic status of the parents that attend the NICU. The study also provides an opportunity to evaluate the effectiveness of the stress model by comparing the stress score generated by the system against stress & anxiety scores generated by PSS: NICU and STAI self-assement instruments. The collection of PSS: NICU and STATI scores will also provides a "gold-standard" since the circumstances for the parent's perceived level of stress & anxiety will be known. By utilising this "gold-standard" it is hoped that the evaluation of the stress model will also be much better fine-tuned in scenarios where the model differed considerably from the self-assement instruments.

By performing such knowledge acquisition [66], on the four aspects outlined above, it is hoped that the proposed ANLG system can be refined/improved from the insights that are gained in the study. These improvements will include the following:

- An understanding of the types of information that will be need be emphasised/de-emphasised when providing additional information to parents and it's relationship with the parents level of emotional distress. In particular, this will help to inform the Additional Information and Level of Detail affective strategies implementations.
- Allowing the stress model to be evaluated and calibrated against a "gold-standard".
- Modifications to the textual output to meet the linguistic preferences of the parents.

The findings from the study will hopefully be implemented into the ANLG components in Month 7 after the study's conclusion.

A Literature Review

There are four main areas that this literature review will be covering. This is Affective NLG, Hedging, Doctor-Patient Communication and Neonatal Intensive Care Units (NICU) and Parent Communication. The review on Affective NLG will look specifically at two relevant past Affective NLG systems and explore the mechanisms that previous researchers have used to add affect into generated text. Whilst the review on Hedging will concentrate on the concept of Hedging, the use of Hedging in medical domain, and related research on the effects of Hedging. Both Affective NLG and Hedging are important subjects in understanding the past developments in these areas and their implications for the development of an NLG system that has the ability to add emotion to text where it is appropriate to do so. The area of Doctor-Patient communication will take a general look at the essence of information needs that patients have and the manner in which doctors convey information. A greater in depth look is taken in the final section that explores the nature of communication between medical staff and parents in NICU's. This last area of the review is particularly important since it helps to set the context for which this research is being conducted in and also help inform the design and implementation of an Affective NLG system that aims to communicate information for parents who have children in NICU.

B Affective Natural Language Generation

Natural Language Generation (NLG) is a process in which computers are able to produce textual output from some underlying non-linguistic source of information [65]. As the use of real-world NLG systems have accelerated there has been a growing interest in getting computers to generate text that allows for the same piece of information to be tailored to the recipient and the situation. One particular way of tailoring such information that has been recently explored by the NLG community is through the use of emotion. This has led to the development of Affective Natural Language Generation (ANLG) which is a particular subset of NLG that has been defined as "NLG that relates to, arises from, or deliberately influences emotions or other non-strictly rational aspects of the hearer" [12]. In other words, it is a form of textual output from a non-linugustic source like NLG, but unlike NLG it takes into account the emotional aspects of the recipient that is being given such text.

A comprehensive review of ANLG was addressed by Anja Belz [3]. However, among the papers that Belz reviewed she found that there has been very few research papers written on the subject. Whilst there have been very few pieces of work on ANLG there is however two particular relevant past ANLG systems in which both contain many ideas on the application of affect in text that will be reviewed in the subsequent sections.

B.1 PAULINE

PAULINE (Planning And Utterance Language In Natural Environments) was an ANLG system created by Eduard Hovy for the purpose of generating natural language that was tailored pragmatically and stylistically to the end recipient's point view on a subject matter that is potentially contentious. By taking the conversation topic, conversation setting, and the interpersonal goals of the recipient that the text is being given to into account the system was capable of producing textual output that was tailored exactly for the given pragmatic constraints. All of these pragmatic constraints were defined through a set of 23 features that determined the scope of the interpersonal goals and conversational setting that was involved. Of particular interest was the interpersonal goals that could be selected in PAULINE was the desired affect to be deployed in the textual output to influence the emotional state of the recipient. This setting could be set to "angry", "no effect", or "calm down". Additionally, PAULINE could also be configured to take account of the emotional state of the recipient as well as part of the conversational settings which could be set either to "happy", "angry", or "calm". These two features in combination with the other 21 features would determine the amount of affect to deployed to slant the textual output to the recipients favour. In particular, PAULINE would use an underlying "Affect Rule" that acted as the basis of all affect-related generation decisions [29]. This "Affect-Rule" used enhancers and mitigators which were words or phrases that either strengthened or weakened the affect of a concept. It's application by PAULINE was defined in the following

manner:

- for a GOOD effect, say GOOD topics with ENHANCERS and BAD topics with MITIGATORS
- for a BAD effect, say GOOD topics with MITIGATORS and BAD topics with ENHANCERS

In addition to the "Affect-Rule", PAULINE also had two strategies for the deployment of affect when determining what content to give to the recipient:

- Evasion
- Selectivity

The Evasion strategy allowed PAULINE to indirectly deal with subject matters that were defined to be contentious to the recipient. This was done by hinting, implying, or referring to something that is in some way related, and trust that the end recipient would make the final inference alone [29]. This particularly strategy used the following rules to wishfully suppress and mitigate particular subject matters [29]:

1. Say GOOD topics
2. Juxtapose NEUTRAL topics with GOOD ones in enhancer phrases
3. Leave out BAD topics altogether, unless they can be mitigated using mitigator phrases and words, or unless they are central to the story

Using a fictional example of where two characters are involved in an altercation, that results in one character getting hurt, Hovy demonstrates how PAULINE would generate textual output for the character that got hurt:

(2) JIM, A NICE FELLOW, COULD NOT EXPECT THAT UNPLEASANT MIKE WOULD BE HURT IF JIM ACCIDENTALLY BUMPED HIM; ALSO, A REASONABLE PERSON COULD NOT FORSEE THAT IF HE BUMPED HIM JIM WOULD HURT HIM. HE HAD NO INTENTION TO BOTHER MIKE. MIKE WAS ONLY BUMPED BY JIM ONCE. [...]

Since for Jim there is very few negative topics PAULINE instead tries to emphasise the severity of the altercation by juxtaposing netural topics aganist good ones with enhancer phrases such as the "*Also*, a reasonable person..." However, for Mike since all of the topics are negative PAULINE directs the wishful suppression and mitigation strategies to say the following single line:

(3) THAT NASTY JIM BUMPED MY FRIEND MIKE HARD AND HURT HIM.

