Does speech prosody matter in health communication? Evidence from native and non-native English speaking medical students in a simulated clinical interaction

Fan Wang

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

School of Rehabilitation Sciences

University of East Anglia

January 2014

Abstract

The impact of the UK's multilingual and multicultural society today can be seen in its healthcare services and have contributed towards shaping communication skills training as a core part of the UK undergraduate medical curriculum. NHS complaints statistics involving perceived staff attitudes have remained high, despite extensive communication skills training. Furthermore, foreign doctors have received a higher proportion of complaints than UK doctors. Finally, how linguistic and social factors shape the conveyance and perception of attitudes related to professionalism in medical communication remains poorly understood.

The ultimate aim of this study was to ascertain if speech prosody contributes to the perception of professionalism in medical communication. Research questions on the role of speech prosody in conveying professional attitudes in medical communication, the prosodic differences between native and non-native English speaking medical students in a simulated clinical interaction, and the influence of prosodic features on listeners' perceptions of professional attitudes were addressed.

A set of acoustic parameters representing the speech prosody of native and non-native medical students in the simulated clinical setting was analysed. A perceptual experiment was then carried out to investigate the factors affecting perceived professionalism in extracts of the analysed simulated clinical interaction.

The examined acoustic parameters were found to be sensitive to the English language background and the task within the simulated consultation. Interestingly, the attitudinal information associated with some of these acoustic parameters were perceived by listeners and were reflected by higher professional scale scores in the perceptual experiment, even after adjusting for the English language background. The factors of training level and consultation task also emerged to be affecting professional scale scores.

Initial findings have confirmed that speech prosody plays a role in terms of contributing towards the perception of professionalism in medical communication. Incorporating *how* messages are delivered to patients into current models of communication skills training may have positive outcomes.

Acknowledgements

I am deeply indebted to my supervisors Dr Christine Raschka, Dr Zoe Butterfint and Prof Jacqueline Collier for all their efforts in helping me advance this research. They have all taught me to become a resilient researcher and a better writer. In particular, Christine has been a role model in attention to details; one cannot have better examples than those from Zoe of how to present ideas accurately, clearly and coherently; Jacqueline's optimism has been invaluable, especially when things looked like dead ends (and SPSS wanted to unfriend me). I have also benefited tremendously from Dr Allan Clark's support in statistics. Insight and encouragement from Prof Sam Leinster, my former supervisory team member, was an inspiration before his retirement and has remained so since then. Discussions with and constructive comments from Dr Jan McAllister and Dr Simon Horton have been extremely helpful. I have been fortunate enough to receive the same kind of help from my panel chairpersons Dr Jill Robinson and Prof Paul Hunter to whom I am grateful.

This research was generously funded by a UEA FMH (then FOH) PGR Studentship and a UEA International Studentship without which this work would not have been possible. There are various people and teams at UEA that I wish to thank sincerely – their help and support has been crucial to the completion of this research. These include the MED administration team, the Survey Office, the Communications Office, the Telecommunication (and the former TV) Studios, the FMH PGR Office, the lovely LCS team, and all the research participants.

I am lucky to have, in various parts of the world, friends who are like family to me and have supported me selflessly. Among them are Ailsa and Phil Wright, Ellie, Chryseis and Alexandros Tsatalas. Ellie has been a special friend who has never given up faith in what I do, even when I could not see any myself. Nada Ali has offered shelter, food and company whenever I am in need. The communities of UEA Chaplaincy and Holy Apostles Parish have given me much more than charity and brotherly love. Riccardo Monni sacrificed much of his festive celebration to proofread this work. Each and every one of my friends from and through the Chinese students community is wonderful in their own ways. Fellow PGR students at UEA, including those who finished before me and those whose journeys were cut short, are all legends! The list can go on and on until reading it becomes boring, and that is simply because words are not adequate to thank them all.

My diagnosis of, treatment for and recovery from clinical depression in the midst of this journey was a humbling experience but also a blessing in disguise. By understanding and experience first-hand the weakness and limitation of the mind, I have been made stronger and more eager to embrace life. For this I owe a lot to my counsellors, therapist and doctors (and for not being terrified upon hearing about my research).

I dedicate this work to my parents Qiuxia Ye and Zhongyuan Wang. Throughout their lives they have worked tirelessly and without complaint to ensure that I have a better one.

Deo gratias, for the gift of wisdom and for what we yet strive to understand.

Contents

List of tables

List of figures

Abbreviations

ANOVA	analysis of variance
CI	confidence interval
dB	decibel
DH	Department of Health
EEA	European Economic Area
ELT	English Language Teaching
F0	fundamental frequency
GLM	general linear model
GMC	General Medical Council
GP	general practice or general practitioner
HSCIC	Health and Social Care Information Centre
Hz	Hertz
ICC	intraclass correlation coefficient
IHC	Institute for Healthcare Communication
L1	first language
L2	second language
M	mean
MBBS	Bachelor of Medicine, Bachelor of Surgery
NHS	National Health Service
OSCE	Objective Structured Clinical Examination
PIS	patient information sheet
RCP	Royal College of Physicians
RP	Received Pronunciation
SD	standard deviation
SSBE	Southern Standard British English
UEA	University of East Anglia
UK	United Kingdom
VAS	visual analogue scale

Chapter 1 Introduction

1.1 Aims of this study

This exploratory study is intended as a first step to explore the role of speech prosody in health communication. More specifically, this study aims to examine:

(1) the importance of speech prosody in doctor-patient communication, especially in the building of the therapeutic relationship,

(2) potential differences between native and non-native English speakers in prosodic features during clinical encounters, and

(3) the influence of prosodic features on perceived professional attitudes in clinical settings.

Ultimately, the study aims to ascertain if speech prosody contributes towards professionalism in terms of medical communication, and if incorporating the active use of speech prosody in current medical communication skills training programmes can be beneficial.

The need for this research is raised in the contexts of the current climate of the healthcare service as well as past research findings in relevant fields. Such background information is given in Section 2 below. The above immediate research aims are based on previous evidence as well as current gaps in knowledge. These are summarised in Section 3 below.

1.2 Research backgrounds

Few today would dispute the importance of good communication in patient-centred medicine, which not only seeks to cure the disease but to care for the patient as a person (see e.g. Francis 2013, IHC 2011). However, good doctor-patient communication is not limited to the simple exchange of information. In a clinical setting, any communication that contributes to the condition’s diagnosis, treatment and management is of course important, but similarly important is the communicative behaviour that attends to the socio-emotional wellbeing of the patient. The former cure-oriented or instrumental communication belongs to the cognitive domain, and the latter care-oriented or affective communication, the emotional domain, both of which are integral to the role of the doctor (Ong et al. 1995).

A constructive and co-operative therapeutic relationship is a partnership between the doctor and the patient (RCP 2005). In other words, the patient does not play a passive role but perceives and responds to the affect (integral to the care provided) in the therapeutic relationship. Therefore the affective elements must be present in the doctor-patient communication. Affective doctor-patient communication does not only contain a verbal exchange on the feelings of the patient, but should include the display of professional attitudes appropriate to the context of the communication and to the role of the doctor. These attitudes can range from general interpersonal stances, such as friendliness, to the compassion expected of the medical profession and to clinical empathy. Therefore, if there is insufficient or inappropriate communication between doctors and patients, in either the cognitive or the affective domain, patient satisfaction with the service can be undermined.

In the period of 2008-2009, many of the complaints received by the NHS do not relate to any aspects of clinical treatment, but are about communication and attitude, amounting to approximately 23% and 24% for hospital and GP services respectively (HSCIC 2009). Such figures further add to the evidence that communication to patients and professional attitudes displayed by medical staff are linked to patient satisfaction; insufficiency or inappropriateness in these areas can lead to the dissatisfaction of patients (Ong et al. 1995, Jangland 2009).

Most of the emotional and attitudinal information in speech is not expressed in words but is communicated in speech prosody (melody and rhythm). Prosodic features, i.e. features of the speech signal related to pitch, volume and tempo, are responsible for conveying attitudes and emotions as well as forming part of the linguistic structure (Nooteboom 1997, Nolan 2006). For example, a rising-falling intonation (the pitch pattern of a sentence) in English can convey a sense of sarcasm, intended or not, to a native ear (Cruttenden 1997, 2008). Similarly, using a falling intonation rather than a rise when an interaction should be softened can give the impression of being unfeeling or harsh (Hewings et al. 1998). The saying “it’s not what you say; it’s how you say it” depicts how misunderstanding of attitudes and emotions can occur, even when the content of speech is perfectly clear. In everyday communicative contexts, such misunderstanding can cause embarrassment or hurt to feelings. However, in the clinical context where the patient is often already vulnerable, such misunderstanding can have even worse consequence: not only does the communication risks breakdown, the therapeutic relationship might also be damaged, leaving patient experience compromised (Ong et al. 1995, Silverman et al. 2005).

In addition, languages vary in the way prosodic features are used, particularly the use of pitch (the highness or lowness of a sound). For example, in some languages the use of quite a narrow pitch range is perfectly acceptable; however in English, a language that uses a wider pitch range, this can be interpreted as demonstrating a lack of interest (Mennen et al. 2007, Gussenhoven 2002, Wichmann 2002). Non-native speakers of English often have patterns of prosodic features based within their respective native languages and thus are different from native English prosodic norms (Collins & Mees 2003). The perception of such differences by native speakers of English or, indeed, by other non-native English speakers, is more likely to cause misunderstanding than when prosodic norms are shared.

As outlined above, there is a complex interplay between need, expectation, production and perception in the communicative exchange between speakers and listeners. This is normal in all communicative contexts and is usually quite successfully negotiated and, if needed, repaired. However, in communicative situations where there may be an inherent imbalance, such as in doctor-patient interactions, the 'more powerful' conversation partner will have to be more aware of the communication structure to avoid being too dominant, or at least perceived to be.

For this reason, communication skills are now firmly embedded in UK medical school curricula (Brown 2008, von Fragstein 2008). However, many of these mainly concentrate on the "what to say" (together with the "when" and "where") rather than the "how to say it". Although both attitudes and skills are acknowledged to be vital in doctor-patient communication, these training programmes often only introduce the importance of displaying certain attitudes in clinical encounters but do not go on to explore the 'how' any further (Kurtz et al. 2005, Silverman et al. 2005). It is not clear whether learning the skills of what to say automatically leads to the ability to know how to say it. Medical students need to develop the ability to convey professional attitudes appropriately in their interaction with patients, even before they begin postgraduate training as junior doctors. Excluding attitudinal topics in the training might not benefit the development of a complete set of communication skills for medical students. Furthermore, in some training programmes, including the widely used Calgary-Cambridge model (Kurtz 2002, Silverman et al. 2005), the scope of attitude can be limited to the cognitive domain, i.e. the ‘underlying intentions, values and beliefs’ (Kurtz et al. 2005:60). Attitude is, however, a rich construct in which cognition interacts dynamically with affect and general evaluations (Millar & Tessar 1986), and a cognitive or affective focus can shape a different overall attitude (van den Berg et al. 2006). Not extending the discussion of attitude to its affective and behavioural aspects can be disadvantageous, because the cognitive intentions, values and beliefs underlying good professional practice as understood by medical students might potentially fail to be transferred to behaviour, especially to affective communication. Yet doctors’ attitudes are most likely to be perceived by patients through communication behaviour. Therefore even if a doctor possesses all professional attitudes in the cognitive domain but fails to demonstrate them in communication, patients might not be able to perceive them and might make unfavourable attitudinal judgements on the doctor based on the communication behaviour.

In summary, good doctor-patient communication is essential in patient-centred care in which professional attitudes must be displayed. Attitudinal information can be conveyed via multiple channels, for instance through non-verbal features in communication, but most importantly through prosodic features such as pitch, volume and rate in speech that modifies or enhances the speech content. When prosodic features are not used appropriately, misunderstanding, especially that pertaining to the speaker’s attitudes and emotions can occur. If this happens during a clinical encounter, it can result in a compromised patient experience. Prosodic features of English can also pose a challenge for non-native speakers who tend to differ from native speakers in these features. However, skills in using prosodic features to convey professional attitudes appropriate to clinical encounters, though sometimes briefly mentioned, are largely absent from medical communication training programmes. It is not known whether medical students can develop such skills on their own. It is also not known whether the absence of targeted training will further disadvantage non-native speakers in health communication who might find it difficult to master the subtlety of English prosodic features in the first place. Section 1.3 below introduces the body of previous research on the areas relevant to the current study.

1.3 Evidence and unanswered questions

Several topics from multiple research disciplines are of interest in this study. These range from medical communication to speech prosody and to attitudinal perception. Consequently relevant literature can be found in health communication research, linguistics and psychology. This section summarises the existing knowledge on these topics and identifies the gaps. A more detailed review of the relevant literature is presented in Chapter 2.

There is a wealth of research showing the benefits of good doctor-patient communication to quality health care. These include patient satisfaction (Wong & Lee 2003), patient adherence (Baumann et al. 2008) and positive patient health outcomes (Stewart 1995). Communication skills training for healthcare professionals has been found to be effective (Gysels et al. 2004, Parry 2008), especially in terms of increased patient satisfaction (Trumble et al. 2006). Communication skills training not only forms a core part of UK medical education (Brown 2008), it is also a requirement for medical graduates stipulated by the GMC (2009). The training outcomes should include the ability to communicate clearly, sensitively and effectively as well as to demonstrate politeness, considerateness, trustworthiness and honesty appropriately in the relationship with patients (GMC 2009). However, it is not fully established whether currently adopted training programmes in UK medical schools can directly lead to the fulfilment of these communication and attitudinal outcomes. Furthermore, relatively few studies in the literature focused on communication features that contribute to the consolidation of the therapeutic relationship, especially in terms of the display of professionalism in communication. Some studies, though, considered the link between the surgeons' tone of voice and their history of malpractice claims (Ambady et al. 2002) or the communication styles of strong and weak candidates in an Objective Structured Clinical Examination (OSCE) (Roberts et al. 2003), for example. None of these studies, however, included prosodic features that could be responsible for the tone of voice or the attitudinal meanings conveyed. Therefore there is a gap in research evidence on the role of speech prosody in communicating attitudes in clinical settings.

In both linguistic and psychological research, however, speech prosody is a well-studied area. There is now accumulated evidence on how prosodic features are used by the speaker to encode affective information and how the listeners decode such information using prosodic features as perceptual cues. A synthesis of such evidence is given in Chapter 2. Previous research also conducted cross-linguistic comparison of the paralinguistic usage of prosodic features and found differences in the perception of affective meanings associated with speech prosody between listeners from different languages (e.g. Chen et al. 2004). In addition, native and non-native speakers of English were reported to differ in their use of prosody to convey affective meanings (e.g. Verdugo 2005), so misunderstandings of attitudes and emotions due to prosodic features used are possible in communication between native and non-native speakers. Hence when studying the use of speech prosody for affective communication in a multilingual/multicultural society such as the modern UK, evidence from non-native speakers must be considered.

Some types of affective information conveyed through speech prosody were studied more often than others. To date, the most comprehensive evidence (see Scherer et al. 2003) lies in how basic emotions (e.g. anger and fear) are expressed in prosodic features. How attitudes, especially those relevant to the contexts of professional communicative, are expressed in prosodic features has been less frequently studied. Linguistic and psychological research also tends to rely on elicited or acted speech when studying affective meanings in prosody (e.g. Banse & Scherer 1996, Gobl & N í Chasaide 2003). Findings using such speech data obtained outside real-life communication are not necessarily generalisable to clinical settings where attitudinal and emotional stakes can be higher than in everyday situations. In order to understand the effects of healthcare professionals' speech prosody on patients' perceptions of professional attitudes in the interaction, investigation must be made on speech materials that reflect as accurate a representation of clinical communicative events as possible in an exploratory study of this nature. While the speech samples and their manipulation are too limiting for the purposes of the current study, the methods used in linguistic and psychological studies of speaker affects in speech prosody may be usefully extended to the examination of clinically relevant speech samples, which is further discussed in Chapter 3.

Although a great many studies were carried out in the past on non-native English prosody, their focus was often on the errors[1] present in prosodic features (e.g. incorrect word-stress placement, Hahn 2004), the resulted impression of foreign accent or the affected intelligibility. The differences between native and non-native speakers in the paralinguistic usage of English prosody were less often studied. However, as pointed out earlier, the different attitudinal or emotional meanings inferred as a result of the prosodic features used can also be a source of misunderstanding and therefore need to be examined. Furthermore, prosodic features of non-native healthcare professionals in their communication with native patients, especially in terms of conveying professionalism, were rarely, if ever, studied in previous research. Previous research on the differences between native and non-native doctors in other aspects of speech (e.g. syntactic and lexical features, Van del Poel & Brunfaut, 2010) yielded interesting results. Similarly, the study of the communication behaviour of non-native doctors should be extended to prosodic features. Since inappropriately used prosodic features can cause misunderstandings of attitudes, studies on them with evidence from both native and non-native doctors seem particularly urgently needed. Currently, 10.4% and 26.1% of the doctors practising in the UK are from other countries in the European Economic Area (EEA) and non-EEA countries respectively (www.gmc-uk.org). Many of these foreign doctors are from countries where English is not spoken as a native language. Foreign doctors are also known to be over-represented in disciplinary actions (Allen 2009) and the outcomes for them are likely to be more serious compared with native doctors (Humphrey et al. 2011) and therefore the need to study the communication behaviour of foreign doctors is particularly urgent. Recently there has been

[1] Some consider using terms such as ‘errors’ for features of World Englishes (those types of the English language that are different from those varieties spoken in Western English speaking countries) in applied linguistic research a neo-colonial view. See Kachru (1992) for a discussion on this.

increasing discussion on restoring a caring, compassionate and committed culture in the NHS since the publication of the Francis (2013) Report. This, combined with huge cultural changes in society at large, further suggest that it is necessary to study how messages from healthcare professionals, both native and non-native, are conveyed to and perceived by patients during clinical encounters.

1.4 Adopting established research methods

This study is embedded in the fields of health communication, speech prosody and attitudinal judgment, and draws on the respective methods used in these research fields. Thus, this section briefly locates the current exploratory study in the existing methodologies of its related fields.