The second strategy that PAULINE had for determining what content to give to the recipient is use of selectivity. Instead of trying to evade unpleasant subjects, PAULINE can select particular parts of said subject whilst suppress any parts that may have a negative connotation. This was implemented through the use of three plans. These plans are called DESCRIBE, RELATE, and CONVINCE. The DESCRIBE plan was used to describe objects; the RELATE plan was for relating events and state-changes; the CONVINCE plan which is used to help win the recipient over to the speakers opinion of the topic when they disagree [29]. It is in the last plan in which affect is used [27]. Hovy defined several rules that defined the affect strategies that PAULINE would use to slant the textual output for the recipient:

- Affect-bearing Phrases
- Affect-bearing Adverbs and Verbs
- Sentence Clause Order
- Verb, Adjectives, and Noun Choice

The affect-bearing phrases used by PAULINE where a "not only X but Y" sentence form to imply that X and Y carry the same affective value through the juxtaposing of two affect-laden sentences. This affect of juxtaposing two affect-laden sentences can be strengthened or weakened through the use of enhancers (Example 4) or mitigators (Example 5) [29]:

(4) "Pete played the game; *in addition*, he hit five home runs"

(5) "*Although* John remembered the cat, he whipped the dog"

In addition to Affect-bearing phrases, a number of adverbs and verbs can also participate in the affective manipulation of text. Adverbial stress words such as "really", "just", and "only" for actions; "very", "extremely", "slightly", and "moderately" for states and adverbs [29] can also induce affect by acting as enhancers or mitigators. The use of these words only allow for the modification of concepts that carry any pre-exisitng affect and cannot be used in extreme states/actions which can't be mitigated effectively.

The third strategy used by PAULINE for creating affect is the use of Sentence Clause Order to organise particular parts of a sentence such as the choice of subject and the order of clauses. The choice of subject enabled PAULINE to choose the subject focus in a given sentence which allows good candidate sentences to be paired with good affect and vice-versa for bad candidate sentences. Clause Order Organisation, on the other hand, allowed PAULINE to place adverbial enhancers or mitigators within a noun group. The placement of the adverbial modifiers could be placed before the noun (pre-nominal), after the noun (post-nominal), or in some cases in both positions. The decision of where to place these modifiers were decided by the speaker's rhetorical goals of style and opinion [29].

The final method for PAULINE to induce affect is to select appropriate Verbs, Adjectives, and Nouns. Verb selection by PAULINE also plays a role in determining the severity of a particular action by either enhancing or mitigating the action that was performed. The following examples illustrates this distinction between enhancer and mitigator verbs [29]:

(6) "Jane *slammed*/*tapped* me on the head with a crowbar"

(7) "Mike *wolfed down*/*nibbled* his supper"

Like Verbs, different Adjectives was used by PAULINE to either enhance or mitigate textual statements by using adjectives like "concerned", "wonderful", "nice", and "nasty" to express the speakers opinion.

Nouns were also used by PAULINE to express the speakers opinion. PAULINE allows the speaker to express an opinion of an actor when building a noun group by selecting a mitigator or enhancer [29]:

(8) THAT JERK, JIM ...

(9) THAT GENTLEMAN, JIM, ...

B.1.1 PAULINE Conclusion

Unfortunately, whilst PAULINE contains ideas of how to induce affect, Hovy didn't demonstrate with any empirical measure the effectiveness of all of these implemented affective NLG strategies. No indication is given by Hovy of how effective textual output generated by PAULINE when given to test subjects. Hovy also never states precisely what the input representations of the texts looks like [27]. Additionally, large parts of the PAULINE system are based on a huge number of complex rules that decide when to activate goals and strategies, which is implemented as an interleaved planning and realisation regime. However, having all these complex rules means that the system was difficult to maintain and extend. The many fine grained results presents another problem; It's difficult to see how interaction between rules and therefore undesirable side effects can be predicted [3].

Nevertheless, PAULINE does contains several interesting ideas of what strategies a possible ANLG system could use for strengthen or weakening textual content through the use of enhancers or mitigators. This could such a potential ANLG system to strengthen claims with enhancers that could have a potential positive affect, whilst weakening any claim/statement with mitigators that could have a negative affect on the recipient. This could be done using affect-bearing phrases, adverbs, and verbs. Sentence Clause Ordering could allow for the structuring of textual content that is favourable to the recipient by helping them to focus on positive subject aspects, whilst de-emphasing negative ones. Also, the use of Selectivity and Evasion strategies could help with content selection for some ANLG domains, but in others such as a medical domain, it could also be a ethical liability if critical information is avoided or selectively placed. However, any implementation of these affect strategies may differ in a ANLG system that uses a more complex emotional model than the simplistic three-mode emotion model that was used in PAULINE.

B.2 Affective NLG for Drug Prescriptions

Research work by de Rosis et al. [13] created an NLG system, called OPADE, that aimed to produce user-adpated text of drug prescription explanations by utilising the pre-exisitng European Drug Database. One of the areas of focus in the OPADE system was to generate texts that were really convincing to the recipient. However, the evaluation of the initial OPADE system found that there were problems with the knowledge representation, which led to problems such as the generated text did not provide additional information when describing severe side-effect of particular drugs. Also the use of pre-exisiting data-sources, in the OPADE system, limited the grammatical and lexical choices and reduced the possibility of the system to add elements of empathy, empathising-de-empahsing techniques, and 'user-tailored' definitions of medical terms [13]. This need of introducing affect came from an analysis of two differing texts provided by two doctors to see the types of "affect-conveying" techniques used by the doctors. From their analysis they found that the first text focused on convincing the patient to follow the specified treat by relying on the patient's sense of responsibility. The second text, unlike the first text, tried to motivate the patient with the goal of reassuring him/her about the perspectives of success of the treatment [12]. However, a similarity between texts was the manner in which both doctors used the "first plural person" form of verbs to give the patient the feeling that caring for the diagnosed disease concerns the patient and doctor jointly [12].

To solve the problems found in the first iteration of the OPADE system de Rosis et al. proposed a new model for the system. Interestingly one of the proposal to improve the system included a series of heuristics, similar to those employed by Haimovitz [24], to produce "emotionally empathetic responses". A set of five rules was defined by de Rosis et al. to produce affective responses which included the following [13]:

1. When describing a the patient status if there are any items that are relevant for patient compliance that may be unfavourable THEN add empathy to the description of those items.
2. When describing a the patient status if there are any items that are relevant for patient compliance that may be unfavourable AND the patient is anxious THEN de-emphasise these items.
3. IF the patient is anxious when trying to make a request to administer an individual drug THEN add empathy.

4. When trying to explain favourable, positive effects of a drug, that is relevant for patient compliance, THEN emphasise these items.
5. When trying to explain unfavourable, negative effects of a drug, that is relevant for patient compliance, AND the patient is anxious or non-cooperative, THEN de-emphasise these items.

In addition to using heuristics to apply affect to generated textual output de Rosis et al. also proposed the use of repetition to generate affect: "Repetition plays a crucial role, for instance, in minimising concern about side-effects, and contributes to the persuasion process [13]." In order to generate repetitions to maximise the persuasion effect two particular heuristics were proposed by de Rosis et al. [13]:

1. When explaining the negative effects of a drug that includes favourable items with equal values THEN emphasise them by stressing these repetitions.
2. When explaining the negative effects of a drug that includes unfavourable items with equal values AND the patient is anxious towards the pain THEN aggregate these repetitions.