A tradition has been established in modern linguistic research to collect empirical data from natural speech for hypothesis testing. The data collection techniques often involve eliciting speech using everyday tasks such as story-telling (e.g. Verdugo 2005) as well as the use of high quality audio recording for acoustic-phonetic analysis (examining the acoustic properties of the speech sounds of interest). Such a data collection method is adopted in the current study, albeit with some adaptations. Firstly, this study intends to observe native and non-native English speaking medical students performing a set of tasks with a same simulated patient. On the one hand, like previous phonetic/linguistic research, this study records speech using high quality audio equipment in a sound-treated environment, and the recording is ready for acoustic-phonetic analysis with the aid of dedicated computer software. On the other hand, unlike speech in everyday communicative contexts in previous linguistic research, this study uses speech in a simulated professional interaction. Such interaction not only executes clinical purposes such as taking the (simulated) patient’s history, but also has attitudinal intentions to display elements of professionalism in communication. The speech data thus collected realistically represent a clinical interaction setting without affecting real patients’ experience in any way.

A wealth of previous linguistic and psychological studies examined listeners’ perceptions of speaker affect. These studies often manipulated the speech data by means of masking (Scherer et al. 1984, Ladd et al. 1985) or re-synthesis (Gobl & Ní Chasaide 2003, Chen et al. 2004). These means of manipulation act as an effective control in the studies, although the naturalness of speech can be affected. The attitudinal perception under investigation in the current exploratory study needs to represent the experience of the potential service user, therefore any manipulation technique that makes the speech less realistic is not appropriate for the needs of this study. In particular, this study explores untrained listeners' attitudinal perception from the recorded and analysed audio clips as stimuli. The listeners in this study are native speakers of English, while the audio clips are of both native and non-native speakers of English. This enables the comparison between the two groups of speakers delivering medical communication in terms of a battery of aspects of attitudes relevant to the clinical context. These attitudinal aspects can be measured on a scale containing appropriate items as in related previous studies (Ladd et al. 1985, Gobl & Ní Chasaide 2003, Chen et al. 2004). Results of the comparison should consequently help to ascertain any effect of speakers' English language background on professionalism as perceived by listeners. Furthermore, the attitude-measuring scale can be used in future health-related, as well as psychological, studies as a basis for developing a refined instrument.

This study also plans to adopt the established method in psychological research of carrying out multiple regression analysis estimating listeners' attitudinal perception in the clinical context with acoustic measures of speech prosody as predictors. This method enables the study to answer the overarching research question as to whether speech prosody matters in health communication. In addition, results of the analysis will also contribute to the further understanding of the vocal expression of speaker affects in psychology.

Each of the fields above can offer considerable methodological strengths, as well as some limitations. This study attempts to adopt their strengths and aims to avoid some of their limitations. In this way, it is anticipated that findings generated from this study will not only support the research hypotheses, but may also add to wider methodological considerations in the respective fields.

1.5 Summary and thesis structure

Communication is at the heart of patient-centred care, and good communication not only facilitates the effective exchange of information in clinical encounters, but also builds and maintains the therapeutic relationship between doctors and patients. Professional attitudes must also be demonstrated in doctor-patient communication. Lack of demonstrated professional attitudes can compromise the patient experience of the communication and indeed the entire healthcare experience. However, much of the attitudinal information present in speech is not conveyed verbally, but in the features of, for example, pitch, volume and tempo, all known as prosodic features. Inappropriate use of prosodic features can result in misunderstanding of speaker attitudes and can, in clinical communication, lead to the lack of professional attitudes or even presence of unprofessional attitudes perceived by patients. Differences in prosodic features between native and non-native speakers of English have been widely reported, but it is not known whether such differences are also observable among non-native speaking healthcare professionals. Nor is it known whether such differences affect professionalism in clinical interactions perceived by patients. Thus, this study aims to address these questions by analysing empirical evidence, using adapted existing research methods.

Ultimately, this study wishes to make a threefold contribution to research, education and practice. Findings from this study aim to inform, based on clinically relevant evidence, medical students, professionals and educators of the importance of speech prosody as an effective communication channel in their interaction with patients. It also endeavours to highlight the potential communication training needs of native and non-native speaking healthcare professionals and students.

The chapters in the thesis are arranged in the following order. After this introduction, Chapter 2 gives a comprehensive review of relevant literature on issues in health communication and speech prosody. Research questions are formulated following the literature review. Chapter 3 reports the methodological considerations for this study, which is essentially interdisciplinary in nature. This study is divided into two phases which are reported separately. Chapter 4 reports the methods and results of an acoustic investigation of clinically relevant speech, and Chapter 5 of a perceptual experiment on attitudes in speech related to professionalism during clinical encounters. The findings of these two phases are then discussed more holistically in Chapter 6, before the concluding remarks, implications, and limitations, are given in Chapter 7.

Chapter 2 Literature Review

Several issues in patient-centred health care are relevant to this study, especially those related to professionalism and communication. Health communication studies and training manuals have almost unanimously advocated the significance of non-verbal communication. However, the channel in speech communication most responsible for the expression of attitudes (including those related to professionalism), speech prosody (melody and rhythm), has largely been ignored or received only passing mentioning in these texts. Prosodic features related to pitch, volume and rate of speech have been extensively studied with regards both to their functions in regulating linguistic structures and conveying attitudinal and emotional information, although such research has often been limited to the disciplines of linguistics and psychology. The omission of evidence on the attitudinal and emotional functions of prosodic features in health communication can be disadvantageous (see O'Toole 2008: 158-162) and a review of such evidence is needed in the context of healthcare settings. The current chapter synthesises evidence from all relevant areas before the research questions of this study are formulated in context.

2.1 Healthcare expectations

2.1.1 Basic rights of service users

The basic rights of patients and the fundamental duties of doctors have their origins in the principles of Hippocrates and the oath under which doctors are obliged to prevent harm to patients and to promote their health (Chin 2002). Safe, quality care with respect, dignity and compassion is incorporated in the values of the NHS in England and safeguarded in the legal rights of the patients (DH 2013a).

Quality care is not limited to a high standard clinical treatment. NHS patients in England are entitled to professional clinical care delivered by qualified staff, with approved medications and programmes, within maximum waiting times (DH 2013a). Furthermore, patients should be treated by their healthcare professionals in a manner that is dignified and respectful (DH 2013a: Section 3a). These basic rights afforded to patients are established by the European Convention on Human Rights and the Human Rights Act 1998 (DH 2013b:50). Any violation of the basic rights of patients to quality clinical care with dignity and respect is therefore unlawful in the UK.

2.1.2 Requirements for communication and attitude

The UK medical profession and its education are further regulated by the statutory authority, the General Medical Council (GMC). The main functions of the GMC include registering doctors, training medical students and setting standards for the medical practice in the UK. Consequently all UK medical students and registered practitioners are required to meet the outcomes of medical education and the standards of medical practice required by the GMC.

The GMC has set not only the standards for the clinical aspects of care but also requirements for both qualified doctors as well as medical students regarding communication with patients. These requirements were established in the capacity of the GMC to protect the public by determining 'the principles and values that underpin good medical practice' (www.gmc-uk.org). Such principles and values are included in *Good Medical Practice* (GMC 2006) in which doctors are required to base their therapeutic relationship with patients on openness, trust and good communication. In particular, doctors are asked to be polite, considerate and honest in their partnership with patients in which attitudes such as being respectful, encouraging and sensitive are mandatory (GMC 2006). Furthermore, these requirements are not for qualified doctors alone. Training outcomes set for medical students in this area are similar to the principles for doctors in practice. In *Tomorrow's Doctors* (GMC 2009), requirements for professional attitudes and communication are made for the medical graduate as a professional and as a practitioner. With regards to becoming professionals, medical students are asked to observe, among others, the following principles:

> 'Be polite, considerate, trustworthy and honest, act with integrity, maintain confidentiality, respect patients' dignity and privacy.' (GMC 2009: 25)

This requirement clearly outlines the target of attitudinal development for a medical graduate which is in the best of patients' interests. In addition, in their professional communication with patients as practitioners, medical graduates are required to be clear, sensitive and effective (GMC: 2009). This requirement also suggests that while the content of communication with patients is important (it must be clear), it needs to be delivered with appropriate (sensitive) attitudes. In short, the GMC guidelines on doctor-patient communication, together with those on clinical standards, are there to protect the rights of patients and to assure the quality of care. Doctors, both those in practice and those in training, are expected to understand the importance of, and to learn the skills of effective communication with patients.

The above GMC guidelines on maintaining the partnership and effectively communicating with patients include several distinctive aspects of professional attitude, such as politeness and sensitiveness, though the word "attitude" is not explicitly used in the documents (GMC 2006, 2009). There are various definitions of attitude, but one that suits the healthcare context needs to include two elements: the object of evaluation (e.g. patient-centredness) and the expression of the evaluation (e.g. positive conveyance of patient-centredness to the patient) (Martin et al. 2002). These elements are reflected in the GMC guidelines in that doctors and medical students are required, for example, to respect patients' dignity as well as to actively show this in the therapeutic relationship. Professional attitudes appropriate to healthcare services therefore include the cognitive and behavioural dimensions just like attitudes in general. Furthermore, although the word "attitude" is not used in the policy documents (GMC 2006, 2009), the GMC have clarified the expected attitudes that are appropriate to the medical profession by setting out 'the principles and values on which good practice is founded' (GMC 2006:4). These principles and values underlie the professional attitudes and form the basis for defining professionalism in the medical practice (GMC 2006, Hilton & Southgate 2007). Medical professionalism which is 'a set of values, behaviours, and relationships that underpins the trust the public has in doctors' (RCP 2005:14) is not optional, but is essential to 'being a doctor' (RCP 2005:44). In other words, professionalism in the medical practice extends beyond technical knowledge and skills, and incorporates at least a set of attitudes and actions that doctors must have and show in their care for patients (RCP 2005).

In contemporary society, the principles of medical professionalism include, among others, altruism and compassion (RCP 2005, Arnold & Stern 2006, Cohen 2006, Hilton & Southgate 2007). These principles were regarded by some as part of the set of non-cognitive skills necessary for medical students to acquire and to convert into corresponding behaviour, and one area of professionalism-driven behaviour is communication (Kirk 2007). Therefore communication skills, just like attitudes such as respect, are inherent in medical professionalism (Arnold & Stern 2006). It can further be seen that the principles and values of altruism, compassion and respect underpinning professionalism are all in fact aspects of patient-centredness, which Section 2.1.3 below shall turn to.

2.1.3 Characteristics of patient-centred care

Part of what patients expect of good medical practice and high quality medical education is reflected in the GMC (2006, 2009) documents. Patients not only expect their doctors to make good diagnoses and to decide on the right treatments but also expect them to be effective communicators and to demonstrate professionalism (see Crellin 2005). It is apparent that patient expectations of care can be beyond the safety and quality of the healthcare service.

In cases where the healthcare service can be said to have failed, it often seems to be related to an apparent lack of caring (see Francis 2013 especially Ch.20). Advances in medical science have hugely improved the chance of a cure for many conditions. However, such advances are not sufficient for better patient experience of the care received. New drugs, therapies or equipment cannot replace putting patients first. Modern medical practice and education sometimes suffers from being dominated by knowledge and dismissing patient experience (Stewart et al. 2003). Yet patient experience which is a dimension of the quality of care is as important as patient safety and effectiveness of service (Francis 2013. When patient experience is less than satisfactory, it can indicate a lack of caring and may suggest that the service has shifted from being patient-centred to economy-driven.

Historically, doctor-patient relationship has not always been patient-centred. Since the late 20th century, the traditional paternalism characterising doctors' relationship with their patients has been increasingly challenged by consumerism in health care (Beisecker & Beisecker 1993, Brown 2008). Consumerism in health care demands rights, equal or greater power (compared to doctors), information giving, and a role in decision-making in one's own healthcare management. In the paternalistic doctor-patient relationship, however, the doctor dominates the patient, makes treatment decisions and controls information about the patient's condition (Beisecker & Beisecker 1993). Such a relationship, though still for the patient's benefit, cannot be regarded as patient-centred, and even in its best form can no longer meet the diverse expectations of patients. Patient-centredness, which may or may not overlap with consumerism in health care, puts patients first, and respects their feelings towards their own health care. They are given relevant information about their conditions and treatments and are empowered to take part in making clinical decisions. Patients have therefore become "partners" in the healthcare service they receive (Ong et al. 1995). A vital ingredient in such partnership between doctors and patients is good communication which therefore has become an indispensable element of patient-centred care.

2.2 Benefits of effective communication to patient-centred care

In patient-centred care, effective doctor-patient communication refers to not only the exchange of information and facilitation of medical decision-making, but also the maintenance of the therapeutic relationship (Ong et al. 1995, Sliwa et al. 2002, Wong & Lee 2003). Benefits of effective doctor-patient communication are manifold, and can be found in three main areas from the perspectives of patient-centredness: patient satisfaction, patient adherence and patient health outcomes (Stewart et al. 2003).

The benefit of effective communication to patient-centred care lies firstly in patient satisfaction (Ong et al. 1995, Roberts 2002, Wong & Lee 2006). However, not all communication behaviours are linked to improved patient satisfaction. Clinical communication can be divided into two types: one that contributes to the diagnosis, treatment and management of diseases (cure-oriented) and one that attends to the socio-emotional wellbeing of the patient (care-oriented) (Ong et al. 1995). Both categories are integral to the roles of the doctor in the therapeutic relationship, with cure-oriented communication fulfilling the instrumental role and care-oriented communication the affective role. These roles function jointly in providing health care to the patient as a whole person. Although instrumental aspects of doctors' communication such as information-giving, discussion of preventative care and discussing plans for discharge can have positive influence on patient satisfaction, research findings generally support the view that affective communication is mainly associated with patient satisfaction towards the care received (Ong et al. 1995). More specifically, affective communication behaviour such as the doctor showing concern and interest, or empathetic listening is linked to higher patient satisfaction (Deveugele et al. 2002, Ong et al. 1995), and dominant, controlling communication styles are linked to lower patient satisfaction (Ong et al. 1995). However, one type of communication might dominate the clinical encounter. In their study of the consultations of 20 GPs, Deveugele et al. (2002) found that they predominantly used instrumental communication with patients, reacted little to patients' utterances/comments about lifestyle, seldom asked for clarification and had limited sub-types of affective communication(e.g. restricted to social talk). Such evidence suggests that although both cure-oriented communication and care-oriented communication are important in patient-centred care, doctors may not have enough care-oriented communication with their patients, and thus potentially limit the level of patient satisfaction. Given that patient experience is an important dimension of the quality of care (see also Section 2.1.3 above), improving care-oriented doctor-patient communication is therefore a pressing area of patient-centred care to work on.

The benefit of effective communication to patient-centred care lies secondly in compliance to treatment (Ong et al. 1995, Roberts 2002, Wong & Lee 2006). Communication is essential in choosing a treatment that the patient believes in and follows (DiMatteo et al. 1994). Bauman et al. (2008) asked 393 patients to rate the perceived communication skills of their GPs on a 15-item scale and also studied their adherence to treatment. The communication skills of the GPs were found to have a strong effect on the patients' adherence to treatment, and better GP communication skills (as perceived by patients) predicted a higher probability of the patients being adherent, after adjusting for the possible influence from patient age (Bauman et al. 2008). The cross-sectional design in Bauman et al. (2008) was not able to determine any causal relationship between communication skills and compliance, and sampling from patients already concerned about their health limited the study's generalizability to a wider population. However, the findings of Bauman et al. (2008) still provided clear evidence that effective communication is positively associated with adherence to treatment. Similar evidence can also be seen in Roberts (2002) that doctors' communication styles had an influence on patients' satisfaction with care, which in turn influenced their adherence to medication. It is important for patients to follow treatment plans, and effective communication thus plays a vital role in patient-centred care.

The third benefit of effective communication to patient-centred care lies in health outcomes (Ong et al. 1995, Wong & Lee 2006) which can have physiological, functional, symptomatic and emotional measures (Stewart 1995). Optimal outcome can be achieved through patient empowerment, which is in turn dependent on effective communication (DiMatteo et al. 1994). Stewart (1995) found in a systematic review of 21 studies that 16 of them reported positive effects regarding effective doctor-patient communication on patient health outcomes. The reviewed studies were conducted in a wide range of clinical settings, and most reported positive effects on patient health outcomes were found during medical history-taking and the discussion of management plan. Stewart (1995) concluded that effective communication positively influences patient health outcomes in the measures of emotion, symptom resolution, functional and physiological status (e.g. blood pressure) and pain control. Since effective communication has direct benefits to a patient's emotional and physical wellbeing, it was recommended that doctor-patient communication be identified as a central clinical function to be trained at all levels of medical education (Stewart 1995).

Patient-centred medicine also produces outcomes at the provider end (Stewart et al. 2003). Many of these outcomes are also related to effective communication. Doctors' communication with patients has been found to be associated with their history of malpractice claims (Levinson et al. 1997, Ambady et al. 2002). A genuine malpractice claim involves at least some patient interests disrespected, and thus suggests a deviation from patient-centredness in care. Therefore the relationship between doctors' communication features and malpractice claims further points to the benefit of effective clinical communication to patient experience and consequently to patient-centred care. In particular, communication features making the consultation structured and not rushed, respecting patients' perspectives and imparting warmth and friendliness were found to be negatively associated with malpractice claims (Levinson et al. 1997). These features should be encouraged to be used in patient-centred care. Furthermore, Ambady et al (2002) found that the effects of surgeons' tone of voice (dominance and concern/anxiety) on their history of malpractice claims were significant, even with the content of communication being controlled for. A dominant tone of voice can be perceived from a deep and unaccented speech, delivered in a loud volume and uttered at a rather fast tempo (see also Section 2.6). This can convey a lack of understanding and empathy and, together with a perceived absence of concern may give the feeling of indifference on the doctor's part (Ambady et al. 2002). Indifference is compatible with neither medical professionalism nor patient-centredness (see also Sections 2.1.2 and 2.1.3), and coupled with poor clinical outcomes can lead patients to file malpractice claims (Ambady et al. 2002). The evidence on the link between unprofessional features of communication and malpractice claims (and thus alleged violation of the principles of medical professionalism) indirectly suggests that the attitudinal factors in doctors' communication with patients can play a significant role in patient experience with clinical encounters. Other benefits of effective clinical communication are found in, for example, the availability of time (Brown et al. 2003) and physician satisfaction (Wong & Lee 2003). Though these benefits are less immediately related to patient-centredness than the benefits at the patient end, they are still important to the healthcare service. In summary, the multiple benefits of effective doctor-patient communication demand further studies, especially in-depth studies concerning communication behaviours that contribute to these benefits.

2.3 The study of patient complaints

This review now turns to situations in which patients' rights are infringed and the standard of care falls below its statutory requirements and rightful expectations. There are occasions on which things can go wrong in a healthcare system that is otherwise functioning well. Formal patient complaints not only give warning signals when the quality of care fall below the expected standard (Taylor et al. 2004) but also highlight areas of service in which improvements are needed (Anderson et al.