Whilst de Rosis et al. have proposed and implemented changes to the initial version of the OPADE system no subsequent evaluation of the improved system was undertaken. One of the main problems with the OPADE system was that it's lack of domain independence. In a related paper, based upon their experiences with the OPADE project, de Rosis and Grasso [12] tried to define what a domain independent ANLG architecture may look like. De Rosis and Grasso comes to the conclusion that one of the primary factors to generate "Affective" texts is that "All phases of the text generation process should be influenced by consideration of emotional factors", which they propose is to be achieved through the modification of a classical pipeline NLG system [65]. However, whilst their architectural proposal considers a theoretical implementation, the real-world results produced are considerably vague due to it being "Work in progress". This is confirmed by the papers conclusion that states that more research is necessary in this particular field.

B.3 Affective NLG Conclusion

In essence the emotional modelling for the above systems have kept fairly simple or are still "works in progress". A lot of the research in Affective NLG gets close to making the connection between emotion and NLG output, but instead boils down to preparing the ground work for such a connection [3]. However, in some cases the approaches to emotion modelling and emotional generation are based upon the researchers' intuitions. In other cases they were based upon existing psychological and cognitive theories of human emotion. However, either case there hasn't been any form for empirical testing to determine which affective NLG techniques have any benefits for the intended target audience. This lack of empirical testing makes it hard to determine the effectiveness of past affective NLG implementations and ideas. Nevertheless, there are many interesting ideas of how to induce affect, but if such ideas are to be used in any future proposal of a possible ANLG system it would need some form empirical testing to asses its effectiveness.

C Hedging

The term "Hedge" was first coined by G. Lakoff in his 1973 paper; *Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts* [34]. The definition that Lakoff gave for hedges, paraphrased by Kristen Prechet, was "vague boundaries and fuzzy edges that [...] will very often be neither true or false, nor nonsensical, but rather true to a certain extent and false to a certain extent" [60]. One interesting observation is that Lakoff briefly points out the possibility that hedges may "interact with felicity conditions for utterances with the rules of conversation" [34, 44]. Later definitions of hedging came from observations by Brown and Levison [6] as expressions that modify the category membership of a predicate or noun phrase [44]. However, the definition has been widened

in part due to the observation that certain verbs and syntactic constructions conveyed hedged performatives [44]. Verbs such as *suppose, guess, think*, which can act as a hedge when inserted in a sentence [44], are examples of these hedge performatives in which the speaker attempts to perform a particular act. However these verbs have also been defined as hedges that affect the amount of commitment which a speaker may have to the truth-value of a particular sentence [44]. Prince et al. [61] in their analysis on hedging found that hedges can be be classified into two distinct groups with respect to their effect on the truth-value.The first are *approximators* which do modify the truth-condition of the proposition (e.g. I *suppose* the sky is blue). The second is *shields* which unlike *approximators* do not affect the truth-conditon of the proposition, but instead show the amount of commitment that the speaker/writer has to the proposition (e.g. I *think* his feet are blue) [44]. It has been observed that other English discourse particles such as, *well*, *oh*, *ah*, *so*, *anyway*, *actually*, *still*, *after*, *all*, can act as 'maxim hedges' [38], a reference to the Gricean maxims of conversation. R. Lakoff [35] showed how one of these discourse particles, well, when used as a front hedge shows that the speaker is unable to meet the requirements of the Maxim of Quantity in full [38].

In the hedging literature distinctions are not always made between the strategies to be applied and the modifying devices used to achieve the hedging [8]. The five distinct strategies defined by Clemen [8] are: *Politeness*, *Indirectness*, *Mitigation*, *Vagueness*, and *Understatement*. It can be argued that hedging is determined by a combination of factors, namely the type of context (discourse type), the colloquial situation, the speaker's/writer's intention and knowledge of the background dialogue/conversation [8].

C.1 Hedging and Culture

Culture is one factor that may play a role in the amount of hedges that are used between different groups of people. Crismore et al. [11] performed a Metadiscourse study in persuasive writing between American and Finnish University students. Analysis of the writing between the two different students found that the Finnish students hedged more than compared to their American counterparts. It was found that Finnish males had the highest percentage of metadiscourse that contained hedges (27.6%), than compared to Finnish females (22.5%), U.S. males (19%), and U.S. females (17%). A possible explanation for this difference was explained by Markkanen et. al [45] as cultural phenomenon due to past Finnish history, in which Finland had been ruled by other countries for several centuries. This combined with their position to a powerful neighbour has taught the Finns to be cautious about expressing their opinion. In other words the Finns have learned to use hedging as a protective device [11]. An alternative explanation is presented by Holmes [28] who suggested that hedging is used as a device for politeness and consideration for others, to give readers a chance to disagree [11].

C.2 Hedging - Effects on Attitudes and Learning

Research conducted Crismore et al. [10] focused on whether it is possible for hedges to move readers in changing their attitudes towards a particular author or subject matter of a given text. This was done through the selection of a 1,000 word passage from a science text book and a social studies textbook aimed for children in junior or high schools. The science text consisted of the work of Charles Darwin's Origin of the species, whist the social passage consisted of two different topics; "Change as progress" for the first half and women's liberation as the second half. Both passages were chosen due to the controversial nature of their content, thus naturally including hedging by the authors. The students were split into two separate groups. The first group, called the "control group" received a version of the two passages of text in which all hedges had been stripped away. The second group received hedged text that had hedges added back to the stripped version in which the hedging clauses were added back according to a predefined criteria by the authors.

The results for the science text showed that when the hedging criteria was applied it produced greater positive attitude change that the control group. Females were twice as more likely than the males to show a positive attitude change. However, with social studies text passage it was the control group that showed a higher positive attitude change than the hedged groups considered together, with there being no difference between male and female participants in positive attitude change. It was noted that some individual hedging criterion when used

together in the social studies text did bring about a greater positive attitude change than the control group.

C.3 Hedging in Physician-Physician Discourse

One of the interesting aspects of hedging is it's use within the medical domain where sometimes there is uncertainty of the consequences of particular actions. Work conducted by Prince, Frader, and Bosk [61] analysed the usage of hedging in the discourse between physicians. Prince et al.'s work focused particularly clarifying Lakoff's [34] original definition by making the observation that hedges makes things "fuzzy" in two distinct ways. The first way is by adding fuzziness *within the propositional content* proper, whilst the second way is how some hedges correlates with fuzziness *in the relationship between the propositional content and the speaker* [61]. This second type of hedging illustrates the speaker's commitment to the truth of the proposition that is being conveyed. These two distinct types of hedges were called by Prince et. al as *Approximators* and *Shields*.