2001, Chavan et al. 2007). The call for a caring, compassionate and committed healthcare service highlighted in the Francis (2013) Report echoes studies of patient complaints within different healthcare systems and at various levels. This section starts with a more general overview of NHS complaints before then moving to discuss some areas of complaints in detail.

2.3.1 Overview of patient complaints in the NHS

Patient complaints can occur for various reasons (HSCIC 2012), particularly when the standard of care falls below the level expected (Taylor et al. 2004). Therefore formal patient complaints can be a useful indication of patient satisfaction (Taylor et al. 2004, Chavan et al. 2007). In a large healthcare system that delivers services to a great number of patients, such as the NHS, it is natural that cases of patient dissatisfaction may occur which in turn result in formal patient complaints (Chavan et al. 2007). The large scale of the services delivered in the NHS, however, also translates to a large volume of complaints received. Complaints about NHS care and treatment are dealt with following a statutory procedure (HSCIC 2012) and the statistics, published annually, can be used to identify patterns of change and areas most needing improvement.

During the period of 2011-2012, the NHS in England (henceforth NHS) received 149,765 written complaints about its hospital and general practice (GP) care. This averages to more than 3,000 written complaints a week, an increase of 8.3% from 2010-11 (HSCIC 2012). The scale of complaints in the NHS has been similarly large since 1997 and has remained so (HSCIC 2009, 2012), and a summary of this general overview is produced in Figure 2.1 below, using statistics from the HSCIC.

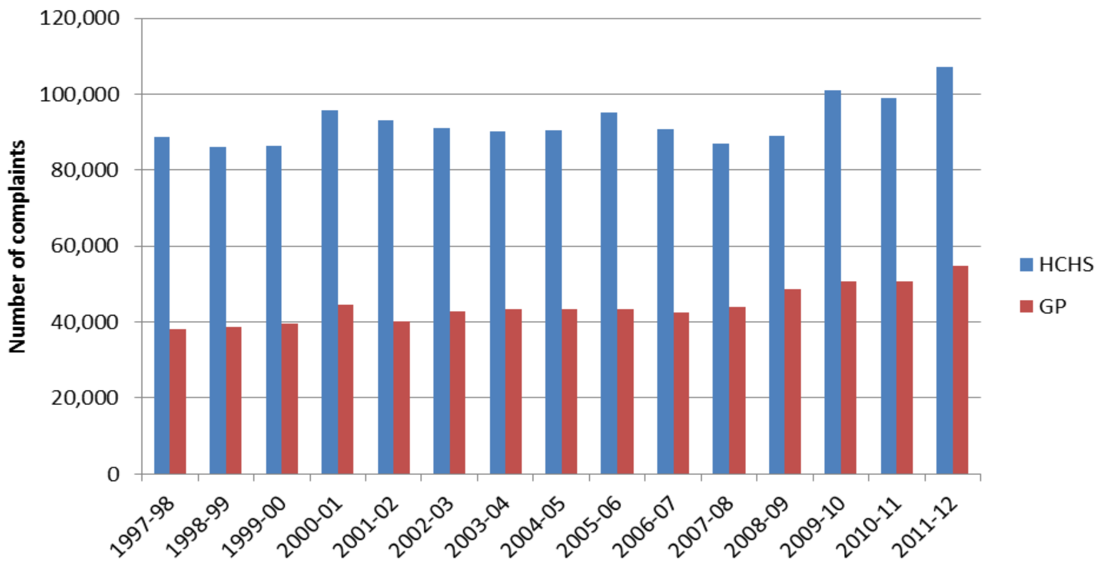


Figure 2.1 Numbers of written complaints to the NHS in England, 1997-2012 (HCHS: hospital and community health service complaints; GP: general practice (including dental) health services complaints. Source: HSCIC 2009, 2012).

It can be clearly seen in Figure 2.1 that the scale of complaints about the NHS services has been substantial for more than a decade or so (HSCIC 2009, 2012). A discussion of the potential link between complaint rate and the expansion of the NHS services is beyond the scope of the current investigation. However, two patterns can be observed in Figure 2.1. Firstly, the volume of complaints about the NHS has remained relatively stable over the years: neither drastic rises nor sharp falls can be found in recent history. Secondly, although both large in volume, the complaints against NHS hospital services have outnumbered the complaints related to NHS GP services over the years. This is not unexpected though, since hospital treatments and hospitalisations can involve more complex or serious conditions, can involve higher emotional stakes and can draw more attention from both the media and the public. A hospital environment (especially a large one) can also be a more intense experience than that of a GP clinic, and it might be more difficult to achieve satisfactory patient experience in an intense environment than in a relaxed one. Furthermore, patients might have higher expectations of hospital services than of GP services which, when not met, can lead to lower satisfaction. Though this chapter does not intend to discuss whether or not the standard of GP services and that of hospital and community health care are different, the urgency of improving patient experience does seem even greater in hospital services than in GP clinics. Nevertheless, it is necessary to examine the particular aspects of care that fail to meet patient expectations in both healthcare settings.

2.3.2 Areas of patient complaint

Examination of the complaints data discussed above revealed that the greatest number of patient complaints in both NHS hospital and community health services and NHS GP/dental services related to clinical treatment (46% and 36% respectively, HSCIC 2012). However, the figures relating to aspects of communication also contributed to a significant proportion of the complaints received: in the hospital and community health services setting, problems categorised as pertaining to 'staff attitude' and 'communication/information to patients (written and oral)' constituted 12% and 10% of complaints received during this period, respectively. Complaints relating to NHS GP/dental services in England during the same period reveal a similar picture, with perceived problems relating to poor 'communication/attitude' resulting in 22% of complaints received (HSCIC 2012). In an earlier report (HSCIC 2009), similar figures can also be found, as summarised in Table 2.1 below.

Table 2.1 Top subjects areas of complaints in the NHS in England (2008-09)

	HCHS		GP	
Subject	**n**	**%**	**n**	**%**
clinical	37,149	41%	14,868	32%
attitude	11,332	13%	11,003	24%
communication	89,678	10 %		

Source: HSCIC (2009)

The official reports (HSCIC 2009, 2012) did not explain the reason for combining communication and attitude into one subject area of complaints in GP/dental services. It may be, however, reasonable to regard attitude as an important aspect of communication, since only when staff attitude is communicated (regardless of the channel) can it be perceived by service users and therefore become a subject of complaint. In the case of hospital and community health services, though, attitude has always been a significant subject area of complaint. In fact, since 2004 (the earliest of these reports), attitude has always accounted for approximately 12% of all complaints and has often been the second largest category of complaints only behind all areas of clinical treatment (HSCIC 2009, 2012).

It is apparent that NHS patients most frequently complained about the clinical treatment that they had received, which is perhaps not surprising. After all, quality treatment is what patients use the service for and what the service pledges to deliver. And when the treatment falls below the expected standard, complaints can follow as a consequence. However, the fact that staff attitude accounted for a significant proportion of the patient complaints in the NHS is quite alarming. It suggests that it was *how* the patients received their treatment that was not up to their expectations, instead of (and/or in addition to) *what* they received, i.e. the treatment itself. Even if there might be times when it was not possible to give a patient the treatment he or she expected, it is less possible to justify any inappropriate attitudes with which the service was delivered. It is also alarming that this situation has remained largely unchanged for almost a decade. Before taking the discussion further, it will be useful to know whether these attitudes that patients have complained about are isolated problems. For this purpose, studies of patient complaints from other parts of the world also need to be included in the review. If complaints about attitudes in health care are widespread, then they are common rather than isolated problems and deserve further study, especially those that include different elements of attitude.

A great number of studies have been conducted on the content of patient complaints, both in the UK and in other countries. Interestingly, despite the differences in the scope of investigation and the reported complaint rates between studies, attitude and communication are found to be two common themes emerging from many of the investigations. Table 2.2 presents a synthesis from the findings of some of such studies carried out in various countries.

Table 2.2 Rates of patient complaints related to attitude and communication in selected research studies

Study	Location	Attitude	Communication
Anderson et al. (2001)	Australia	45% (*n* = 621)	
Chavan et al. (2007)	England	18% (*n* = 17)	
Hunt & Glucksman (1991)	England	38% (*n* = 46)	12% (*n* = 14)
Jangland et al. (2009)	Sweden	11% (*n* = 194)	
Jimenez-Corona et al. (2006)	Mexico	>80% [a]	
Montini et al. (2008)	US	28% (*n* = 347)	17% (*n* = 212)
Taylor et al. (2004)	Australia	29% (*n* = 7809)	
Wofford et al. (2004)	US	23% (*n* = 404)	

Note. Complaints related to communication between professionals are not included.
[a] Identified by the authors through the qualitative analysis of a total of 639 complaints.

It can be seen in Table 2.2 that the proportion of attitude-related complaints reported in these studies is mostly larger than that for NHS hospital and community health services (12%) in England in 2011-2012. With the complaints about attitude and communication combined, these studies mostly reported a proportion similar to or higher than those for NHS hospital and community care (22%) and GP/dental services (24%) in England in 2011-2012. It is not certain whether the type of healthcare settings in these studies played a role in influencing the amount of patient complaints that were linked to attitude and communication. However, the rates in hospitals delivering specialised care, e.g. the eye hospital in Chavan et al. (2007) and the 1100-bed university hospital in Jangland et al. (2009) are markedly lower than the general NHS ones. Nevertheless, Hunt et al. (1991) reported a notably high rate of attitude-related complaints (38%) in their study of an inner-city accident and emergency department (A&E) in London over seven years. This might reflect the usually tense atmosphere associated with A&E where patients' expectations of professional attitude could be even higher than in other hospital departments. Even so, this highlights the significance of professional attitudes in patients' expectations in terms of their experience with the healthcare service. If professional attitudes are not demonstrated in these encounters, or are not perceived by the patients, complaints can be expected to follow. Furthermore, within the very large proportion of communication-related complaints reported in Jimenez-Corona et al. (2006), misunderstanding occurred more frequently than lack of information. It seems that even though sufficient information is exchanged in clinical encounters, if the information exchange is not performed in an appropriate manner, misunderstanding and its resulting complaints are still possible. It should also be pointed out that some demographic factors in the Mexican study are noteworthy: complaints were made predominantly by female patients and most complaints were from patients paying fees (through social security). These groups of patients would most likely have higher healthcare expectations due to the costs incurred by their treatment or may be more capable (due to their relatively high social/educational level) of making complaints (Jimenez-Corona et al. 2006). Future studies with a more robust design should be able to consider or control for these factors.

Studies in patient complaints essentially rely on data collected from patients' perspectives. How communication and attitude are judged can vary greatly because of this. However, patient complaint studies do not generally intend to analyse the communication behaviours and attitudes that are reported in depth. Instead, these studies often describe complaint rates in categories according to subject areas, specialties of care, or patient demographics as related to the complaints, for the purposes of quality assurance or risk identification (Taylor et al. 2004). In addition, a complaint can be made about multiple subject areas and can be recorded more than once in some reports (see e.g. HSCIC 2012). Nonetheless complaints related to attitude and communication are quite clear categories. In other words, although some variation can be seen in the figures of attitude and communication-related complaints in the studies reviewed, partly due to the different methods used (some studies considered attitude as an aspect of communication), the attitude and communication issues can be quite clearly identified in these studies.

Examples of inappropriate attitudes or unprofessional behaviours occuring in clinical encounters can be found in the studies reviewed. These include offensive attitudes and lack of care (Anderson et al. 2001), or even arrogant, rude and condescending behaviour (Wofford et al. 2004) experienced by patients or their representatives. Such reports contradict the fundamental pledges of the healthcare service as well as basic public expectations (see Section 1 above). These reported attitudes and behaviours are therefore examples of a lack of professionalism present in doctors' interaction with patients.

While it may be that some of the health professionals involved in the complaints did indeed possess or display inappropriate attitudes, some other complaints may be due to patients' misperceptions of the attitudes expressed, rather than the attitudes themselves always being inappropriate in themselves. That is to say, for a professional to have an appropriate attitude is one thing, but for this to be fully understood and comprehended by a patient is quite another. Attitudes not clearly conveyed are a potential source of misunderstanding. Therefore in order to deliver patient-centred care, professionals not only need to have the appropriate attitudes in terms of their cognition (i.e. beliefs etc.) but also must have the essential skills to display them in behaviour (communication) so that the misunderstanding of attitudes in clinical encounters can be prevented (Martin et al. 2002). Sometimes even innate skills need training to reach a proficient level, and communication skills that work to display professional attitudes are no exception.

2.4 Medical communication skills training

2.4.1 Curriculum development, training content and evaluation

Although training is needed to develop the necessary medical communication skills for conducting patient-centred consultations, and to display professionalism in practice, formal communication training is relatively new in the history of medical education. Before the 1980s, formal communication skills teaching did not exist in undergraduate medical programmes in the UK (Brown 2008), although by the mid-1990s the majority of UK medical schools had incorporated communication skills teaching into their undergraduate curricula (Hargie et al. 1998). The progress of such development, however, was not an easy one. Among the 19 UK medical schools that responded to the study by Hargie et al. (1998), the proportion of communication skills teaching as related to the overall undergraduate programme was uneven, and varied between 5-25%. Furthermore, most medical schools experienced problems during the course of implementing their communication skills teaching; they were faced with inappropriate teaching accommodation and related facilities, the cost of simulated patients and tutors, and staff matters including availability, lack of experience, overt staff opposition, lack of curriculum time given to the training, and negative student attitudes (Hargie et al. 1998). Since then efforts to address these problems, that could potentially affect the training outcome, have been made in order to assure the effectiveness of medical communication (Brown 2008).

Medical students' negative attitudes towards communication skills training can particularly hinder their achievement of the targets required to become future practitioners and professionals. Several themes in terms of the negative attitudes towards communication skills learning among medical students have been reported. In a study at the University of Nottingham, UK, Rees & Garrud (2001) found that some first year students failed to comprehend the importance of such learning, that they failed to take such learning seriously, viewing it as an easy subject that is not assessed in the same way as, say, subjects involving pure science, and that they failed to admit their personal limitations in communication skills. These negative attitudes are clearly incompatible with the objectives set by the GMC, that require medical students to be self-directed and lifelong learners who must be clear regarding understanding their own limitations and to be capable of seeking help with these limitations (Rees & Garrud 2001, see also GMC 2009). Since communication skills are inherent in medical professionalism (Arnold & Stern 2006), negative attitudes towards learning communication skills can also hinder the development of professionalism among medical students.

These negative attitudes towards communication skills training were also found in medical schools outside the UK. Anvik et al. (2008) found that Norwegian medical students' cognitive attitudes towards communication skills remained stable throughout the course of 6 years in education, but their affective attitudes towards communication skills declined as they progressed in their medical education. This means that their "feelings towards and experience from" the communication skills training received became negative (Anvik et al. 2008: 276). On the one hand, it is essential that medical students undertake communication skills training, the training itself can be a laborious and challenging process that best benefit from positive, both cognitive and affective, attitudes[1] towards it (Anivk et al. 2007, Anvik et al. 2008). On the other hand, negative attitudes such as not taking communication skills learning seriously or a lack of interest or motivation in communication skills learning can persist and be manifested as poor communication behaviour (Anvik et al. 2007). Attitudinal change similar to that found in Anvik et al. (2008) was also reported in a cohort study on Canadian medical students by Woloschuk et al. (2004). Two self-reporting instruments measuring a range of attitudinal aspects were administered to a University of Calgary Medical School cohort at their entry into medical school, at the end of their preclinical training and the end of their clerkship. The communication sub-scale of one of the instruments recorded a sharp decline in the students' attitudes towards communication in clinical encounters, even with gender differences in attitudes controlled for. The observed significant fall on the communication sub-scale scores, between the end of preclinical training and the end of clerkship, could be associated with the absence of communication skills teaching during clerkship (Woloschuk et al. 2004). The above studies identified negative attitudes that might prevent medical students from sufficiently developing communication skills and, consequently, professionalism in practice. Medical communication skills training programmes need to be able to help students overcome these negative attitudes that can risk making them poor communicators lacking professionalism.

In the UK, communication skills training has been integrated within the core medical undergraduate curriculum as established by the GMC, thanks to continuous development despite some challenges (Brown 2008, Hargie et al. 1998). *Tomorrow's Doctors* (GMC 2009) specifically requires medical students to acquire an essential set of skills in communicating effectively with patients. Such skills as breaking bad news and discussing sensitive issues are particularly highlighted in addition to the ability of carrying out consultation tasks such as taking medical history and providing explanation and advice. These are also in addition to the attitudinal principles specified in the same document, as discussed in Section 2.1.2.

[1] The cognitive, affective and behavioural elements of attitude represent a classic tripartite view of the structure of attitude (Rosenberg & Hovland 1960), which though challenged by some (see e.g. Eagly & Chaiken 1998), still finds influence in recent work (e.g. Martin et al. 2002, Anvik et al. 2007).

Medical communication skills training therefore needs to go beyond the exchange of information during the consultation and cover aspects of professional attitudes that are essential for establishing and maintaining the patient-centred, therapeutic relationship, too. Depending on how teaching is delivered, the aspects of professional attitudes as specified by the GMC and pledged in the *NHS Constitution* (DH 2013a) can be incorporated into the learning objectives of clinical skills in various ways. All UK medical schools have worked together and have agreed on the domains of the clinical communication curriculum under the authority of the UK Council of Clinical Communication Skills Teaching in Undergraduate Medical Education (von Fragstein et al. 2008), and professional attitudes cannot be separated from the core or any domain of the agreed curriculum. For example, the widely-used Calgary-Cambridge model (Silverman et al. 2005) of communication skills teaching provides a framework that incorporates both the process and the content of the medical interview that students must be able to command.

This framework aims at teaching learners a set of skills needed for conducting a consultation with patients in tasks including initiating the session, gathering information, giving explanation, closing the session, as well as providing structure and building their relationship (Silverman et al. 2005). The process guide of the model lists 71 individual skills that serve as a training manual according to which medical students should demonstrate particular behaviours, such as asking open-ended questions and using empathy (Silverman et al. 2005, Kurtz et al. 2005). In addition to the process guide, the Calgary-Cambridge model of medical communication skills teaching also includes a content guide that conceptualises the information required in the consultation and the medical record (Kurtz et al. 2003). This content guide adds new content e.g. patient's perspective to the traditional ways of recording medical information e.g. past medical history (Silverman et al. 2005). These two specific guides, according to the authors of the Calgary-Cambridge model, should work in unison to develop the skills essential to a consultation, incorporating both biomedical and patient perspectives (Kurtz et al. 2003). The Calgary-Cambridge model is very clear about its emphasis on the specific skills listed in the communication process guide (Kurtz et al. 2005), although not all skills in the Calgary-Cambridge model are discussed in detail (see Section 3.4.2).