Approximators were given a general definition by Prince et al. as a type of hedge that affected the truth condition of a given proposition. Thus, if an Approximator is added to a proposition $\mathbf{P}_i$ it results in the formation of a proposition $\mathbf{P}_j$, where $\mathbf{i} \neq \mathbf{j}$ [61]. However, a closer look by Prince et al. revealed that not all Approximators work the same way and that there are two different types of Approxmiators; *Adaptors* and *Rounders*. Adaptors occurred when a speaker tries to adapt around a given prototypicalness to express a vague approximation. For example, a person who has undetermined amount of crust in their ears would be described simply as "*just sort of* crusted". Rounders on the other hand appear quite frequently in the domain of measurements, where reference to a precise term is not relevant or immediately available, or, what one intends to convey is a range of items [61]. A example of Rounder is shown below [61]:

(10) "I didn't try a larger tube, because he came in with a two and half from the other place that was ***almost*** falling out, which is why we changed it right away."

Shields, unlike Approximators, do not affect the truth conditions of a given proposition. Prince et al. gave examples of prefixes for sentences such as *"I guess"* or *"It seems that"*, in which the truth conditions of the proposition remain unchanged [61]. It could be said that a simple definition for Shields is the expression of some doubt for a given proposition. However, analysis conducted by Prince et al has shown, that like Approximators, Shields also have two different types: *Plausibility Shields* and *Attribution Shields*. Plausibility Shields are a form of Shields in which a speaker indicates that s/he is less than fully committed to the proposition being conveyed [61]. In other words the speaker inserts a particular belief acquired via *plausible* reasoning [61]. An example of this type of Shield is illustrated below, in which the person answering the question tries to add a plausible reason for the given fact stated in the answer [61]:

(11) Q: "Has he had uh episodes of aspiration in the past?"
A: "He: never– *according to the mo– as far as I could tell from the mother*, never had *documented* aspiration."

Attribution Shields on the other hand differ from Plausibility Shields in that the speaker is speaking from knowledge or beliefs acquired via hearsay [61]. In other words, Attribution Shields attribute the truth of a given proposition to another. Sometimes this attribution is explicitly mentioned and sometimes it isn't. A example of this form of Shield is given below , in which the truth of proposition is attributed to another person[61]:

(12) "*According to Dr. Smith*, there was a dramatic response after medication."

Prince et al. in their study also focused on the distribution and occurrence of these different types of hedges within the corpus they had of physician-physician discourse. In particular the authors focused on the notion of

"self-repair" within their corpus and it's effect on the distribution and occurrence of the different hedge types. Self-repair was defined by the authors as to occur when a speaker breaks off before a completing a sentence, using characteristic repair intonation [61]. Their findings showed that within self-repairs, no form of the Rounder hedge type occurred. This finding was expected by the authors since Rounders do not involve any uncertainty or fuzziness. When comparing the amount of Shields and Approximators in self-repair and non self-repair texts, they found that Shields occur significantly more often in self-repairs than Approximators [61]. Overall, in self-repair and non self-repair corpus texts, the most prevalent type of hedge was Plausibility Shields (252), followed by Adaptors (158), Attribution Shields (54), and finally Rounders (24). No explanation was given by the authours for the high amount of Plausibility Shields in their findings. However the fact Attribution Shields occurred more significantly in self-repairs than any other type of hedge was explained by the authors "as a need on the co-participants to hedge some proposition/belief that they have uttered in a unhedged form [...] by explicitly attributing that belief to another" [61].

Overall this work provides a valuable insight into the possible types of hedges that are used in the medical domain and their frequency/distrubtion within a given a corpus. It remains to be seen if these findings by Prince et al. have any applicability to the type of discourse used between Consultants and Patients in a Neonatal Intensive Care Unit. However, whilst Prince et al. have provided through definitions for the different types of hedges, they have nevertheless based such definition on their own postulation. Thus, the lack of any empirically based method of testing the different hedging definitions does give some weakness to the work conducted by the authours.

C.4 Hedging Conclusion

A common theme in the review of Hedging literature is the lack of agreement and differing perspectives of what defines "Hedging". Whilst there has been an attempt by Clemen [8] to group together the various devices, strategies, and definitions of Hedging, there still remains the case that there is no formal agreement on this type of linguistic phenomenon. The lack of agreement of a definition and a taxonomy of hedge words/phrases means that any investigation that looks into the possibility of using Hedging as one particular device for Affective NLG would have to address these deficiencies.

While there are questions about the definition of Hedging, there is no question about its existence and it's effect. In particular the work by Crismore et al. shows how Hedging can influence and change people's attitudes [10]. In addition to this, the amount of Hedging used can depend on cultural and gender based factors as well [11]. These two particular pieces of work show this particular linguistic phenomenon at work. However, the work done by Prince et al. in their study of hedging in physician-physician discourse gives many ideas of the types of hedges that exist in this domain and their frequency/distrubution in their usage. Whilst this work may be of use, the lack of empirical testing for the differing hedging definitions does leave some room of doubt as to the usefulness on such definitions. In addition to this, it remains to be seen whether the findings by Prince et al. may or may not be applicable to the discourse between consultants and parents in an neonatal intensive care unit.

D Doctor-Patient Communication

There is considerable interest within the medical community to have a better understanding of the type of interactions that occur between doctors and their patients. The British Medical Journal noted in an editorial [47] that most complaints by patients and the public about doctors deal with problems of communication not with clinical competency. Patients want more and better information about their problem and the outcome, more openness about the side effects of treatment, relief of pain and emotional distress, and advice on what they can do for themselves [47]. The reason for such intensive focus on this aspect of medical care is evidence that several studies and reviews which have clearly shown a correlation between effective communication and improved health outcomes [47]. A comprehensive literature review on the subject matter was undertaken by Ong et al. [52]. In particular the authors noted that particular aspects of doctor-patient communication that seems to have an influence on the patients' behaviour and well-being. These aspects included the satisfaction of care, adherence to treatment, recall

and understanding of medical information. coping with disease, quality of life, and state of health.

Ong et al. noted that creating a good inter-personal relationship between the doctor and the patient was "a prerequisite for optimal medical care" [52]. Other researchers have noted that there are several basic 'core conditions' which have an effect on the on the relationship between a doctor and their patient. These 'core conditions' are: empathy, respect, genuineness, unconditional acceptance, and warmth [36, 19]. Empathy in particular was considered by Ong et al. to be one of the more important 'core conditions'. In particular, empathic doctor-patient relations consist of: eliciting feelings, paraphrasing and reflecting, using silence, giving encouragement, listing to what the patient is saying, but also to what the patient is unable to say and non-verbal behaviour [52].