Communication skills training for healthcare professionals, as well as students, is generally found to be effective (Gysels et al. 2004, Parry 2008). However, the measures of effectiveness varied greatly among the studies reviewed in Gysels et al. (2004) and Parry (2008), and so did the research methods used. While it was not possible for meta-analysis to be conducted in these systematic reviews (Gysels et al. 2004, Parry 2008), distinctive themes can be seen in the reviewed studies. It emerged from the literature that although no training is perfect in terms of effectiveness, there are training topics and methods that can improve its benefits. In particular, the effect of training could be observed in emotionally laden areas, such as communication about life-threatening conditions (Wilkinson et al. 1999, also cited in Gysels et al. 2004). The 110 nurses in cancer and palliative care studied in Wilkinson et al. (1999) were found to have, after a communication skills training programme, improved on their discussions with patients about emotional issues such as diagnosis, prognosis and psychological impact. Such improvement is also benefiting to patient experience. However, in order for the effect of training to be maintained, a sustained effort is needed to prevent the trained communication skills from declining (Maguire et al. 1996), in agreement with the GMC requirement for lifelong learning. Furthermore, for the effectiveness to be maximised, the training should address attitudes and beliefs, and the trainees must be interested in the training and be open to change (Parry 2008). In previous research, the effectiveness of communication skills training was often measured with the presence/absence or amount of certain behaviours, but rarely incorporated the appropriateness of such behaviours in clinical interaction (Parry 2008). The same can be said about the Objective Structured Clinical Examination (OSCE) communication stations, in which the communication skills of the candidates can be assessed with the presence or absence of a set of behaviours observed by the examiner. Based on the theory that behaviour (or its antecedent intention) can be determined by attitude (Ajzen & Fishbein 2005), it can be concluded that the observed communication behaviours in OSCEs or similar assessments can be an indicator of candidates' underlying attitudes. A discussion of the measures of expressed attitude is presented in Chapter 3, on research methodology. However, it is apparent that the study of medical communication, especially doctors' expressions of their attitudes, is best conducted with a view to the appropriateness of communication behaviours and the attitudes indicated by them in a clinical context.

As well as a method of assessment, the OSCEs can be a tool for generating research data to be used in medical communication studies. Using transcriptions for 28 of 309 video-recorded OSCE consultations that they viewed, Roberts et al. (2003) identified some different communication styles between stronger and weaker candidates. This study adopted the interactional sociolinguistics approach to discourse analysis and focused on “how individuals differ in the ways in which they interact with and understand one another” (Roberts et al. 2003:194). The good candidates were found to use empathetic communication styles overall, which elicited patients’ agendas, involved the patient in joint problem solving, and avoided embarrassment, resulting in solidarity and persuasiveness. In contrast, the communication styles of the failing candidates generally drove through the medical agenda, were judgmental, confused the patient, and created interactional discomfort, leading to distance and lack of mutuality (Roberts et al. 2003). It was pointed out that although the weaker candidates were taught the ‘rapport words’ which they used when communicating with the patient in an attempt to establish the therapeutic relationship, they failed to achieve the ‘rapport work’, that is, they did not seem to know how to use the communicative formulae effectively (Roberts et al. 2003:194). While the strong candidates sensitively tuned in to the contexts of the communication to achieve its empathetic goals, it was noted that the weaker ones did not benefit from their ‘trained empathy’ (Roberts et al. 2003:200). The ‘trained empathy’ is hardly more than a reproduction of the empathetic words taught in training, and not surprisingly seems counterproductive. It is the contextual appropriateness that contributes to the empathetic sense in an expression; formulaic empathic words out of context may not be perceived as genuine empathy at all (Roberts et al. 2003). Perhaps communication skills teaching needs to equip medical students with the ability and readiness to apply the subtlety of language to practice instead. In sensitive communicative situations, mechanical, inflexible expressions are not likely to be effective: one size cannot fit all. Therefore it cannot be sufficient to only set attitudinal goals, such as empathy, and expect natural attitudinal development and a positive behavioural outcome for all medical students after only limited classroom input. Specific training on the display of professional attitudes in communication behaviour is likely to lead to a more positive outcome, especially for medical students with weaker communication skills.

2.4.2 Gaps in current curricula

From the literature reviewed above, it seems that current medical communication skills training, especially that which utilises the Calgary-Cambridge model, mainly focuses on the content and structure of clinical consultation. That is, more attention is given to *what* should be said rather than *how* things should be said. While it is true that the Calgary-Cambridge guide to communication process skills (Silverman et al. 2005) intends to complement the traditional teaching of consultation content that is solely based on the biomedical perspective, it mainly provides a set of skills to be grasped in order to conduct the consultation session. In other words, such training primarily aims at providing students with the skills needed to perform each task in a structured, patient-centred consultation rather than discussing the ways how each task should be approached in order to be patient-centred.

Although professional behaviours in the consultation are among the expected training outcomes of the Calgary-Cambridge model, it does not actively explore the whole range of professional attitudes that can be displayed. The Calgary-Cambridge model is a predominantly skills-based approach: although attitudes are admitted to be vital in communications with patients, these do not occupy a substantial part of the training. Instead, attitudes are regarded as a natural development of the skills acquisition process, and without this development the learners may be prevented from using the already acquired, appropriate skills for the consultation (Kurtz et al. 2005). It should be pointed out that this training model mainly emphasises the cognitive domain ('underlying intentions, values and beliefs'; Kurtz et al. 2005:60) of attitude. However, the domains of attitude can also extend to affect and behaviour. A narrow definition of attitude might not facilitate learners to develop a full set of professional attitudes. Even some attitudinal aspects required in the Calgary-Cambridge guide to communication process (Silverman et al. 2005) are, in fact, not free from affective attributes. For example, empathy, sensitiveness and support (although discussions on how these attitudes related to medical professionalism can be shaped and displayed during clinical encounters provided by the Calgary-

Cambridge model are scarce). It remains unknown how skills for conveying professional attitudes in medical communication can be acquired by students using this training model.

Last but not least, non-verbal communication (e.g. eye contact, facial expression and voice) is encouraged in the Calgary-Cambridge model. Non-verbal communication plays a preponderant role in conveying “attitudes, emotions and affects” (Silverman et al. 2005: 124) and is associated with positive clinical experience and outcomes (Hall et al. 1995, Griffith et al. 2003, Deladisma et al. 2007). Furthermore, both verbal and non-verbal communication must work cooperatively in medical consultation, since when there is contradictory information, non-verbal communication will prevail against verbal communication (Silverman et al. 2005). However, the model only introduces non-verbal communication as a skill for relationship building. According to the model, the building of relationships should be pursued throughout the process of the consultation, but the learners may not necessarily appreciate that non-verbal communication can take place at all stages of the consultation, in order to make the interaction effective. Although a range of non-verbal features are touched on in the model (for example, eye contact, facial expression, posture and vocal cues), the introduction is brief and the discussion does not go into great detail. Students could benefit from more in-depth discussion on non-verbal behaviour, such as the use of a specific feature to convey an attitudinal message, for instance. However, this is not provided in the Calgary-Cambridge text. An example can be found when the skills of giving the “accepting response” are introduced, which include accepting patients’ expressions non-judgementally and acknowledging their ideas and feelings (Silverman et al. 2005:129). Among these skills, the use of ‘appropriate non-verbal behaviour to make space for the patient to say more’ (Silverman et al. 2005:130) is suggested, but the model does not specify any particular behaviour (e.g. nodding etc.). In another section on communicating empathy, examples of empathetic statements are given, but the appropriate tone of voice for these expressions is not discussed. Recalling the findings in Roberts et al. (2003) on “trained” empathy, the formulaic empathetic statements, if reproduced without tone of voice appropriate to the communication context, might even compromise the rapport-building.

From a linguistic point of view, tone of voice in speech is conveyed by a range of features related to pitch, volume and tempo. These are introduced in the Calgary-Cambridge model (Silverman et al. 2005) when non-verbal behaviour is discussed. However, communication skills texts, Silverman et al. (2005) included, do not generally explore this topic in depth, with very few exceptions. One such exception (O'Toole 2008) includes more detailed description of how these features (e.g. different speaking rates) can characterise various emotions. O'Toole (2008) classifies non-verbal communication into several categories, separating the vocal features (or more specifically prosodic features in speech, as reviewed in Section 2.6) from those such as gesture and space. This approach can be very helpful since the prosodic features can convey messages in doctor-patient communication differently from gesture and space. The prosodic features (e.g. pitch, volume and rate) are responsible for communicating attitudinal and emotional information in speech by enhancing or modifying the verbal meanings, while gestures etc. can function independently of any speech. Without a discussion of these prosodic features in at least some depth, medical communication skills training risks generating a significant gap in terms of achieving its outcomes, in that its learners might not have the skills necessary for displaying professional attitudes in speech, as intended.

2.5 Foreign language in health care

Both medical education and medical practice need to address the cultural diversity in the British society today where 'both patients and doctors now came from a range of backgrounds that may well not share a common culture or language' (Brown 2008:272). Clinical interactions in a non-native language, whether spoken by the patient or spoken by the doctor, can be challenging. Roberts et al. (2005) found that language and cultural differences due to patients' non-native English caused over 20% of the misunderstandings in 232 consultations they studied in four inner London practices where the sociolinguistic diversity is particularly apparent (the authors identified native speakers of 30 languages other than English). Therefore extra effort is needed to minimise misunderstanding during the consultation with patients of limited English proficiency. Similarly doctors who practise in a non-native language can be challenged where effective communication in their professional relationship with patients, including native speaking patients, is vital.

There was a large volume of media coverage of a case in which the unlawful death of a British patient was the result of the injection of 100mg of diamorphine (instead of the recommended dose of 10mg) by a German doctor flown into the UK to cover out-of-hours service. The doctor's English language proficiency was questioned and the media further revealed that several similar overdose cases also involved foreign doctors in out-of-hours service (see Smith 2010). It was also reported that the GMC suspended more foreign-trained doctors than British-trained doctors, and approximately two-thirds of the doctors erased from the register were qualified outside the UK (Devlin 2010). In fact many of the top 20 countries of qualification for the GMC registered doctors are not considered as English-speaking (e.g. Germany) or use varieties of English that are distinctively different from Western native ones (e.g. India ; see www.gmc-uk.org). Even though there might be political issues around EU doctors practising in the UK, any problems about their English language skills must not be ignored by the health care systems of a more mobile Europe.

Although foreign doctors, both from EU and non-EU counties, are reportedly more likely to face more disciplinary actions from the GMC than British doctors, the reason behind this has not been clarified (Allen 2009). The mobility of doctors in Europe is now safeguarded by EU-wide regulations, which have been influenced by a complex set of factors (Jinks et al. 2000). However, difficulties for doctors to practice in a foreign country (e.g. EU doctors in the UK) exist in several areas, namely: healthcare systems and local practices and languages (Allen 2009). The language issue about foreign doctors is apparent, and while this can be addressed through adequate training and robust assessment, currently a clear proficiency standard, especially for EU doctors, has been slow to establish. This situation has led to a recent recommendation in the Francis (2013) Report that all healthcare professionals coming to the UK to practise should be subject to an English language proficiency requirement. In addition, there can be language factors other than the language proficiency of non-native doctors that influence their communication with patients. Compared with proficiency in the non-native language, these factors can be subtler and more complex to test.

A study by Van del Poel & Brunfaut (2010) of doctors speaking Dutch and Swedish, respectively, as their second languages (L2s), found several differences in language use between native speaking doctors and non-native speaking doctors. For instance, the L2 Dutch-speaking doctors were found to differ from the native Dutch norms in syntactic modification (e.g. should, would), which, combined with a lack of politeness-marking lexical items (e.g. please, thanks), and particles (e.g. perhaps, just), could be experienced, in a medical context, as overly 'direct, impolite, crude, brutal or even unkind' (Van del Poel & Brunfaut 2010:118). Compared to this, the L2 Swedish-speaking doctors used a significantly larger amount of linguistically complex syntactic modification compared to native norms, which could be associated with patients' perception that foreign doctors in Sweden took more time and were genuinely more interested in their patients than native doctors (Van del Poel & Brunfaut 2010:119). While admitting the limitations present in using written rather than spoken language data, Van del Poel & Brunfaut (2010) recommended that more studies should be carried out on L2 doctors and their clinical interactions in order to investigate the effects of pragmatic-linguistic differences on communication.

This recommendation of Van del Poel & Brunfaut (2010) can be interpreted as providing several messages for both health communication research and communication skills teaching. Firstly, there is a great amount of underrepresentation present in existing literature concerning studies on the communication behaviours of non-native speaking doctors or medical students. Given the global nature of medical practice and education, the urgency of including these groups in health communication research is apparent. Secondly, the studies of non-native speaking doctors and medical students need to go beyond language proficiency, and focus on features of communication that matter in terms of actual patient experience. In UK medical schools, at least, the English language proficiency requirement is strict for applicants with first languages (L1s) other than English. The Norwich Medical School at the University of East Anglia, for example, requires these applicants to demonstrate proven English proficiency with a minimum IELTS score of 7.5 or equivalent. English proficiency at this level should be comparable to or above that of good and very good users of English (*IELTS Guide for Educational Institutions, Governments, Professional Bodies and Commercial Organisations*).

However even such high level of English proficiency does not translate directly into excellent skills of communication with patients in English, nor can this guarantee satisfactory patient experience with their communication. This suggests a need to use data from real communication in a study of non-native English speaking medical students or doctors. Thirdly, as introduced in Section 2.4 above, patient experience with communication in the clinical encounters is not only the result of actual words spoken, but is affected by a wide range of communication features, including non-verbal communication. Since prosodic features such as pitch, volume and rate factor into the communication of attitudes, emotions and affects, evidence on these in health communication should be collected from both native speaking and non-native speaking medical students or doctors. Consequently, communication skills teaching informed with such evidence is likely to benefit both of these groups, who have common goals in terms of effective communication with patients. Section 2.6, below, is devoted to a discussion of prosodic features with regards to their use in health communication.

2.6 The prosodic features of speech

In speech communication, ideas are not only encoded in spoken words that are in turn made up of speech sounds such as vowels and consonants, but also rely on the melody and rhythm of speech for information (e.g. tone of voice) beyond the verbal content. All these speech signals work coherently to deliver messages, explicit and implicit, in natural sounding utterances (Nooteboom 1997). Although advancements in speech technology has been able to produce quite natural sounding synthesised speech (Schröder 2009), for instance, Apple's "Siri" voice interface, historically, machine generated speech sounded unnatural and robotic. An example of such an older generation speech synthesizer is that used by Stephen Hawking, Speech Plus CallText 5010 (www.hawking.org.uk). While Hawking has an enormous amount to communicate, the rather unnatural sounding speech produced by his synthesiser can demand some extra effort, on the listener's part, to follow, especially when the subject matter is unfamiliar to the listener. For speech synthesisers to express basic emotional meanings, such as joy and sadness, parameters of prosody (i.e. melody and rhythm of speech) can be manipulated to produce perceivable expressive output (Schröder 2001). What is interesting about expressive speech synthesis, in particular for this study, is how prosodic features can be used to the advantage of increasing the acceptability for the listener. Transposed here, if a patient, who may or may not be familiar with the technical details of the illness he or she is consulting their doctor about, he or she might also have difficulty taking in the message from the doctor.

The fact that the attitudinal or emotional state of the speaker is more often reflected in his or her voice than in the words spoken can be intuitively felt by the listener. Such intuition is also supported by a wealth of research evidence (see e.g. Scherer et al. 2003). Of course, there are communication channels other than voice that are also capable of conveying attitudes and emotions (such as body language) but prosodic features in one's voice can enhance and modify the content of one's speech (Mozziconacci 2001). In this section, the nature, functions and perception of prosodic features, as well as characteristics of non-native English prosodic features, are reviewed before a summary is given in Section 2.7 to place the discussion of speech prosody within a context of health care, where patient experience is paramount.

2.6.1 The nature of prosodic features

Whether the speaker speaks softly or loudly, whether the speech is slow or fast, and whether they use a particular tone of voice can all make a difference to the listener's understanding or interpretation of the speaker's intent. These properties, that make the difference in speech communication, are not limited to the levels of phonemes (e.g. vowels and consonants), syllables or words, nor can they be readily transcribed in normal orthography. They are examples of prosodic features in speech that belong to a group of suprasegmental properties (i.e. beyond the segments of speech, see Nooteboom 1997). While sometimes the terms 'suprasegmental' and 'prosodic' are used interchangeably in research literature, this study reserves 'suprasegmental' to refer to the mode of description and 'prosodic' the actual features in speech, such as pitch, loudness and tempo (c.f. Fox 2000). Prosodic features are not restricted to the domains of pitch, loudness and tempo, yet these three domains and their linguistic and affective functions have received extensive treatment in previous research and are central to this study. The features of pitch, loudness and tempo can all be measured objectively using software for acoustic-phonetic analysis, and their measurable acoustic properties are discussed in the subsections below.

Adopting a perspective of phonetics that studies the physiological, acoustic and auditory attributes of speech sounds, this study defines prosodic features as features of speech that are not confined by the place and manner of articulation and can extend beyond one segment (e.g. a vowel or a consonant, see Laver 1994, Nooteboom 1997, Fox 2000). In terms of the physiology of articulation, prosodic features originate in the subglottal and laryngeal components and are thus more basic than segmental features which are controlled by the supralaryngeal component, since the latter can only modify the airstream already specified by prosodic features at the source. That is to say, the pulmonic airstream which is regulated by the lungs, trachea and associated muscles together with voicing and pitch regulation by the larynx precedes pharyngeal, nasal and oral features which serve as a variable filter in speech production (Fox 2000; see also Johnson 2003, Hewlett & Beck 2006 on source-filter theory of speech production). In this study in particular, prosodic features refer to the primal pitch, loudness and temporal signals underlying a speech string, that are responsible for not only linguistic meanings (i.e. literal content) but also paralinguistic information (emotion, mood, attitude etc.).

2.6.1.1 Pitch

Since pitch is related to all kinds of sounds in life (e.g. voices, instrumental music and even noise), the idea is almost self-explanatory (Hewlette & Beck 2006). What is perceived by the human ear as the pitch of a voice in physical reality is the fundamental frequency (F0) used by the speaker. The F0 refers to the rate at which the vocal folds vibrate; the faster the vocal folds vibrate, the higher the F0. The effect of a faster vocal fold vibration (i.e. a higher F0) on the listener’s perception is a higher pitched voice, and the unit of measurement for F0 is Hertz (Hz), i.e. cycles per second, as for any frequency.

For communicative purposes, pitch variations or movements are more important than absolute pitch values (Laver 1994). Therefore, the level and extent of pitch movements can be characterised as the mean and range of fundamental frequency values. There are physiological pitch variations due to the anatomy of the vocal folds. Generally speaking, women have higher pitch ranges than men, and young children can have even higher pitch ranges. The typical pitch range in conversation is about 80-200 Hz for men and 180-400 Hz for women (Nooteboom 1997), while the maximum pitch range in conversation is approximately 50-250 Hz for men and 120-480 Hz for women (Laver 1994). During puberty, the pitch range change is generally more noticeable in boys than in girls. Fundamental frequency, or its perceptual correlate, pitch, is considered to be the major difference between male and female voices and thus can have implications for methodological consideration in research design, which will be addressed in later chapters of this thesis. Aside from the physiological factor, pitch range differences can also be used in utterances for paralinguistic (e.g. attitudinal) purposes. The paralinguistic information encoded in speech is intertwined with the linguistic information in that utterance as well as the cultural norm in the speech community (Laver 1994). This paralinguistic function of pitch range is of interest and is discussed in Section 2.6.2.2.

While the Hz scale measures exactly the physical property of the F0, psycho-acoustic scales for the description of pitch measurement also exist which usually attempt to correspond better to human perception (Nooteboom 1997, Nolan 2003). These include semitones, mels, Bark and Equivalent Rectangular Bandwidth (ERB-rate, or simply ERB for short), but the most useful measurements for comparing pitch distances are reported to be semitones (logarithmic) and ERB-rates (near-logarithmic) (Nolan 2003). The usefulness of psycho-acoustic scales in pitch distance (the size of difference between two points in a pitch contour) is clear: the same tune sung by a man and a woman will be perceived as the same tune, but the pitch distance measured in Hz will be different for the man and for the woman because of the linearity of the scale. Since the perception of pitch is conditioned by not only the fundamental frequency used by the speaker, but also the gender and age of the speaker (affecting vocal folds size) as well as factors of the transmission and the listener (auditory system), a psycho-acoustic scale should be used when studying pitch perception (see Nolan 2003). However, measurements on all of the above scales can be derived from the measurements in Hz, and the Hz scale is considered appropriate to use when the acoustic property of pitch is of interest, as it is in this study.

2.6.1.2 Loudness

Loudness is the perceptual correlate of intensity, which is proportional to the root mean square (RMS) amplitude of the sound wave (Ladefoged 1962, Laver 1994). The sound wave is the pressure fluctuation travelling from the original source through the acoustic medium (usually the atmosphere), and the average amplitude (calculated using the RMS method) of the wave is used to represent the magnitude of the sound pressure fluctuation (Johnson 2003, Hewlett & Beck 2006). The intensity of the speech sound is usually measured on the decibel (dB) scale, with 0 dB corresponding to the sound pressure level (SPL) of the threshold of hearing at 1,000 Hz. Below are some everyday examples of speech intensity:

- Shouting at a very close range: 100 dB
- Normal conversation: 70 dB
- Quiet conversation: 50 dB
- Soft whisper: 30 dB (Laver 1994).

These typical values also reflect the general expectation of average speech intensity appropriate to different environments. For example, in a quiet library where sounds are usually only as loud as 40 dB (Hewlett & Beck 2006), one might want to lower one’s voice to almost a whisper if he or she wants to talk without causing disturbance, but at a rock concert which can be as loud as 110 dB (Hewlett & Beck 2006), one probably needs to shout in order to be heard. These intensity values reflect the norms determined by particular environments, but speech intensity can be voluntarily controlled by the speaker for different functions in communication. Such communicative functions of speech intensity will be discussed in Section 2.6.2.

2.6.1.3 Tempo

Essential to the temporal organisation of speech is the concept of duration, which is the ‘amount of time taken up by a speech event’ (Laver 1994:431). The human auditory system is sensitive enough to detect very fine durational differences in speech, which can be as short as 10-40 milliseconds (Laver 1994). The linguistic function of durational features is to make length contrasts in speech segments (e.g. short vs. long vowels), and in paralinguistic communication the speech (including silence and hesitation) duration can indicate the psychological state or trait of the speaker. Further discussion of the functions of durational features as well as those of other prosodic features will be reserved for the next section.

Related to the notion of duration is the rate of speech. This is the speed at which utterances are produced. Two kinds of speech rates, namely articulation rate and speaking rate, can be distinguished. Articulation rate is the rate of speech sound production (i.e. vocalisation) in an individual utterance, including linguistic units (e.g. words) as well as non-linguistic speech sounds (e.g. the uh's and the um's, also known as filled pauses). Speaking rate describes the overall tempo of a multi-utterance speaking-turn, including all silent pauses that make up the speaking-turn. The unit of measurement for both rates can be syllables per second or words per minutes, though the latter is said to less accurate than the former (Hewlett & Beck 2006). More importantly, in order to have a meaningful comparison of speech tempo, a reference to a medium rate is needed. This, in English, can be a speaking rate of 5-5.5 syllables per second, although variation across different sociolinguistic communities is possible (Laver 1994). Such variation is, perhaps, particularly interesting when comparing native and non-native speakers, since a slower speaking rate of a highly proficient non-native speaker is unlikely to reflect his or her English fluency, but may be an indicator of his or her sociolinguistic background, instead.

2.6.2 The functions of prosodic features

2.6.2.1 Linguistic functions

Prosodic features not only carry out paralinguistic functions in speech to indicate speakers' attitudes, emotions or moods, but also perform linguistic functions in terms of their role in grammatical structures. This section begins with an overview of the linguistic functions of prosodic features.

Bearing in mind that the F0, intensity and duration of speech have their perceptual correlates as pitch, loudness and length, it will not be difficult to understand how prosodic features mark the stress (prominence) and the intonation (pitch pattern) in utterances. Due to their importance in terms of linguistic structure, the use of stress and intonation has traditionally been discussed extensively in English language teaching (ELT) and its related literature (Collins & Mees 2003, Cruttenden 2008).

a) Stress

Word stress or lexical stress is stress placed on the prominent syllable of a word, and sentence stress is the placement of the nucleus (see later in this section for further detail) in a sentence (Laver 1994). In most languages, pitch, loudness and duration function jointly to make stress (Laver 1994), but pitch is the most efficacious in giving syllables prominence in English (Cruttenden 1997).

Some of the world's languages, unlike English, have (predominantly) fixed word stress. For example, word stress falls on the initial syllable in Finnish, the penultimate in Spanish, and the final in French. French, however, loses word stress in connected speech and shifts the stress to the end of the sentence (Cruttenden 1997). There are two points that can be made about stress placement. Firstly, word stress can be more interesting to study in connected speech than in isolation, due to the potential effect of the word position in the sentence. Secondly, word stress in English might be a difficult area for non-native learners of the language, especially if their first language (L1) is a language with fixed word stress. Stress errors among non-native learners of English are therefore widespread and can cause embarrassment in communication or even affect intelligibility (Collins & Mees 2003, Hahn 2004). The resulted misunderstanding can have more serious consequences in a clinical communication context than in other communicative situations.

Furthermore, some languages (e.g. Greek) mark word stress in orthography. Although English does not do so, native speakers do not generally have much difficulty in guessing the stress of unfamiliar words (Collins & Mees 2003). This could suggest that there are underlying rules governing word stress in English, even though the rule system seems to have a large number of exceptions. Whilst accepting these exceptions, the English word stress rules can be determined by word classes, final syllable structures, suffixes etc. (Cruttenden 1997). The influence of a suffix on English word stress in particular, poses a challenge for non-native learners (Cruttenden, 2008). Furthermore, in addition to pitch, loudness and length, vowel quality can also play a role in English word stress, in that vowels in stressed syllables retain their full forms while those in unstressed syllables take reduced forms (e.g. schwa) instead (Laver, 1994, Cruttenden 1997). Therefore, English can differ from other languages in not only stress placement, but also vowel quality in stressed syllables. It is easy for non-native learners of English to have stress forms different from native patterns, and due to such differences, misunderstanding between native and non-native speakers can potentially occur.

Another notion related to stress is the rhythm perceived in a language, with Standard English being an example of a stress-timed language and French a syllable-timed language. Stress-timing is often characterised as equal interval between stressed syllables, and syllable-timing as equal syllable duration, although such idealisation can be rather different from the physical reality (see Laver 1994). That said, stress-timing and syllable-timing can be convenient terms when comparing the rhythmic characteristics across different English varieties. When compared to Standard English, the English varieties spoken in, for example, India, Singapore and parts of Africa can give a syllable-timed feeling. This feeling can make such non-native English varieties sound "exotic" to native listeners, and if this exotic feeling is not perceived favourably, the communication experience between non-native speakers and native listeners might be judged negatively. Some more details on the stress differences between native and non-native varieties are discussed in Section 2.6.5.

b) Intonation

Intonation, in a strict sense, is the modulation of pitch beyond the word level, which can also be called 'intonation proper' (Hirst & Di Cristo 1998:7) or 'melody' (Nolan 2006: 433), which is the definition adopted in this study. Unlike this definition, other authors may choose to use "intonation" to refer to a range of other prosodic features, in addition to the pitch activity in utterances (c.f. Nolan 2006). However, the aspects of intonation that concern this study are the pitch movements in speech that are purely melodic.

An obvious interaction between intonation and grammar can be observed in different sentence types. Languages seem to have a typological preference for signalling questions. When nearly 250 languages were investigated, 70% of them were found to use a rising terminal (an uptrending pitch movement at the sentence-final position) and 30% a higher overall pitch for questions (Bolinger 1978 cited in Hirst & Di Cristo 1998). When yes/no questions are pronounced without a final rise in languages that use rising terminal for questions by default, they are often perceived as having added connotations (e.g. surprise). In the case of WH-questions, a rising intonation in many languages, including English, can give a feeling of being friendly (Hirst & Di Cristo 1998). Collins & Mees (2003) gave an excellent example of some such questions when asked if different intonation patterns could create very different messages:

What's your name?
Where do you live?
What were you thinking of doing this evening?

Using falls to ask the above questions could sound a bit as though the police were interviewing a suspect, while using rises to ask the same questions could add friendliness and even sound like the first stage in asking someone out for a date (Colins & Mees 2003).

Before taking the discussion further, it would be helpful to have a convention by which to describe and analyse intonation patterns. There is, already, an established tradition used to analyse sentence intonation, especially those in British English, in terms of a set of nuclear tones. As explained earlier, pitch, together with length and loudness, can make syllables prominent in sentences. The nuclear tone is characterised by the pitch pattern of the most prominent syllable (the nucleus) in a sentence (Cruttenden 1997). Usually seven nuclear tones are identified in English, which are all summarised in Table 2.3 below:

Table 2.3 English nuclear tones[a]

Movement	Nuclear tone	Type
Falling	High fall	Simple
	Low fall	
	Rise-fall[b]	Complex
Rising	Fall-rise	
	High rise	Simple
	Low rise	
Level[c]	Mid-level	

a. See Cruttenden (1997 §4.4.1) for a comprehensive account of English nuclear tones
b. Not common; absent in some varieties (Collins & Mees 2003)
c. Mid-level is the most common variant; it can be seen as a form of rise (Cruttenden 1997)

Nuclear tones are the analysis convention adopted in this study when discussing the intonation patterns in sentences and are used because of their intuitiveness and their prevalence in literature on British English intonation.

Different nuclear tones can add to or modify the literal meanings in sentences, as Cruttenden (2008: 283) summarised:

(i) Falling nuclear tones are seperative, matter-of-fact and assertive; the higher the fall is, the more vigorous is the degree of finality involved.

(ii) Rising nuclear tones and level tones are continuative, implicative and non-assertive.

The above can be generalised even further, as in their earlier treatment in Cruttenden (1997):

Table 2.4 Overtones of English nuclear tones

Category	Nuclear tone	Overtone
'Serious'	High-fall	More 'involved'
	Low-fall	More 'business-like'
'Lighter'	High-rise	The most casual
	Low-rise	

Adapted from Cruttenden (1997:51)

What needs to be pointed out here is that these generalisations are variable according to their context, situation or personality. In fact, a polite intonation in one situation might sound offensive or patronising when used by another person or on another occasion (Cruttenden 2008).

The contextual sensitivity of the intonation overtones can be applied to the setting of professional communication, including that between doctors and patients. That is, in order to convey the message clearly and effectively, the choice of intonation needs to accord with verbal content and be appropriate to the context.

Another way of generalising meanings of English nuclear tones is in terms of their attitudinal markedness: some nuclear tones are generally attitude-neutral while others generally mark particular attitudes, as in Table 2.5 below.

Table 2.5 Attitudinal markedness of English nuclear tones

Neutral tones	high fall	low rise
Marked tones	fall-rise	rise-fall
Exaggerating tones	low fall	high rise

Adapted from Collins & Mees (2003:125)

As can be seen here, the complex tones fall-rise and rise-fall are usually reserved for attitudinally marked messages. The former possibly implies ‘doubt, correction, reservation, appealing to the listener to reconsider’, and the latter suggests ‘impressed, arrogant, confident, self-satisfied, mocking, putting down’ attitudes on behalf of the speaker (Collins & Mees 2003:125).

It can be said further that English sentence types have their ‘default’ and ‘alternative’ intonation patterns. In other words, the attitudinal meanings of intonation do not function independently, but interact with grammar (Cruttenden 2008). The following (Table 2.6) is a summary of the interaction between the syntactic sentence types and the attitudinal meanings conveyed in intonation.

Table 2.6 Default and alternative English intonation patterns

Utterance type	**Default pattern**	**Other patterns**
Statements	Fall	(1) Rise (adds non-finality or questioning) (2) Fall-rise (adds non-finality with an additional but unspoken message)
Commands	Fall	Rise (turns command into request)
Wh-questions	Fall	Rise (adds warmth, interest)
Yes-no questions	Rise	Fall (turns question into exclamation)

Reproduced from Collins & Mees (2003:129)

As can be seen here, nuclear tone analysis can be understood as resting upon a set of two rules about intonation. Firstly, a language only accepts a limited set of basic intonation patterns (Nooteboom 1997), which is true even for a language like English that uses intonation elaborately and flexibly. Secondly, some intonation patterns are more marked (attitudinally) that the others, and ideally using the marked patterns should only be on a "when needed" basis in communication.

These principles, in turn, have multiple implications for non-native speakers of English. Since intonation patterns are often used in conjunction with verbal meanings, when subtle attitudinal information needs to be conveyed, using intonation patterns in effective communication can be understandably complex. Non-native speakers may be influenced by their native languages and produce English intonation patterns that are very different from native English varieties. Due to the contextual sensitivity of the attitudinal information conveyed in intonation mentioned above, when the attitudinal interpretation by native listeners differs from their expectations in that context, misunderstanding can occur as a result. However, such potential misunderstanding, due to inappropriate intonation, is not limited to non-native speakers and can happen to those whose first language is English as well. Anybody, regardless of their English language background, may occasionally produce intonation patterns that are not expected in the communicative context and thus create the chance for misunderstanding. In a medical context, when questions are asked to gather information, if a doctor uses an intonation that can have an arrogant (e.g. low fall used) or sarcastic (e.g. rise-fall used) overtone, or fails to use intonations to convey warmth and interest, the patient may perceive this as expressing indifference. In addition to marking sentence types and conveying attitudinal meanings via interaction with sentence types, intonation is also used to highlight information and to regulate turn-taking in speech (Nolan 2006). If any of these roles fails to be carried out as expected, communication may potentially be frustrated or even deteriorate into miscommunication.

c) Lexical tone

Intonation exists in all languages, though some languages, unlike English, make further use of pitch as a means to distinguish word meanings (see Laver 1994, Cruttenden 1997). The product of such pitch use is the lexical tone, and languages that exploit lexical tones in their sound systems are known as tone languages. Some tone languages such as Yoruba use pitch levels (e.g. high, mid, or low) to distinguish between word meanings, and these are called level/register tones. Others, such as Mandarin Chinese, rely on pitch contours (e.g. level, rise or fall) for word meanings, which can be referred to as contour tones (Laver 1994, Gussenhoven 2004). Unique to tone languages, also, is the sometimes complex interaction of intonation and tone systems. For example, in Thai, the intonation patterns are superimposed on the syllable tones (Laver 1994). In Mandarin Chinese, however, the language's intonation patterns seem to influence the lexical tones on the sentence-final syllables more than on the other syllables (Yuan et al. 2002). Therefore, non-native learners of English whose native languages are tonal may face an additional challenge when attempting to master English intonation patterns, due to a likely influence from the native phonological systems (systems of the structured organisation of speech sounds).

2.6.2.2 Paralinguistic functions

Paralinguistic behaviour can be defined as the coded communication of the speaker's affect, attitudes, emotions and turn-taking that can be non-verbal and non-linguistic, and much of this is conveyed through tone of voice (Laver 1994). It is precisely the paralinguistic functions of prosodic features responsible for communicating a speaker's tone of voice that are the focus of this study. This section synthesises previous research on the paralinguistic functions of prosodic features with a view to consider the implication of the evidence on the aims of this study

A clarification is needed, here, of what non-verbal communication refers to in this study. According to Laver (1994), verbal communication is limited to the use of spoken words, while non-verbal behaviour can incorporate any features of communication other than words. Therefore, non-verbal communication can be through both vocal and non-vocal channels (see also DePaulo & Friedman 1998). From this perspective, non-verbal vocal behaviour can include any prosodic features not serving linguistic functions. For example, word stress would be classified as verbal, but tone of voice is a non-verbal matter. Examples of non-verbal, non-vocal communicative behaviour include gesturing, posturing eye-contact and facial expressions. Although the prosodic features of interest in this study are largely not verbal behaviour, when discussing their functions in communicating information, such as attitudes, the term 'paralinguistic' is preferred rather than the term 'non-verbal', which is reserved, in this study, for those features of communication that are purely non-vocal (e.g. body-language).

Some paralinguistic functions of prosodic features are culturally universal. For example, tenderness in most cultures is marked with a quieter (lower intensity) than normal voice. Others such as tone of voice through intonation (F0 modulation) can interact with linguistic functions, as pointed out above, and can be language- or culture-specific (Laver 1994). The summary below will focus on the paralinguistic functions of prosodic features that can be more or less generalised across cultures, and the individual features will be discussed first before the evidence is synthesised. Cultural and linguistic specificities in the paralinguistic use of prosody will be introduced when prosodic perception is discussed.