The main purpose of any communication between a doctor and their patient is to exchange information. However, the purpose and reason for the exchange of information differs greatly between the two parties. For the doctor, the information is needed to establish the right diagnosis and treatment plan [52]. The patient on the other hand two particular needs from the doctor: "The need to feel known and understood" [52]. These two needs from the patient side tends to lead a desire to want as much information as possible [52]. However, doctors seem to underestimate this need for information [52]. In the case of Cancer patients it has been found that 92% desire all information about their disease, regardless of whether that information is good or bad to their prospects [52]. A recent study conducted by Hartmann et al. [26] found the usage of a web site enabled to patients to be inform of their condition and ask questions to their physicians had positively influenced the interactions between the two parties.

In medical communication an important distinction is made between instrumental and affective behaviour [52]. Instrumental behaviour is mainly cure oriented, whilst affective information is mainly care oriented in other words a set of behaviours that are "designed to establish and maintain a positive relationship between the doctor and his patient" [52]. Examples of Instrumental utterances made by doctors include: Giving information, asking questions, counselling , giving directions, identifying future treatment or tests, discussing side effects of tests or treatment, discussing test results with the patient, and explaining reasons for treatment or non-treatment [52]. Affective utterances including the following: Encouraging the patient, being very relaxed, extremely friendly, open and honest, showing concern, giving reassurance, showing approval, showing empathy, introducing self to the patient, addressing the patient by his/her first name, and encouraging small talk [52]. However out of all of these behaviours the most frequent was 'addressing the patient by name' was the most frequent and occurred in 71.8% of interactions. However, it is noted by Ong et al. that it is unclear if patients can discriminate between instrumental and affective communication [52].

Within Doctor-Patient communication literature, the amount of information given to patients by doctors remains unclear. Roter et al. [67] conducted a meta-anaylsis of physician communication and found that information giving was the most frequent form of communication making up 35.3% of total communication [52]. However, this is contrasted by work done by Waitzkin [73] who found that doctors only gave 1 minute of information in encounters lasting 20 minutes. It is noted by Ong et al. that Oncologists tend to withhold information from patients on the basis that total disclosure may cause strong negative emotional reactions on the side of the patients [52]. Different personal characteristics on the patient side may have any influence on the amount of the information they may receive from the doctor. Patients who asked more questions, expressed more concerns, were more anxious, received more information than patients who asked fewer questions, expressing less concern and were less anxious [52]. It was noticed by Waitzkin [73] that female patients tended to ask more questions and also received more from their doctors that compared to their male counterparts [52].

Doctors in essence are bilingual in terms of type of vocabulary they use, which can change from their everyday language (EL) to their medical language (ML) [52]. Patients typically are unfamiliar with ML and thus communicate with their doctors in their everyday language [52]. However, some patients may have some basic understanding for ML and may use it in lieu of equivalent EL terms for communicative effectiveness [52]. A similar concept was explored by Bromme et al. [5] on the use of language between fictitious patients and medical professionals on the subject on the diabetes questions and advice. It was found by Bromme et al. that, unlike face-to-face conversations where the standardisation on vocabulary takes place quite quickly, Internet communication can lead to misunderstandings on the part of the medical expert on the type of vocabulary to use when responding to a given question. In particular, it was found that if a medical professional was given a question that used the medical technical language vocabulary they would be to respond in kind with the same language with the presumption that the questioner understands the concepts involved [5]. Like wise, if a medical professional received a question that

contained medical everyday language vocabulary they wold be more likely to respond in kind [5]. An interesting finding by Bromme et al. was the fact that medical experts were prepared to give more explanations of important concepts, introduce more critical issues, and gave more behavioural tips in their replies to inquiries written in medical everyday language [5]. Bromme et al. did comment on the fact that "experts working in internet-based health counselling have to realise how dangerously easily it is to succumb to the temptation of using medical technical terms which have been used by patients, even if the have not understood their meaning" [5]. The understanding of common health terms by doctors, nurses and patients was examined by Hadlow and Pitts [23]. Their findings showed that the level of correct understanding was highest for physicians (70%) and the lowest for patients (36%). Similar high levels of misunderstanding of common medical language terms amongst patients was also found by Lerner et al. [37].

Ong et al. noted that patients are frequently dissatisfied with the information they receive, which has remained at a constant level for the last 25 years [52]. The median percentage of dissatisfied hospital patients is 38%, for general practice patients it is 26%, and for psychiatric patients it is 39% [39, 52]. Such dissatisfaction is due physician often underestimating the patient's desire for information [52].

D.1 Doctor-Patient Communication Conclusion

Research into Doctor-Patient communication is still an on going area of research within the medical domain. However, from the literature reviewed above there does seem to be several common themes that seem to emerge. Firstly, among patients there is a strong desire to be high informed about the medical that may concern them. Whilst doctors to underestimate this strong desire for information. In some cases there can be an argument made for withholding information from a patient if it may cause undue emotional stress and effect the patient's recovery chances [52]. However, while the literature has indicated a reason for reducing the amount of information that patient may receive, it isn't particular clear the manner and strategies in which doctors go about reducing the information for patients. The second theme is how empathy is one the more important core-condtions for a "healthy" interpersonal relationship between a doctor and a patient. An open question that remains to be answered is whether the usage of empathy in a NLG system for patients would be substantially as it is in a doctor-patient relationship? The final theme is the usage medical vocabulary between medical staff and patients. The high amount of misunderstanding of common medical terms for patients may mean a future NLG system would have to compensate for potential misunderstanding by rephrasing a medical term that may be misunderstood by the end user. However, whilst there is high level of misunderstanding of medical terminology among patients, there still nevertheless may be a small minority of patients who are comfortable with usage of such medical terms and an NLG system would have to adapt to such preferences.

E Neonatal Intensive Care Units and Parent Communication

In a special article on "Family-Centered" Neonatal care Helen Harrison noted that parents voiced their concern and anguish over their difficulty on obtaining accurate information about their babies' conditions [25]. This was one particular aspect led Harrison to define ten principles that emphasise the need for open and honest communication between parents and professionals on medical and ethical issues [25]. To this end the first principle makes it clear that information for parents must have the same facts and interpretation of those facts as possessed by the medical staff. In addition to this information must be complete, specific, detailed, and meaningful [25].

Even though "Family-Centered" care emphasises equality in terms of the information given to consultants and parents, work by Perlman et al. has shown that there is a difference in the type of information emphasised by the two parties [54]. Parents were more likely to emphasise information concerned the prognosis and medical treatments pertaining to their child. Whilst medical consultants on the other hand preferred to emphasise information that described the nature of the problem and its diagnosis. Perlman et al. conclude that even though parents want specific types of information, what is most clear is that they want to feel they have free access to all medical information [54].

The manner in which parents obtain information can also change during the course of care given to a premature child in NICU. Before delivery, parents were observed by Brazy et al. to be passive recipients of information, in which the learning is done on a one-to-one basis and initiated by others [4]. Whist parents find medical staff as the most accurate and knowledgeable source of information, they also frequently presented information for parents that was too complex [4]. The stress and strain of before and after giving birth also played a role in the amount of information parents could asborb. Brazy et al. found at times that parents would feel that they were overloaded with more information than they could assimilate [4, 68]. This aspect is particular important since one of the challenges faced by parents with babies in NICU is the knowledge gap, since all of the information that is been given is new to the parents [40]. Such a lack in knowledge has been cited by parents as a source of stress [2]. However, over time parents gradually developed more confidence in their knowledge base and a better capacity to evaluate information and resources. This gradual change suggests that parents become proficient in dealing with the information pertaining to their child's care and thus their information needs may change. This was alluded to by Brazy et al. who found that parents concern changed over time as it went from infant survival to infant care and personal coping [4]. A form of this personal coping was identified by Bass et al who found anticipatory coping as one of the important coping mechanism used by parents with babies in NICU [2]. Parents would try to gain a sense of power by frequently visiting the NICU and attempting to gain an intellectual understanding of their infants condition and treatment [2].

The dynamic nature of parental knowledge pursuit and stress alleviation was a particular aspect explored by Loo et al. In particular, Loo et al. looked at the relationship between knowledge of NICU monitoring by parents and it's affect on stress. Parents who were knowledgeable about such aspects of NICU monitoring attributed less stress alleviation than compared to parents who weren't familiar of such technical aspects [40]. Loo et al. postulates that the reason for such low levels of stress alleviation might be due to the fact that such understanding may not have a significant role in helping such parents to cope with the stress [40]. Additionally, the parent's perception of how severe the infant's illness can also be a powerful variable that determines the stress of the parents [72]. However, the parent's perception of the infant's morbidity is not just defined by the physiological condition of the child. A child with a debilitating condition, such as Down's syndrome, maybe physiological fine, but because of the presence of such a condition the parents could possibly be quite distressed [72]. Parents can also become distressed by seeing their child surrounded by evasive medical equipment [30], noticing colour changes such as jaundice, and also witnessing episodes of apnea or respiratory distress [49]. A similar finding was also found by Bass [2] who found that the infant's appearance was a major stress factor for parents as well.

E.1 Past Information Systems for NICU Parents

Studies have shown that parents want easily accessible information that is clear and concise. One avenue for parents to gain access to such information is through the Internet. Kowalski et al. found that 67% of parents in their study used the internet one hour per weeks for researching information about their baby on the internet [33]. However, Dhillon et al. found that only 10% of parents find such information on the Internet reliable [15]. e-Health systems on the other had have the ability to provide consistent and reliable information to that parents that is contextually relevant. Two particular examples of past e-Health systems that will be looked at are Baby CareLink and BabyLink.

E.1.1 Baby CareLink

Baby CareLink was the first successful e-health system to give Parents in the NICU domain. It gave parents of Very Low Birth Weight (VLBW) infants enhanced medical, informational, and emotional support during and after the infants time in the NICU [22]. Parents were given access to the e-health system through a secure website that acted as a portal which enabled parents to access six particular areas of clinical content [22]. Parents were able to obtain a daily clinical report that provided updates on the infant's clinical care and status, send messages to NICU staff, view daily digital images of their infant, access targeted informational resources, video conferences with medical staff, and finally access information dealing with infants care after NICU discharge [22]. Clinical reports and Informational resource tailoring for individual parents was done by the clinicians themselves rather than through

any automated process [22]. In the context of health education after NICU discharge Safran et al. postulated that the act of displaying or reading a single intensively personalised piece of information at the right time could have a much greater affect on health-related outcomes than compared to volumes of less specific web pages [69]. Most interestingly, was the fact that parents were given emotional support on the Baby CareLink system through a series of information resources which helped parents to understand the various issues confronting families of high-risk newborns. However, over a five month period Gray et al. found only a small percentage of parents actually visited such informational resources like the emotional support resource [22]. Such low usage of informational resources could be for various reasons. Gray et al. state that the total number of times a resource was used "should not be equated to it's potential impact even a single reading might have for the family [sic]."

Parents reaction to this e-health system was overwhelmingly positive and use of the system was enthustically embraced [22]. The use of such a system was found to empower parents to be more activate participants in their child's care [68, 69]. Thus, showing an overwhelming need for such a system to fulfil the communicative needs of parents.

E.1.2 BabyLink

The BabyLink project was co-developed by the Royal Infirmary of Edinburgh and software company Clevermed Ltd to give clinical reports for NICU parents. This was done by using an existing clinical information system to gather the information that had been entered for the babies to produce the automated reports. The clinical reports for parents, whilst automated, was not personalised in anyway and strictly focus on communicating the physiological and clinical aspects of the child's care. Hyperlinks were added to the clinical report which linked to a medical glossary of terms for words that a parent may not be familiar with. Doctors were also able to make manual modifications to the clinical reports if they wish to document their discussions with parents. In addition to the clinical report parents were also able to access an informal "Baby Diary", which was written manually by nurses. These reports were more informal than the clinical reports and contain a written story narrative from the child's view of the medical and non-medical events that have occurred during the day. Baby Diaries also contained digital images of the baby as well. Like Baby CareLink, parents using the BabyLink system also gained access through a secure website, which in addition to the clinical report and baby dairies, also contained a set of informational resources for parents [18].

The BabyLink system showed that there is a need for providing parents with not only technical physiological information about the condition of their child, but also information like the "Baby Diaries" that cater to the emotional needs of parents as well.

E.2 NICU-Parent Communication Conclusion

One of the clearest things that comes from NICU literature is the diverse informational and emotional needs parents have during the time their child is NICU.The relevance of information for parents is an important factor for helping parents to cope with stress and the importance of particular types of information may change during the course of a child's time in NICU. However, one of the major factors for the level distress that parents feel is the physiological condition of their child. Other pertinent factors can also have an emotional impact with parents and the importance of these factors may also change as well during neonatal care. Work by Bass [2] has shown that parents have indicated that how things that were said to them was as important as what was said. In addition to this, parents also have emotional needs to be taken into consideration when communicating information in such a sensitive domain. Parents want to receive positive information whenever possible and for any negative possibilities to be presented in a more positive way [2]. However, past information systems such as Baby CareLink and the BabyLink systems have only attempted to address the informational aspects of parents needs with and have not attempted to adapt information according to the parents emotional needs. ANLG systems however presents an opportunity to tailor information to the parent's differing emotional and informational needs.

List of Figures

List of Tables

References

[1] P. Baines and A. Haslam. *Type & Typography*. Laurence King Publishing, second edition, 2005.

[2] L. S. Bass. What do parents need when their infant is a patient in the NICU? *Neonatal Network*, 10(4):25–33, December 1991.

[3] A. Belz. And now with feeling: Developments in emotional language generation. Technical report, NECA, 2003.

[4] J. E. Brazy, B. M. H. Anderson, P. T. Becker, and M. Becker. How parents of premature infants gather information and obtain support. *Neonatal Network*, 20(2):41–48, March 2001.