Intensity manipulation can convey a range of affective information such as enjoyment, fear or rage (Banse & Scherer 1996, Scherer et al. 2003). Within this paralinguistic role of intensity exist several idiosyncratic differences, that is, different speakers can have different intensity settings when expressing the same emotion (Pfizinger & Kaernbach 2008, Spackman et al. 2009). However, emotion seems to have a greater effect on speech intensity than speaker idiosyncrasy in a study by Pfizinger & Kaernbach (2008) who attempted to explain the inter-speaker variation by the fact that individuals have different personalities, after carefully controlling the intended emotions and the recording setup.

It needs to be pointed out that multiple acoustic cues (including intensity features) might be at the disposal of a speaker when expressing a specific emotion (Pfizinger & Kaernbach 2008, Spackman et al. 2009). Spackman et al. (2009) adopted a theoretical model (see Section 2.6.4 for a discussion on the Brunswikian lens model) that would not rely on the existence of a one-to-one match between emotions and singular acoustic profiles, so that multiple acoustic cues can be studied at one time. Spackman et al.'s (2009) theoretical stance did not assume a conception of discrete emotions (see Scherer 2003, Scherer et al. 2003 for a comparison between discrete emotion theories and appraisal theories of vocal expression of emotion) either, and therefore the study design did not have to rely on the encoding of emotion to be pure and error-free.

F0 (perceived by listeners as pitch) features have also been found to have a strong association with vocally portrayed emotions. In their study, Banse & Scherer (1996) found that emotion can explain a large proportion of variance in the F0 measures, after adjusting for speaker attributes (gender and speaker identification) and linguistic information (sentence type). Based on this finding, Banse & Scherer (1996) suggested that predictions could be made on the F0 profiles for accompanying emotions (e.g. increased mean F0 for intense emotions such as panic fear). Therefore, in speech communication, a listener can expect to hear specific pitch configurations when certain emotions of the speaker are being presented.

It is important to remember that inter-speaker variation does not invalidate the F0 profile for emotion, similar to the case of intensity variation reported in Spackman et al. (2009). The same authors also found emotion by speaker interaction in their modelling of F0 features in emotional portrayals which did not affect the accuracy of listeners' identification of emotions (Spackman et al. 2009). Such findings are interesting since they could suggest the F0 profiles for emotional expression are a result of speaker appraisal of the emotional need in the communication event rather than direct specification by a categorical set of emotions (see also Scherer et al. 2003 on the discrete emotion theory vs. the appraisal theory). The importance of this information to the current study is therefore that in the health communication context, instead of determining the specific emotional meanings of differing prosodic, including F0 features, focus should rather be placed on a range of aspects of professional attitudes for which the evaluation could be influenced by prosodic features.

Banse & Scherer (1996) and Spackman et al. (2009) also reported a close relationship between speech tempo features and encoded emotions. In particular, observed temporal profiles for vocal emotional portrayals by Banse & Scherer (1996) confirmed their predictions in most cases. For example, sadness is associated with a decreased rate of speech, and several intense emotions are characterised by an increased rate of speech (Banse & Scherer 1996, see also Laver 1994). An analogy can be made comparing the relationship between a speaker's rate of speech and emotional expression with tempo and emotion in music, where sadness is reliably marked with a slow tempo and high arousal emotions, such as fear, with a rapid tempo (Krumhansl 2002). This analogy in music suggests that a (near) universal link between the rate of speech and the expression of emotion might also exist.

The above are simply examples of previous research that relate to the prosodic features of interest in the current study. Among them, Banse & Scherer (1996) is a comprehensive study including multiple members of several emotion families and a wide range of acoustic properties based on evidence in encoding and decoding. Accumulated evidence from Banse & Scherer (1996) and other studies (e.g. Johnstone & Scherer 2000) was synthesised in Scherer et al. (2003), which gave a rather full picture of the acoustic profiles of some basic emotions in speech communication. Only a selection of the global acoustic properties in the synthesis by Scherer et al. (2003) is investigated in this study, and this is reproduced in Table 2.7 below:

Table 2.7 Effect of emotion on prosodic features in speech

Perceptual correlates	**Acoustic parameters**	**Arousal/stress**	**Happiness/elation**	**Anger/rage**	**Sadness**	**Fear/panic**	**Boredom**
Pitch	F0 mean	>	>	>	<	>	<=
	F0 range	>	>	>	<	<>	<=
Loudness	Intensity (dB) mean	>	>=	>	<=	>[a]	<=
Rate	Syllables per second	>	>=	<>	<	>	<

Adapted from Scherer et al. (2003:439)

Note. > : increased compared to non-emotional speech, < : decreased compared to non-emotional speech, = : not different from non-emotional speech

a. Not included in Scherer et al. (2003); included in Scherer (2003), in which the unit of measurement is not specified

Table 2.7 shows a clear and expectable pattern regarding the acoustic profiles of emotions: high-arousal emotions such as stress, elation and rage tend to have a high F0 mean, high F0 range, high intensity and an increased rate of speech, while sadness and boredom usually possess the reverse acoustic values in speech. Therefore the prosodic features investigated in this study can all signal, if not mark, basic emotions. On a side note, emotions such as sadness or fear are called basic because there is said to be evidence for universal, particularly facial, expressions to distinguish them (Ekman 1992). Furthermore, the paralinguistic communicative functions of prosodic features are not limited to the expression of basic emotions.

Most of the reviewed studies in this section were carried out using acted portrayals of basic emotions such as anger and joy. These emotions are so basic that one can claim that even untrained speakers should be able to express them vocally, consciously or unconsciously, with some degree of faithfulness (Scherer 2003). Indeed, in the study by Spackman et al. (2009), training did not result in a clear advantage in terms of the identification accuracy of the vocal portrayals of emotions. Furthermore, although acted emotions resemble naturally occurring emotions closely enough in terms of recognisability, the former is likely to be more intense, more exaggerated, or more stereotypical than the latter (Scherer 2003, Pfizinger & Kaernbach 2008). Although the validity of using acted, simple, basic emotions in research has already been established in previous studies, they may lack the complexity and context of naturally occurring emotions in speech (Scherer 2003). An emotional speaker in a real life situation might as well adapt their expressive features according to their communicative environment (Pfizinger & Kaernbach 2008). Therefore it is important to conduct similar research in a more realistic interactional setting in order to obtain empirical evidence for the emotional or altitudinal role of intensity in communication.

The paralinguistic information investigated in the current study does not include basic emotions, but focuses instead on aspects of professional attitudes (see Section 2.1.2) relevant to doctor-patient communication. Differences exist between attitudes and emotions, although they share certain core psychological processes. This study adopts a ‘conventional’ (Albarrac ń et al. 2005:4) definition of attitude: ‘attitude is a psychological tendency that is expressed by evaluating a particular entity with some degree of favor [sic] or disfavor [sic]' (Eagly & Chaiken 1998:269). On one level, it can be said that attitudes, being affective as well as cognitive and behavioural (Eagly & Chaiken 1998), belong to a more complex construct and thus can incorporate emotions as a component. On a deeper level, attitudes and emotions share the similarity in that they both have an evaluative focus, but their difference lies in the temporal constraint. In other words, an emotion is an evaluative state but an attitude is an evaluative tendency: an emotion can only last as long as one is in that particular state while an attitude can be enduring and does not have to disappear when one no longer thinks about the attitude object (Clore & Schall 2005). It can be further said that despite their difference in terms of temporal constraint, a study of the vocal expression of attitudes can borrow from the type of methodology used in studying the vocal expression of other affective constructs, such as emotions, because they are both evaluative activities and they can both be expressed in shared channels. This study therefore finds common ground with previous psychological research on vocal emotional expression. Since attitudinal display and emotional expression are similar evaluative processes using shared perceptual cues including prosodic features, research on them can utilise such similar study designs leading to comparable findings. Before such methodological issues are addressed in Chapter 3, prosodic features as perceptual cues are discussed in Section 2.6.3 below.

2.6.3 The perception of prosodic features

Human perception appears to be very sensitive to prosodic features. For example, the perceptual threshold for pitch difference in speech was reported to be as low as 1.5 semitones (Rietveld & Gussenhoven 1985 cited in Nooteboom 1997). However, when comparing the distance of pitch movements cross-linguistically, languages seem to demonstrate some size differences. British English would need 12 semitones (a logarithmic measure of pitch, see Glossary) to characterise its pitch movements, but German would only need 7.5 semitones, and Dutch, 6 semitones (Nooteboom 1997). This, perhaps, can explain the stereotypical British impression that the German language has a flatter pitch pattern (see Mennen et al. 2007).

Aside from cultural-linguistic factors, cross-linguistic difference in prosodic perception might have more fundamental origins in the neurological processes.

There is evidence from brain scanning studies showing that the right hemisphere is primarily responsible for the auditory processing of complex pitch signals, including those relating to speech, while the left hemisphere can be involved when the categorical processing of phonological, syntactic or semantic information in language is needed for communication (Gandour et al. 2004). Such evidence seems also to support the claim by Seddoh (2002) that intonation can be modified by emotional processes. In other words, linguistic and emotional processes can interact within the brain, and the resulting speech products (prosodic features) should be viewed not as separate, opposing entities (i.e. "linguistic prosody" vs. "emotional prosody") but as the same set of features for different communication purposes. Furthermore, the lateralisation occurring in the left hemisphere, when the brain processes speech signals (including prosody) beyond the auditory level, depends on language experience (Gandour et al. 2004). Consequently, limited language experience, for example, in the case of non-native learners, might contribute to difficulty with speech prosody on a neurological basis. In fact, due to linguistic typological differences, native speaker of tone-languages (e.g. Mandarin Chinese) may be less familiar than native speakers of intonation languages (e.g. English) when it comes to using intonation to express emotions (Seddoh et al. 2002).

2.6.4 Prosodic features as perceptual cues to speaker affects

Due to their paralinguistic functions introduced earlier on in this chapter, prosodic features serve as important cues in communication for listeners, in particular, being used to infer the attitudes and emotions of speakers. As introduced in Section 2.5.3, some such prosodic cues to affective perception can be universal, that is, listeners can decode emotions in even unfamiliar languages through vocal features at a level above chance (Thompson & Balkwill 2006). This is possibly thanks to the capability of the brain to use acoustic properties (such as F0 and intensity) as cues to speaker affects. Like non-verbal features (the non-verbal proper, e.g. facial expression), these cues can operate independently of the verbal content of speech. However, the evidence that listeners decode affective information with the greatest accuracy when hearing their own language, suggests an in-group advantage due to language experience (Thompson & Balkwill 2006). That is, some prosodic cues to speaker affects cannot be regarded as purely acoustic; such prosodic features are designed to work in conjunction with linguistic structures. As paralinguistic features, therefore, unlike the non-verbal proper, prosodic features can interact with or modify the verbal meanings present in speech (Laver 1994). In doctor-patient communication, prosodic features and non-verbal proper features can have different places in bringing forth professionalism and patient-centredness during clinical encounters. Non-verbal features such as body language and eye contact can immediately convey attitudinal messages in a medical interview (Hall et al. 1995). For example, a gentle nod while maintaining eye-contact with a patient during the taking of their medical history can acknowledge the patient's concerns and encourage them to explore their own perspective on their condition, even without words being spoken. However, prosodic features work in conjunction with the verbal content of a doctor's utterances (Hall et al. 1995). In the same example of history taking, a slow speaking rate with pitch and, possibly, intensity kept low can be appropriate in uttering an 'accepting response,' performed to acknowledge and empathise with the patient's sadness. Therefore non-verbal proper features and prosodic features in medical communication can benefit from separate studies designed to answer different research questions. This study explores the role of prosodic features in conveying attitudes related to professionalism in clinical interactions, while leaving any questions on non-verbal proper features to further studies.

Importantly, the linguistic significance of prosodic features cannot be separated from their paralinguistic communicative value, especially in the case of pitch phenomena (Laver 1994). Take English declarative intonation, for example, for a same statement said in the falling intonation (default pattern, see Table 3.5), the lower the pitch level and the narrower the pitch range, the more distance is created between the interlocutors. Therefore the perception of speaker affects encoded in prosodic features can sometimes overlap with the perception of the linguistic function served by the same prosodic features. Furthermore, multiple types of affective information can be perceived through speech prosody, including speaker attributes, emotions and moods (Gobl & N í Chasaide 2003). Perception of these attributes can involve processes relating to the interaction between the linguistic and paralinguistic functions of prosody. For example, perceiving an overtone in a sentence one has heard is more likely to be due from a coordination between the sentence type and intonation rather than entirely from intonation (Collins & Mees 2003). Other prosodic features, such as overall F0 mean and range, alone, can influence affective judgments and are therefore completely paralinguistic (Scherer et al. 1984). For this study, then, it is clear that the design needs to incorporate both differing linguistic categories (e.g. command vs. question) and a range of paralinguistic meanings possible in clinical interactions.

Voice quality (e.g. whispery, breathy, harsh etc.) achieved by varying phonation settings (Laver 1994) is a further example of how different types of speaker affects can possibly be perceived through different prosodic features. Gobl & N í Chasaide (2003) found that voice quality (e.g. whispery, breathy, harsh etc.) alone can change the affective overtone of an utterance. However, the authors also pointed out that voice quality differences appeared to be more effective in signalling milder affective states, while large F0 excursions might be more responsible for cuing stronger emotions. One-to-one mapping between voice quality and speaker affect was concluded to be unlikely, and individual voice qualities could be associated with a cluster of affective attributes, instead (Gobl & N í Chasaide 2003). Such findings imply that the study of affective perception through prosodic cues can be more likely to benefit from listener judgment on multi-item scales than from categorical identification/recognition of affective attributes. This therefore must be considered when deciding on the instrument to be adopted in the current study to examine attitudinal perception.

Another implication is also methodological in nature. Ladd et al. (1985) suggested that the paralinguistic/acoustic use of prosodic features reflected states of arousal (strong and gradient affects), while the choice of intonation types signal distinct cognitive attributes. Consequently, to study the functions of prosodic features from a perceptual perspective, one needs to first determine whether to focus on the affective meanings or on the cognitive significance. Only after that could perceptual measures be dictated. Such methodological considerations will be discussed in more detail in Chapter 3.

It is beneficial to make reference to a theoretical framework in the discussion of prosodic features as cues to perception. One linguistically oriented theory to explain the complexity of prosodic features as perceptual cues to speaker affects is the

Biological Codes theory (Gussenhoven 2002). In this framework, languages share some universal form-function relations regarding interpreting affective meanings in intonation through three biological codes: the Frequency Code (pitch height), the Effort Code (pitch excursion) and the Production Code (pitch declination). The origin of these codes can be traced to the universal aspects of the speech production mechanism, which affect the rate of vibration of the vocal folds under articulatory control (phonetic implementation). However, during the process of grammaticalisation in which these codes become part of the phonological/grammatical structure, the original form-function associations might not be retained and might be subject to the force of phonological change. Such a framework of biological codes is therefore capable of accounting for the universal use of prosodic cues to speaker affects across languages, as well as for certain cross-linguistic or cross-cultural differences in the paralinguistic functions of prosodic features. The influence of the framework can be seen in this study in that the design involves the comparison of the use of prosodic features by native and non-native speakers in a clinical context, but when discussing the paralinguistic meanings perceived in their speech, their English language backgrounds can be accounted for.

Another theoretical framework that Scherer (2003) suggested adopting is a modified Brunswikian lens model. The lens model can serve as a basis of non-verbal communication research paradigms that aim to analyse both the perception and the expression sides (DePaulo & Friedman 1998). These research paradigms are possible because the lens model is capable of distinguishing between the expression (encoding) of affective information on the sender side, the transmission of the sound signals (through space and psycho-acoustically transformed in human perception), and the impression (decoding) on the receiver side (Scherer et al. 2003). For studying vocal expression of emotion, the lens model is able to determine what acoustic cues the perceiver uses to infer the sender's state from (Spackman et al. 2009).

The modified Brunswikian lens model, according to Scherer (1982), "specifies that a particular trait or state of a sender is externalized or expressed in distal indicator cues, that is, characteristics of the sender's non-verbal behavior [sic] that can be objectively measured. These distal cues are perceived by an observer and represented as proximal percepts, the observer arrives at an attribution about the sender's traits or states (which the observer assumes, are reflected by the distal indicator cues) on the basis of a set of inference rules" (Scherer 1982: 141). The advantage of the lens model for non-verbal communication research is that both the non-verbal behaviour of the sender (as perceptual cues) and the perceiver's judgments based on these cues can be independently measured and the association between them established. The implication of the lens model for this current study is, therefore, that the design needs to be capable of measuring prosodic features (communication behaviour) and attitudinal ratings independently so that the relationship can be examined.

It can be seen from the perceptual mechanism, the empirical evidence and the theoretical modelling that while the perception of affective meanings in speech communication relies largely on prosodic features, the perceptual process can lead to unintended results. When inaccuracy occurs, miscommunication can be a potential consequence. This can be due to several factors described in the Brunswikian lens model, one of them being the distal cues at the sender end, which can either be absent, or inadequate in their externalisation. Non-native speakers of a particular language, for example, can contribute to the inaccuracy in distal cues due to their limited experience in their non-native language: they can be perfectly competent in executing the linguistic functions of it, but underperform in its paralinguistic expression, according to what might be expected by a native speaker of that language. . The current study is carried out with highly proficient non-native speakers of English. Therefore the focus is on the paralinguistic information pertinent to the clinical settings.

2.6.5 Characteristics of non-native English prosody

It is perhaps not surprising that non-native speakers often exhibit prosodic features deviating from native patterns, due to the primal significance of prosody in speech and the complexity of the prosodic structure and functions. English is a language in which many such examples can be made. Aside from the differences in vowels and consonants between native and non-native English, the rhythmic characteristics of some African English varieties and the melodic characteristics of some Indian accents, for instance, are immediately distinctive. Furthermore, non-native English generally has a lower speaking rate than native English (Yuan et al. 2006). One might think that these characteristics could not affect communication greatly, apart from sounding exotic (see Anderson-Hsieh et al. 1992). However, evidence in non-native English prosody research, especially in an ETL context, suggests otherwise.