[5] R. Bromme, R. Jucks, and T. Wagner. How to refer to 'diabetes'? language in online health advice. *Applied Cognitive Pyschology*, (19):569–586, 2005.

[6] P. Brown and S. C. Levinson. *Politeness: Some Universals in Language Useage*. Cambridge University Press, 1987.

[7] M. Charchuk and C. Simpson. Hope, Disclosure, and Control in the Neonatal Intensive Care Unit. *Health Communication*, 17(2):191–203, 2005.

[8] G. Clemen. chapter The Concept of Hedging: Orgins, Approaches and Definitions, pages 235–248.

[9] C. Conati. Probablistic Assemenet of User's Emotions in Educational Games. *Applied Artificial Intelligence*, 16(7):555–575, 2002.

[10] A. Crismore and W. J. V. Kopple. chapter Hedges and Readers: Effects on Attiudes and Learning, pages 83–114.

[11] A. Crismore, R. Markkanen, and M. S. Steffenson. Metadiscourse in persuasive writing: A study of texts written by american and finnish unversity students. *Written Communication*, 10(1):39–71, January 1993.

[12] F. de Rosis and F. Grasso. Affective natural language generation. *Affective Interactions, Springer Lecture Notes in AI*, pages 204 – 218, 2000.

[13] F. de Rosis, F. Grasso, and D. C. Berry. Refining instructional text generation after evaluation. *Artificial Intelligence in Medicine*, 17(1):1–36, 1999.

[14] R. L. DeMier, M. T. Hynan, R. F. Hatfield, M. W. Varner, H. B. Harris, and R. L. Manniello. A Measurement Model of Perinatal stressors: Identifying Risk for Postnatal Emotional Distress in Mothers of High-Risk Infants. *Journal of Clinical Psychology*, 56(1):89–100, 2000.

[15] A. S. Dhillon, S. G. Albersheim, S. Alsaad, N. S. Pargrass, and J. A. Zupancic. Internet Use and Perceptions of Information Reliablity by Parents in a Neonatal Intensive Care Unit. *Journal of Perinatology*, 23:420–424, 2003.

[16] M. Fleischman and E. Hovy. Towards emotional variation in speech-based natural language generation. In *Proceedings of the Second International Natural Language Generation Conference (INLG02)*, pages 57–64, 2002.

[17] L. S. Frank, S. Cox, A. Allen, and I. Winter. Measuring neonatal intensive care unit-related parental stress. *Journal of Advanced Nursing*, 49(6):608–615, 2005.

[18] Y. Freer, A. Lyon, B. Stevenson, and C. Russell. Babylink - improving communication among clinicians and with the parents of babies in intensive care. *British Journal of Healthcare Computing and Information Management*, 22(2):34–36, March 2005.

[19] S. L. Garfield. *Psyschotherapy - An Eclectic Approach*. John Wiley & Sons, New York, 1980.

[20] B. G. Glaser and A. L. Strauss. *The Discovery of Grounded Theory*. Weidenfeld and Nicolson, 1968.

[21] D. Goleman. *Emotional Intelligence*. Bantam, 1995.

[22] J. E. Gray, C. Safran, R. B. Davis, G. Pompilio-Weitzner, J. E. Stewart, L. Zaccagnini, and D. Pursley. Baby care link: Using the internet and telemedicine to improve care for high-risk infants. *Pediatrics*, 106(1318-1324), 2000.

[23] J. Hadlow and M. Pitts. The understanding of common health terms by docotors, nurses, and patients. *Social Science and Medicine*, (32):193–196, 1991.

[24] I. J. Haimowitz. Modelling all dialouge system participants to generate empathetic responces. *Computer Methods and Programs in Biomedicine*, 35(4):321–330, August 1991.

[25] H. Harrison. The principles for family-centered neonatal care. *Pediatrics*, 82(5), November 1993.

[26] C. W. Hartmann, C. N. Sciamanna, D. C. Blanch, S. Mui, H. Lawless, M. Manocchia, R. K. Rosen, and A. Pietropaoli. A website to improve asthma care by suggesting patient questions for physicians: Qualitative analysis of user experinces. *Journal of Medical Internet Research*, 9(1), January-March 2007.

[27] W. Hoeppner. Review of "Generating natural language under pragmantic constraints" by Edward H. Hovy. *Computational Linguistics*, 16(3):186–189, 1990.

[28] J. Holmes. Expressing dobut and certainty in english. *Regional English Language Center (RELC)*, 13(2):9, 1982.

[29] E. Hovy. *Generating Natural Language Under Pragmatic Constraints*. Lawrence Erlbaum Associates, 1988.

[30] K. Jämsä and T. Jämsä. Technology in neonatal intensive care — a study on parents' experiences. *Technology and Health Care*, 6(4):225–230, March 1998.

[31] S. Klein, J. Aeschlimann, D. Balsiger, S. Converse, C. Court, M. Foster, R. Lao, J. Oakely, and J. Smith. Automatic novel writing. Technical Report UWCS Tech Report 186, Department of Computer Science, University of Wisconsin Madison, 1973.

[32] T. H. H. G. Koh, P. N. Butow, M. Coory, D. Budge, L.-A. Collie, J. Whitehall, and M. H. Tattersall. Provision of taped conversations with neonatologists to mothers of babies in intensive care: randomised controlled trial. *British Medical Journal (BMJ)*, (334):28–31, January 2007.

[33] W. Kowalski, K. Leef, A. Mackley, M. Spear, and D. Paul. Communicating with parents of premature infants: who is the informant? *Journal of Perinatology*, 26:44–48, 2006.

[34] G. Lakoff. Hedges: A study in the meaning criteria and the logic of fuzzy concepts. *Journal of Philosophical Logic 2*, 1973.
[35] R. Lakoff. *Issues in Linguistics: Papers in Honour of Henery & Renée Kahane*, chapter Questionable answers and answerable questions, pages 453–67. University Ilinois Press, 1973.
[36] R. S. Lazarus. *Personality*. Prentice-Hall, Englewood Cliffs, NJ, 1979.
[37] E. B. Lerner, D. V. Jehle, D. M. Janicke, and R. M. Moscati. Medical communication: Do our patients understand? *American Journal of Emergency Medicine*, 18(7):764–766, 2000.
[38] S. C. Levinson. *Pragmatics*. Cambridge University Press, 1983.
[39] P. Ley. *Communicating with Patients. Improving Communication, Satisification and Compliance*. Chapman and Hall, London, 1988.
[40] K. K. Loo, M. Espinosa, R. Tyler, and J. Howard. Using Knowledge to Cope with Stress in the NICU: How Parents Integrate Learning to Read the Physiologic and Behavioral Cues of the Infant. *Neonatal Network*, 22(1):31–37, January-February 2003.
[41] C. Ma, H. Prendinger, and M. Ishizuka. Emotion Estimation and Reasoning Based on Affective Textual Interaction. In J. Tao, T. Tan, and R. W. Picard, editors, *Affective Computing and Intelligent Interaction*, pages 622–628. Springer, October 2005.
[42] S. Mahamood, E. Reiter, and C. Mellish. A comparsion of hedged and non-hedged nlg texts. In *Proceedings of the 11th European Workshop on Natural Language Generation (ENLG'07)*, 2007.
[43] S. Mahamood, E. Reiter, and C. Mellish. Neonatal Intesive Care Information for Parents – An Affective Approach. In *Proceedings of 21st IEEE International Symposium on Computer-Based Medical Systems*, 2008.
[44] R. Markkanen and H. Schröder. chapter Hedging: A Challenge for Pragmatics and Discourse Analysis, pages 3–18.
[45] R. Markkanen and H. Schröder. Hedging and its linguistic realizations in german, english, and finnish philosophical texts: A case study. In *Proceedings of the VAKKI-Seminar VII, Erikoiskielet ja käännösteoria*, pages 47–57. Vaasa, Finland: University of Vaasa, 1987.
[46] J. M. McGrath. Building Relationships with Families in the NICU: Exploring the Guarded Alliance. *Journal of Perinatal & Neonatal Nursing*, 15(3):74–83, 2001.
[47] S. Meryn. Improving doctor-patient communication. *British Medical Journal (BMJ)*, (316):1922–1930, June 1998.
[48] M. S. Miles, S. G. Funk, and M. A. Kasper. The Neonatal Intensive Care Unit Enviroment: Sources of Stress for Parents. *AACN Clinical Issues in Critical Care Nursing*, 2(2):346–354, 1991.
[49] M. S. Miles and D. Holditch-Davis. Parenting the Prematurely Born Child: Pathways of Influence. *Seminars in Perinatology*, 21(3):254–266, June 1997.
[50] M. Ochs, R. Niewiadomski, C. Pelachaud, and D. Sadek. Intelligent expressions of emotion. In J. Tao, T. Tan, and R. W. Picard, editors, *Affective Computing and Intelligent Interaction*, pages 707–714. Springer, 2005.
[51] J. Ogden. *Health Psychology - A Textbook*. Open University Press, fourth edition, 2007.
[52] L. M. L. Ong, J. C. J. M. de Haes, A. M. Hoos, and F. B. Lammes. Doctor-patient communication: A review of the literature. *Social Science and Medicine*, 40(7):903–918, 1995.
[53] A. Pavia. *Affective Interactions – Towards a New Generation of Computing Interfaces*, chapter Affective Interactions – Toward a New Generation of Computer Interfaces?, pages 1–8. Springer, 2000.
[54] N. B. Perlman, J. L. Freedman, R. Abramovitch, H. Whyte, H. Kirpalani, and M. Perlman. Informational needs of parents of sick neonates. *Pediatrics*, 88:512–518, 1991.
[55] R. Picard. *Does HAL Cry Digital Tears?: Emotions and Computers*, chapter 13, pages 279–303. MIT Press, Cambridge (Mass.), 1996.
[56] R. Picard. *Affective Computing*. MIT Press, Cambridge (Mass.), 1997.
[57] R. W. Picard. *Affective Interactions – Towards a New Generation of Computing Interfaces*, chapter An Interview with Rosalind Picard, Author of "Affective Computing", pages 219–227. Springer, 2000.
[58] F. Portet, A. Gatt, J. Hunter, E. Reiter, S. Sripada, and F. Gao. A Core Architecture to Summarise ICU Data as Tailored Text. In *21st International Congress of the European Federation for Medical Informatics (MIE 2008)*, Göteborg, Sweeden, May 2008.
[59] F. Portet, E. Reiter, J. Hunter, and S. Sripada. Automatic generation of textual summaries from neonatal intensive care data. In *Proceedings of the 11th Conference on Artificial Intelligence in Medicine (AIME '07)*. LNCS, July 2007.
[60] K. Prechet. *Patterns of Stance in English*. PhD thesis, Northern Arizona Univeristy, 2000.
[61] E. F. Prince, J. Frader, and C. Bosk. *Linguistics and the Proffessions*, chapter On Hedging in Physician-Physician Discourse, pages 83–97. Ablex, 1982.
[62] M. Redshaw and A. Harris. Maternal perceptions of neonatal care. *Acta Paediatrica*, 84(6):593–598, 1995.
[63] T. Reid, R. Bramwell, N. Booth, and A. M. Weindling. A new stressor scale for parents experiencing neonatal intensive care: the NUPS (Neonatal Unit Parental Stress) scale. *Journal of Reproductive and Infant Psychology*, 25(1):66–82, February 2007.
[64] E. Reiter. An architecture for data-to-text systems. In *Proceedings of the 11th European Workshop on Natural Language Generation (ENLG'07)*, June 2007.
[65] E. Reiter and R. Dale. *Building Natural Language Generation Systems*. Studies in Natural Language Processing. Cambridge University Press, 2000.
[66] E. Reiter, S. Sripada, and R. Robertson. Acquiring Correct Knowledge for Natural Language Generation. *Journal of Artificial Intelligence Research*, 18:491–516, 2003.

[67] D. L. Roter, J. A. Hall, and N. R. Katz. Patient-physican comunication: A descriptive summary of the literature. *Patient Education and Counseling*, 12(2):99–119, 1988.

[68] C. Safran. The collaborative edge: Patient empowerment for vulnerable populations. *International Journal of Medical Informatics*, 69:185–190, 2003.

[69] C. Safran, G. Pompilio-Weitzner, K. D. Emery, and L. Hampers. A medicaid ehealth program: An anaylsis of benefits to users and nonusers. In *AMIA Annual Symposium Proceedings*, pages 659–663, 2005.

[70] J. A. Sánchez, I. Kirschning, J. C. Palacio, and Y. Ostróvskaya. Towards mood-oriented interfaces for synchronous interaction. In *CLIHC '05: Proceedings of the 2005 Latin American conference on Human-computer interaction*, ACM International Conference Proceeding Series, pages 1–7, New York, NY, USA, 2005. ACM.

[71] R. Y. Seideman, M. A. Watson, K. E. Corff, P. Odle, J. Haase, and J. L. Bowerman. Parent Stress and Coping in NICU and PICU. *Journal of Pediatric Nursing*, 12(3):169–177, June 1997.

[72] D. Shields-Poë and J. Pinelli. Variables Associated with Parental Stress in Neonatal Intensive Care Units. *Neonatal Network*, 16(1):29–37, February 1997.

[73] H. Waitzkin. Doctor-patient communication. clinical implications of social scientific research. *Journal of the Americal Medical Association*, 252(17), November 1984.

[74] T. Wehrle and S. Kaiser. *Affective Interactions – Towards a New Generation of Computing Interfaces*, chapter Emotion and Facial Expression, pages 49–63. Springer, 2000.