The difference in English prosodic features between native and non-native varieties can be observed in form as well as in function. In terms of the form of prosodic features, for example, Mandarin speakers of English were found to be able to produce intelligible stressed vs. unstressed contrasts at a percentage similar to that of American native speakers of English (Zhang et al. 2008). Although similar acoustic parameters (higher F0, greater intensity and longer duration) were exploited by both speaker groups for stressed syllables, Zhang et al. (2008) found that these were exploited in different ways. The acoustic analysis of word stress production of 10 speakers from each group revealed these following details. The Mandarin speakers of English were reported to produce significantly higher F0, but 2dB lower intensity, in the stressed syllables than the native speakers of English, and differences between the groups were also found in F0 peak alignment and vowel reduction when making stressed vs. unstressed contrasts. In terms of the function of prosodic features, this contributed to the acceptability of the non-native pronunciation being deemed to be much lower by native speakers than that of native pronunciation (Zhang et al. 2008). Acceptability is not strictly a measure of communication success if the exchange of information is the only concern. However, it can be an indicator of listener experience, and according to the lens model (see Section 2.6.4), represents an attribution about the speaker resulting from the listener's perception. Of course, this can be complicated by other factors in pronunciation, (e.g. foreign accents in vowel and consonant pronunciation, though non-native prosodic features, can also contribute to the judgement of foreign accents). The research design therefore needs to incorporate consideration to control for such influence, if foreign accents are not to concern the research question. This can be economically achieved by recruiting only non-native speakers whose foreign accents are very mild and not immediately recognisable. Such control can be reinforced by designing the outcome measure so as not to include direct judgements on the acceptability of pronunciation.

The non-native use of English intonation can also differ from the native norm (Collins & Mees 2003, Wennerstrom 1994, Verdugo 2005). Wennerstrom (1994) reported that the non-native (Japanese, Spanish and Thai) speakers of English in the study failed to use pitch accent in order to signal new-old information contrast, a main function of intonation in a native-like fashion. The Japanese and Thai speakers of English were also found, in a description task, to use a lower pitch at phrase boundaries than the Spanish and native speakers of English. Such differences, especially the latter one, could be due to linguistic typology (Spanish is closer to English in terms of its intonation structures than is Japanese or Thai) and language experience (the Spanish speakers had longer exposure) (Wennerstrom 1994). The findings in Wennerstrom (1994) could indicate that non-native speakers of English with a very different typological background (e.g. tone languages) might have more potential difficulty with English prosodic features and might need more training to reach a native-like mastery (e.g. through longer exposure)of them. It will also be interesting to know if highly proficient non-native speakers of English, like the ones included in the current study, differ from native speakers in terms of their prosodic features.

Functional consequence of non-native English prosody can also be observed in the experience of the listeners. This can be the perceived intelligibility (Munro & Derwing 1998, Hahn 2004, Zhang et al. 2008). London Jamaican English, for instance, was said to be sometimes difficult to understand, due to its distinctive use of pitch, loudness and rhythm (Local et al. 1985). The consequence can also be in terms of the paralinguistic function of prosody. For example, Verdugo (2005) found that Spanish speakers of English and native speakers expressed the contrast between certainty and uncertainty through the use of prosodic features differently. In the study, prosodic cues used by the native speakers for conveying a sense of certainty were high-falling intonation, as opposed to the use of falling-rising or low-rising patterns for ‘vulnerability, reservation [and] indefiniteness’. Compared to this, the non-native patterns adopted by the Spanish speakers were obviously different: their frequent use of mid-level intonation was not adequate to convey the intended sense of certainty (Verdugo 2005: 2109). The contrast between certainty and uncertainty is just one (modality) of three types of meanings (mood e.g. statement vs. question, modality e.g. the possibility or probability of a proposition, and key e.g. politeness, assertiveness etc.) related to the tone (i.e. the nuclear tone choice) according to the systematic-functional approach to intonation (see Halliday & Greaves 2008). Therefore all mood, modality and key meanings conveyed through the intonation patterns can be a source of misunderstanding, with the former two types more likely to be linked to potential confusion, and the latter one linked to the undermining of communication or even the causing of offence. This study of doctor-patient communication should, consequently, consider intonational meanings because they can be related to patient experience.

Therefore it can be seen that non-native English varieties can have unique prosodic characteristics that distinguish them from native ones. Although sometimes such characteristics only inspire amusement (Collins & Mees 2003), often they can be potential barriers to communication especially when conveying paralinguistic information between native and non-native speakers of English (Wennerstrom 1994). It does not take much imagination to think of such consequences in the communication in sensitive contexts where even small frustration can develop into major misunderstandings. This is in addition to the possibility that, in their communication with patients, some doctors might actually have unprofessional attitudes which result, consciously or unconsciously, in certain prosodic features of their speech. In either case (misunderstanding or unprofessional communication), patient-cetredness is compromised, causing the quality of care to fall below the expected standard.

2.7 Summary and research questions

Existing knowledge has already been accumulated on the research topics of speech prosody and health communication. Previously, however, speech prosody as a channel of attitudinal communication has not been studied with empirical evidence in the healthcare context, nor has medical communication training addressed the use of prosodic features when communicating with patients in adequate depth. However, existing evidence strongly suggests that the use of prosodic features directly influences listeners' perception of attitudes. Therefore the patients' experience can also be influenced by the doctors' prosody. Furthermore linguistic research literature seems to agree that non-native English prosody can differ significantly from the native norm. Given the increasing globalisation of UK health care and medical education, it is imperative to investigate the potential link between speech prosody and patient experience with evidence from both native and non-native English speaking professionals (in practice or in training).

Therefore three general research questions are asked in the current study:

1. Do medical students use speech prosody to convey professional attitudes in medical communication?
2. Are there any differences between native and non-native English medical students in the use of prosodic features in the simulated clinical interaction?
3. Do prosodic features influence the listeners' perception of the speakers' professional attitudes in medical communication?

These, then, all jointly answer the overarching question as to whether or not speech prosody matters in health communication in terms of expressing professionalism with evidence from both speaker groups. Methodological considerations, and the subsequently adopted general research design, in order to answer these questions, are discussed in Chapter 3. The specific methods utilised in each stage of this study are described accordingly in Chapters 4 and 5, in which the results of the investigations are also reported.

Chapter 3 Methodological considerations and research design

This research integrates at least two research disciplines, namely health communication and speech prosody. Therefore it is necessary to consider the research methods commonly used in these disciplines in order to design a robust methodology for the current study. The chosen methodology must be able to answer the overarching research question raised in Section 2.7, in the previous chapter.

This chapter will introduce the established research methods in the disciplines of health communication and speech prosody, respectively, before discussing the adopted research design used for this study.

3.1 Research methods in health communication

As reviewed earlier, the aims of previous health communication research vary from identifying the nature and subject of patient complaints (e.g. Wofford et al. 2004) to evaluating the effectiveness of communication skills training (e.g. Parry 2008) and analysing strong OSCE candidates' communication behaviours (Roberts et al. 2003). As such, a range of methodologies have been employed and are discussed below.

3.1.1 Communication behaviour as a measure of communication skills

One of the important topics in health communication is communication skills. However, measuring communication skills directly may be difficult and, on some occasions, not immediately possible (Griffith et al. 2003, Brinkman et al. 2006). As defined by the Oxford English Dictionary, a skill is one's 'capability of accomplishing something with precision and certainty', or 'practical knowledge in combination with ability', or 'an ability to perform a function, acquired or learnt with practice'. Skills, therefore, have an inherently practical side and are closely related to performance, and as a result many skills (e.g. language skills, driving skills etc.) are often assessed through performance. Communication skills in health care are no exception, in this sense.

Self-reporting questionnaires are sometimes used as a measure of communication skills and/or the behaviours that result from it. Several studies in Parry's (2008) systematic review on the effects of communication skills training among allied health professionals relied either entirely or in part on self-reports as measures of improvement. However, as Parry (2008) pointed out, the limitation of such methods of measuring communicative abilities lies in their inability to measure long-term improvement in real life behaviour. In other words, their reliability can be questioned in that one will not always do exactly what is reported on a self-administered questionnaire.

Communication skills can be more reliably assessed through observed communication behaviours. In fact, measuring communication behaviours in real life situations is argued to be the gold standard (Parry 2008) of methods for gauging communicative proficiency, due to its apparent validity, though several challenges remain for using it effectively. The first and main challenge is that it is not possible to have a standardised measuring environment in real life, so some target behaviours can be missing due to the uncontrolled environment and not the candidate's actual skills. The second, related challenge is that each clinical encounter in real life is unique, and communication behaviours need to be adjusted accordingly, making it difficult to provide accurate measurements (Parry 2008).

One solution to these challenges is to use simulated patients and scenarios for measuring communication behaviours in clinical settings, so that the encounters to be evaluated can be standardised (Klamen & Williams 2006). Deladisma et al. (2007) even extended the method of standardising clinical encounters for medical students to interviewing a life-size virtual patient in a teaching setting. The authors found that the medical students studied demonstrated non-verbal communication behaviours and empathetic responses to the virtual patient, though both the amount and quality of these were not as high as that shown to a simulated patient (Deladisma et al. 2007). Communication behaviours observed in simulated clinical interactions can therefore reliably be used in the study of communication skills.

Measuring the presence or absence of particular communication behaviours in real life might not be the most helpful for health communication research, unless a checklist including a set of appropriate behaviours can be pre-established for the purpose (Epstein et al. 2005). Compared to the presence/absence measures, a measuring of the appropriateness of the observed behaviours in the context of a particular clinical encounter may be more informative, although appropriateness measures have been the less commonly used in previous research (Parry 2008). In order to have more meaningful results, it is important that further studies should adopt appropriateness as a measure by which to judge an individual’s competency (or lack thereof) at communicating effectively.

While including appropriateness of behaviour measures can be beneficial to health communication skills research (Parry 2008), a single measure of appropriateness can also be highly subjective. The challenge of achieving objectivity can be overcome by measuring the amount or quality of the communication behaviours and the satisfaction of the service users on separate scales. When both observed behaviours of healthcare professionals/trainees and perception/satisfaction of patients (whether real or simulated) are measured, correlation between the two can then be analysed. In a cross-sectional study including 27 doctors and 257 patients, Ruiz-Moral et al. (2006) identified a significant correlation between communication behaviours and patient satisfaction. Furthermore, their study demonstrated that the consultations rated as satisfactory by patients were characterised by patient-centred communication, although the doctors’ use of patient-centred skills was found to be scarce (Ruiz-Moral et al. 2006). However, the tool used by Ruiz-Moral et al. (2006) only recorded the presence or absence of the communication behaviours under investigation but not the degree of these behaviours, making it difficult to determine to what extent the communication behaviours can predict patient satisfaction. Also by measuring both observed non-verbal communication behaviours and patient satisfaction, Griffith et al. (2003) were able to explain the proportion of variance in (simulated) patients’ satisfaction predicted by the communication behaviours of internal medicine residents, which ranged from 19% to 32% depending on the scenario of the encounter. In this way, the regression result can be seen as an indirect measure of appropriateness of the behaviours. Thus, studying professional communication behaviour alongside perception by service users can have added potential of correlational findings. The appropriateness of communication behaviours in patient-centred care as discussed here is largely indexed by patient satisfaction. Alternatively, appropriateness can be defined by other criteria such as the comprehension of information, depending on the purpose of research. From the perspective of patients' perception of professionalism, especially patient-centredness, patient satisfaction is an important and valid component of the measures.

Though there exist multiple approaches to medical communication skills training, for an objective method of evaluation that is readily applicable to practice, the communication behaviours of the trainees need to be measured, for example, in an OSCE-like manner. Therefore the communication behaviour or performance can be seen as a measure of communication skills. The communication behaviour can either be measured by the presence/absence/amount or by its appropriateness to the context judged by real or simulated patients or external raters. In addition, the observed communication behaviour can also be examined as an indicator of patient satisfaction with the communicative event.

3.1.2 Measures of attitudes

Similarly to the level or the improvement of communication skills, the strength or the development of attitudes in healthcare research can be measured through self-reports, some of which can be rating scales (see e.g. Cushing et al. 2005). Using self-report measures to investigate attitudes is quite established in the literature. Anvik et al. (2008) investigated both cognitive and affective attitudes (see the previous chapter on the anatomy of attitude) towards communication skills on a self-administered scale. Woloschuk et al. (2004) also used a similar way to measure attitudes towards social factors and skills in medicine and the changes in these attitudes among medical students. The attitudes focused on in these studies are, therefore, the underlying beliefs and feelings. However, attitude, as discussed in Section 2.4.2, can also be reflected in as well as influence behaviour. Consequently, attitudes evaluated through self-reports in these studies are not behavioural in nature. Although self-reporting scales are useful as a measure of the affective and cognitive elements of attitude, and are easy to administer, they are not sufficient to measure the behavioural element of attitude. In other words, the aspects of attitude reflected in behaviour which can be perceived by observers need to be measured using scales of a different design. Since the current study focuses on those attitudes directly relating to patient experience, the measurement tool needs to be able to capture the attitudinal meanings conveyed to the patient in communication.

As Martin et al. (2002) commented, attitudinal measurement is, however, inherently difficult, and is best served by a multidimensional approach. In addition, the evaluation of professional attitudes is not necessarily restricted to formal assessment or rating only. In order to understand medical students' attitudes towards communication skills learning, Rees & Garrud (2001) conducted a focus group with five first-year medical students. This allowed an in-depth exploration of the students' ideas about communication skills and communication training, as well as their expectations, feelings and beliefs regarding such matters. As a result, both positive and negative attitudes towards medical communication skills learning were identified, together with the relationships between these attitudes and the students' backgrounds. Rees & Garrud (2001) used this method in their exploratory research to achieve a detailed understanding of the themes present in the medical students' attitudes towards their communication skills training, and, more importantly, to generate hypotheses for further research. More specifically, the authors suggested that many of the negative attitudes towards communication skills training could be associated with some of the students' demographical backgrounds (Rees & Garrud 2001). It is clear that any hypothesis on the attitudes demonstrated in doctor-patient communication must be motivated by research evidence. The current study formulated its hypotheses on the attitudinal role of speech prosody in health communication, on the basis of known themes present in existing literature. The wealth of previous research in health communication has also informed the current study of the development of its instrument of attitudinal measurement. This instrument, namely a professionalism in health communication scale, is a multi-item visual analogue scale (VAS) aiming at measuring perceived attitudinal aspects relevant to clinical interaction experience from communication features, which in the current study are those of speech prosody. Details of this instrument are reported in Chapter 5 which discusses a perceptual investigation to understand the link between the perceived level of professionalism and the prosodic features used in clinical communication.

Since measuring observed or perceived attitude can be done in the form of rating by another individual, for example service users, researchers or educators, attitude displayed in behaviour can be more objectively studied than using self-reporting scales. Recruiting participants who are potential users of the healthcare service to evaluate attitudinal information that healthcare professionals or trainees convey in their communication is therefore a useful method. Some form of attitudinal evaluation can be incorporated in OSCE communication skills assessment stations as an objective measure of communication behaviour (see Kurtz et al. 2005) as well as studies of medical communication skills. In Brinkman et al. (2006), for example, the evaluation of communication skills and professionalism of paediatric junior doctors, the observed behaviours under investigation also included several attitudinal aspects such as friendliness, respect, honesty etc. Parents of patients, nursing staff and attending physicians were found to give different evaluations on some aspects of the residents' communication behaviours, which showed that different observer groups can each offer unique perspectives into the residents' communication skills and professionalism (Brinkman et al. 2006). From a patient experience perspective, the attitudes of doctors as perceived by patients (or potential patients) in the clinical interaction best indicate doctors' professionalism in medical communication. This is because the cognitive elements of attitude (i.e. underlying beliefs etc.) can either lead to corresponding behaviour or be concealed by the individual, and therefore not directly known to the observer (patient in the case of the current study). Instead, the patient will use the doctor's communication behaviours to infer attitudinal information and make judgement on professionalism in communication. The current study needs to consider this in designing the instrument that best reflects attitudinal information perceived from the patient's perspective.

3.1.3 Summary

The messages from previous health communication research with different methods indicate that no single method is powerful enough to measure all aspects of health communication. Research methods are established to capture different aspects of the communication behaviours observable and of the attitudinal information conveyed in the clinical interaction. This varied from self-reported behaviours and attitudes to measures of attitudes on an appropriate scale. Research hypothesis on the communication of attitudes in clinical settings generated based on previous studies in particular should be tested using a suitable quantitative method. Furthermore, if the professional attitudes conveyed through health communication are measured via service users' perception of the communication behaviours present during the interaction, the measures are more likely to link to patient experience in health care. This can be made possible by measuring both the communication behaviours and the service users' perception on separate scales with a view to establish a link between the two.

No previous research reviewed here was dedicated, entirely, to the measurement of specific communication features during clinical encounters. The closest to this could be seen in Ambady et al. (2002) in which the surgeons' tone of voice was measured. Nonetheless, the measure was through judges' rating and was therefore impressionistic in nature. In other words, it was the effects rather than the properties of communication behaviours that were measured. Thus using this method can reveal what happens but not how things happen. Depending on the research aims of a study, such level of analysis might suffice but not in this study, since the research aims to consider both the properties and the effects of speech prosody, the communication behaviour in question. Before elaborating on the adopted methods to investigate prosodic features in this study, this chapter reflects on how the properties and effects of speech prosody are examined in research outside the healthcare context.

3.2 Research methods in speech prosody

3.2.1 Impressionistic measures of prosodic patterns

Impressionistic measures of the prosodic features can be taken in speech communication. These measures can use intuitive terms such as "a loud voice", "a high pitch" or "fast speech". Such measures have an ecology in everyday human communication experience. People usually use these descriptions to comment on their observation of the features and phenomena in their experience in daily life. However, speech communication research can take advantage of these descriptions, especially when subjective and categorical judgements are sufficient for the measures of the features of interest. For studies more interested in the gradient nature of prosodic features, more systematic measures are needed. For example, the nuclear tone analysis (see Section 2.6.2.1) use impressionistic terms in a systematic manner to investigate English intonation patterns. Other systems of transcribing intonation exist, including the Tone and Break Indices (ToBI, see Silverman et al. 1992) that adopts a two-tone treatment to intonation (Pierrehumbert 1980). The ToBI was designed to be a transcribing system capable of handling intonation and phrasing together while being machine-readable. Since then it has focused more on prominence and phrasing and has been widely used in linguistics and natural language processing. However the limitation of ToBI can be found in that being descriptive in nature the system can fall short in capturing listeners' perceptual experience in speech communication (Wightman 2002). To use a system that best describes a listener's experience is the main reason why this study chooses nuclear tone analysis over ToBI in discussing intonation patterns in clinical interaction. Furthermore, the use of nuclear tone analysis in the study of English intonation especially in the standard British variety is firmly established (see Cruttenden 1997, 2008). This is especially true when intonational meanings are concerned in the study (see Chapter 2 for a summary of the meanings of English nuclear tones) therefore reference to the meanings can be readily made in the discussion.

3.2.2 Instrumental analysis of prosodic features

Like other signals in speech, prosodic features in speech can be analysed using an instrumental approach. Using speech analysis computer software, researchers can rely on the visual aids of waveform and spectrogram to analyse the speech signals, which has been adopted as a standard practice in relevant research fields including linguistics and psychology. These enable the objective measurement of the acoustic properties such as F0, intensity and duration. Measuring these acoustic properties with a view to ascertain the affective information encoded through speech prosody has become an established method (Scherer 2003) and has been used in many of such studies (e.g. Scherer et al. 1984, Ladd et al. 1985, Banse & Scherer 1996). The F0, intensity and durational measures can be over a short or aggregated passage of speech and can be readily obtained in speech analysis software and do not require extensive phonetic interpretation (Scherer 2003). These measures are also well defined without controversy and are linked closely to speaker affects at least on the arousal level (Scherer 2003), and are referred to in the currently study as global features of speech prosody. More local acoustic-phonetic features e.g. peak value and alignment in primary accents can also be measured using instrumental techniques, and while they might be linked to speaker affects beyond the arousal level, they are often less consensually defined and require more phonetically orientated interpretation (Scherer 2003). Although examining these local prosodic features can lead to very interesting research findings, they are beyond the exploratory nature of this study. Furthermore, there was no indication from existing literature (Scherer 2003, Stern 2006) to suggest that the attitudinal aspects concerning professionalism in medical communication are expressed by these local features of speech prosody. The most obvious advantage for the current study to adopt the global measures of F0, intensity and speaking rate is that these are the most studied features in previous vocal emotional expression research (see Scherer 2003). Therefore the results of the current study can be compared with those from such previous research.

3.2.3 Measuring affective meanings in speech prosody

The acoustic properties of speech prosody in affective communication (encoding) are often investigated together with probing the affective meanings in speech prosody (decoding). Several methods have been used to study the decoding of speaker affects using either categorical or scale-based measures. The categorical measure of speaker affects can be in the form of recognition/discrimination of emotions. For studies aiming to determine if certain speaker emotions can be decoded through vocal cues by listeners beyond chance accuracy, this can be the preferred outcome measure (e.g. Johnson et al. 1986, Banse & Scherer 1996, Thompson & Balkwill 2006). A drawback of such a categorical measure is that the human perception of emotion or attitude is often not limited to a simple set of alternatives, but can be gradient and context-dependent too. That is to say, the opposing affective constructs (e.g. unfriendly vs. friendly) as perceived by a listener are likely to be a continuum and a measure is likely to be a point in this continuum and the value of the measurement is likely to be influenced by the context of communication. The gradient paralinguistic meanings (see Section 2.5.4), therefore, are not suitable to be measured categorically.

Furthermore, there might not be a one-to-one match between prosodic features and paralinguistic meanings. Instead, a certain prosodic feature is likely to be associated with a cluster of paralinguistic meanings (see Sections 2.5.2 and 2.5.4). This suggests that a scale-based measure containing the relevant attitudinal aspects as scale items is more suitable for the current study.

The scale-based measure of paralinguistic meanings can work in ways similar to the attitudinal evaluation measures in the health communication research reviewed above. This can be a Likert-type scale or other scales such as the Visual Analogue Scale (VAS). Although the Likert scale and the VAS are said to have comparable reliability (van Laerhoven et al. 2004), on some occasions, the VAS can show better responsiveness, for example in measuring anxiety in a one-item question, due to the greater resolution than the usually 5-point Likert scale (Heather et al. 2007). The VAS can also be useful for measuring on a continuum the qualities that are otherwise not directly measurable (Chen et al. 2004). Ladd et al. (1985) used 8-point rating scales for both emotions (e.g. aroused, arrogant, etc.) and attitudes (emphatic, reproachful, etc.) in their experiments and then associated the measures with acoustic properties in speech to understand the effects of prosodic features as signals to speaker affects. Similarly, Goble & Ní Chasaide (2003) used 7-point scales with bipolar descriptors to measure emotion, mood and attitude (relaxed/stressed, content/angry, friendly/hostile, sad/happy, bored/interested, intimate/formal, timid/confident, afraid/unafraid) in order to understand the role of voice quality in these attributes. The instruments in these studies can be regarded as forms of numeric VAS (see van Laerhoven et al. 2004) on which the response takes a numeric form (e.g. from 1 to 10). This differs from the simple VAS on which the response is marked on a straight line usually 100mm in length. The latter was also used in a study (Chen et al. 2004) to examine cross-linguistic difference in the paralinguistic meanings of intonation that utilised four scales for friendly, self-confident, emphatic and surprised measures. Although both forms of the VAS can be regarded to be able to capture gradient scale measures, the simple VAS was said to show better bias control (Yiu & Ng 2004).

There is both good rationale and practicality for the current study to adopt the simple VAS to measure the attitudes related to professionalism in clinical interaction. While it is possible to find accurate descriptors for attitudes, it is difficult to verbally specify the strength of attitudes in categories with accuracy and good theoretical support. Therefore for measuring attitudes on a continuum rather than in categories, the VAS is a valid and sensitive tool.

3.2.4 Speech samples for studying prosodic features

There are several methodological issues to take into account when considering the speech samples needed to enable study of prosodic features. These include the content and context of the interaction, the length of the speech material, and the encoding of the paralinguistic meanings.

In additional to healthcare, non-verbal behaviour has been studied in other professional communication contexts such as business interactions (Fatt 1998), library studies (Radford 1998) and politics (Grebelsky-Lichtman 2010). These studies, however, often included a wide range of non-verbal behaviour and did not focus on prosodic features. Therefore they did not only rely on speech data. In previous studies of prosodic features, the speech data were often obtained from educational settings, particularly those involving non-native speakers. This can be well justified because researchers often want to know the performance of foreign learners of a language (e.g. English) by comparing it against the native norm. Several studies reviewed in the previous chapter fall into this group, including Wennerstrom (1994) on the intonation patterns of Japanese, Thai and Spanish learners of English and Verdugo (2005) on the expression of certainty by Spanish learners of English. As Lieb & Shah (2010) pointed out in their study of non-verbal communication in political campaigning, communication behaviour is bound by cultural and contextual rules. Although the studies above produced interesting findings in the paralinguistic use of prosodic features and their investigation techniques can be considered in the current research, their speech samples used are essentially representative of everyday interaction scenarios covered in the English language classroom. As pointed out in Chapter 2, healthcare interactions often involve vulnerable patients and can have a heavy emotional role. Thus the intents of health communication can also differ from those in everyday life. Therefore for the current study, the material for collecting speech data must be immediately relevant to the clinical settings.

When considering emotional speech for the study of vocal expression of emotions, the data are often collected through the portrayal of (usually basic) emotions by actors. These can be seen as stereotypical snippets of vocalised affects in real life (Scherer et al. 2003) and their usually brief length may be sufficient to use as stimuli for the studying perception (see Ambady et al. 2000 on using thin-slice judgment of behaviour). However, such stereotypes may not fully represent the subtlety of natural expression of affects (Scherer et al. 2003). Thus some more subtle aspects of attitudinal attributes might not be simulated accurately by actors and cannot be studied with data collected using this method. Nevertheless, usable speech data to study prosodic features, especially with a view to examine attitudinal and/or emotional expression in professional communication can be obtained through simulated interactions. The speech samples in Scherer et al. (1984) were obtained in this way, where the interviews between German social agency workers and actors role-playing the clients were recorded. Since the interaction was highly natural, and the paralinguistic meanings in the professional speech are not scripted, Speaker affects of the professionals observed could be regarded as naturally occurring. Thus by using simulated clinical interaction with an actor role-playing the patient, the current study can obtain realistic speech samples from medical students with the attitudes they intend to convey in the interaction encoded. Speech samples obtained this way are therefore valid data to study speaker affect.

There is currently a methodological niche in research on the attitudinal meanings in speech prosody in a health communication setting. Although the methods for investigating the speaker affects through prosodic features are well established in linguistic and psychological studies, the attributes studied are most likely to be limited to basic emotions that are not contextual (Scherer et al. 2003). Applying these methods to the study of expression of attitudes in health communication context is still novel. As pointed out in Section 3.2.3, attitudes expressed in speech can be seen as of a gradient rather than a categorical nature, so measurement methods such as listener recognition or discrimination do not work very well for this type of speech data nor for the purpose of the current research. On the contrary, attitudinal rating is not only appropriate for this type of data but also possible to be used for establishing the statistical association between the acoustic measures of prosodic features and the attitudinal meanings expressed (see Scherer 2003). Therefore this study occupies the aforementioned methodological niche by studying attitudinal information conveyed in prosodic features in a simulated clinical interaction.

When studying attitudinal or emotional expression in speech, the control for the verbal content also needs to be considered. One easy way to do so is to use nonsense syllables (see Scherer et al. 2003), but this lab controlled environment results in data samples that cannot be generalised to real life situations. Standardised texts have almost the same limitation: in real life the same words just would not always be repeated over and over by different speakers on different occasions. More elaborately, the verbal content can be masked by software manipulation to reverse, randomly slice or low-pass filter the speech. These techniques each have pros and cons in terms of retaining prosodic features while controlling for speech content. All these techniques can render the verbal content unintelligible. Reversed speech keeps the original F0 and intensity means and ranges while changing the F0 contour. Random slicing keeps the same original properties while making the speech sound choppy. Low-pass filter retains much of F0, intensity and temporal characteristics in the original speech but loses some acoustic energy in the high frequencies that are filtered out (see Scherer et al. 1984). Therefore the justification of using this technique as a means of control again depends on the research aims; if losing the affected characteristics and naturalness is not a concern, then they can be safely used in the study. Alternatively speech samples can be obtained from a single speaker to modify the prosodic features by means of resynthesis (see e.g. Chen et al. 2004, Gobl & Ní Chasaide 2003). This method can be used most effectively when there is a strong, ideally theory-driven, linguistic hypothesis to test as is in these example studies. In the current study, however, the hypotheses are mainly generated in the clinical and English as a second language backgrounds. Because of this, the most suitable way of controlling for the verbal content of speech is not via synthesis or resysnthesis but by standardising the simulated patient and consultation scenario.

3.3 Summary and methodology adopted

The interdisciplinary nature of the current study is reflected in its research questions as well as methodological considerations. Each division of research methods discussed above serves an area of investigation in this study. Firstly, examining the global prosodic features in affective speech is well-justified by accumulated empirical evidence. Secondly, measuring perceived attitudes by listeners can take the form of well-established method of rating scales, and in this case the VAS for response on a continuum. Good statistical tools also exist for establishing the link between the two measures, i.e. the acoustic and the perceptual. Thirdly and most importantly, listeners' rating sheds light on evaluating service user experience in healthcare research by adding to the empirical evidence on perceived professionalism. The current section reports the methodology adopted in this study as an effective way to serve the purposes of the interdisciplinary study.

As the research questions in this study are organically connected to one another, the chosen methods according to each research area also need to be integrated. Therefore the methodology adopted in this study needs to be considered holistically. This study intends to ascertain the role of speech prosody in health communication as a means to convey aspects of professionalism with evidence from two speaker groups training to become doctors. As such, in consideration of the discussion of potential methodologies above, this study seeks to achieve this in a two-step approach, one acoustic and one perceptual, that is, to investigate the acoustic properties of speech prosody and the perceived attitudes in health communication in turn. Each stage of investigation has its own questions. The acoustic study is designed to examine the differences in prosodic features when different attitudes are to be expressed in speech as well as the differences in prosodic features between the two speaker groups.. The perceptual study is designed to explore the effects of the prosodic features in use in such communication on the level of professionalism perceived by potential service users. Ultimately there is a third stage that answers the overarching research question whether or not speech prosody matters in health communication using the empirical evidence collected. This stage draws on the discussions in both the acoustic investigation and the perceptual investigation as well as results of the statistical analysis that are able to predict the level of professionalism conveyed/perceived with both the prosodic features and the speaker attributes.

Details of the methods in each stage of the current study are reported separately in Chapters 4 and 5. However, these are summarised below so that an overview of the research methodology adopted is provided here before the investigations are discussed in due detail in the corresponding chapters.

3.3.1 Designing the acoustic study

Answering Research Questions 1 and 2 (see Section 2.7) relies on an acoustic investigation that adopts an observational design to examine both the within- and between-subject effects on the prosodic features of interest.

Medical students rather than qualified doctors form the population under study for several reasons. Firstly, studying qualified doctors for this research topic does not warrant more validity than studying medical students, since the theme of the study is neither the professional knowledge nor the professional qualification in medicine but the professional "voice" in medical communication. Medical students are required to display a level of professionalism similar to qualified doctors in their communication with patients. Secondly, the medical student population suits the purpose of this study better than the doctor population. In the doctor population, there are a few "noise" factors that are likely to confound the investigation, for example, the specialty of practice and the environment of the clinic. Sampling from the student population instead can avoid these potential confounders. Thirdly, the medical student population enables convenient control for the level of communication skills training due to the stage of the medical education. This is more difficult to control for among qualified doctors. Fourthly, findings on the communication behaviour of medical student population can have immediate implication for real medical practice, since the medical students are training to become future doctors, and during their placement they already interact extensively with real patients. Therefore medical students are considered to be a suitable population for studying prosodic features in health communication.

The within-subject factors can be investigated within a repeated measures design in the speech materials for data collection. Participating medical students perform two tasks in this study.

The first task is a consultation role-play in a standardised, OSCE-style scenario in which the medical students interact with a simulated patient (played by a professional actor). This scenario does not aim to test the students' medical knowledge but to allow them, instead, to use all their trained communication skills in a structured consultation (i.e. initiating the session, gathering information, explanation and planning, closing the session; see Silverman et al. 2005). The scenario for simulated consultation serves as an approximate control over the communication content. This would be difficult to achieve if the speech data were to be collected from consultation with real patients, let alone the fact that there would be insufficient justification for real patients to be involved at the early stage of study of this topic in which benefits are not yet certain. Clear instruction is also given to the simulated patient, including detailed information about the character's background in a format familiar to training and assessment occasions, in order to assist the preparation of the acting.

The second task is sentence reading. In order to explore the extent to which speakers manipulate prosodic features to convey attitudinal information, a list of five clinically relevant sentences is compiled with two distinct set of attitudinal instructions. These are sentences that would be frequently observed in medical consultations and can be found in medical communication textbooks. Therefore they have apparent face validity. The attitudinal instructions are designed to induce two sets of attitudes, one appropriate to the professional communication in the consultation room and the other inappropriate to that context. These are summarised in Table 3.1 below. These attitudinal instructions are by no means a complete list of the attitudes required of the doctor in patient-centred care or those that the doctor should avoid. Rather, they are designed so that the speech prosody in the read sentences may be varied and may be observed due to the function of the attitudes encoded (amongst other functions).

Table 3.1 Encoded attitudes in read sentences

Professional	**Unprofessional**
respectful/attentive	casual/unpremeditated
non-judgmental	accusing
empathetic	patronising
encouraging	demanding
interested	harsh

The professional attitudes shown in Table 3.1 reflect the requirements in *Good Medical Practice* (GMC 2006) and the training objectives in the Calgary-Cambridge guide (Silverman et al. 2005) and *Tomorrow's Doctors* (GMC 2009). In other words, these attitudes are part of the expectations for both student and practising doctors in their relationship and communication with patients. Failing to demonstrate professional attitudes, or more seriously, demonstrating the unprofessional attitudes such as those in Table 3.1 present would mean failing the GMC's expected standards.

The reason for designing the sentence-reading task is to explore whether induced attitudes can be encoded in speech (see Scherer et al. 2003) with a range of prosodic features (acoustic parameters/vocal cues). While it is true that a particular attitude can be indexed with a combination of acoustic parameters, the combination is at the speaker's disposal. That is to say, a speaker does not necessarily have to follow the stereotypical vocal expression but can choose the cues in an idiosyncratic way from the combination, although the decoding of the chosen cues is then at the listener's disposal. The implication of this on the design of this study is that comparing the acoustic parameters used can reveal intra-speaker variation due to the function of the encoded attitudes as well as inter-speaker variation due to group attribute differences.

The acoustic parameters examined in this study (which are introduced in Chapter 2) are summarised in Table 3.2 below.

Table 3.2 Acoustic parameters under investigation

Acoustic parameter	Perceptual correlate	Base unit
F0 mean	Pitch level	Hz
F0 range	Pitch range	Hz
Mean dB intensity	Loudness	dB
Speaking rate	Rate of speech	Syllables per second

These parameters in Table 3.2 can be subject to reliable instrumental measurement using dedicated speech analysis software. As discussed above (see Sections 2.6.2 and 3.2.2), these parameters are global features of speech prosody. They are likely to be nearly universally employed in the vocal communication of affects, and therefore interpreting the difference in them does not need to depend on linguistic theoretical frameworks, unlike local features the interpretation of which often does.

3.3.2 Designing the perceptual study

In order to answer the Research Question 3 in Section 2.7, a perceptual study is designed to explore the effects of speech prosody on the level of professionalism listeners perceive in health communication. Such a study resembles psychological experiments in design but differs from typical clinical trials in randomisation and blinding. The idea is to use speech materials differing in encoded attitudes as well as acoustic properties as stimuli to test listeners' perception by means of rating the perceived attitudes in the stimuli.

Given the notable level of patient complaints about communication and staff attitudes within the NHS, as well as the proportion of foreign doctors and medical students in the UK, perceptual data is best collected from potential users of the services provided in the NHS. Ideally a sample should be the cross-section of a rather diverse community of which the members have entitlement and access to the service. This sampling technique can make the findings generalisable to the experience of an average potential service user.

The speech data collected in the acoustic study are suitable for making the voice stimuli to use in the perceptual study. There are two reasons for this. Firstly the speech content of the acoustic study data are essentially controlled for, and the attitudes in the speech are induced in two different styles. The attitudes induced in the sentence reading task are under explicit instruction. Therefore the vocal display of them is expected to be clear and strong. The attitudes induced in the simulated consultation task are without explicit instruction. However, they are expected to be present in any interaction with patients. Therefore their presence in a simulated consultation is almost naturally occurring and should represent the same attitudes displayed in actual clinical encounters with real patients. Secondly the acoustic properties of the stimuli materials will be known. This is important because by knowing the acoustic properties of the expressed attitudes, it is then possible to study both the encoding and the decoding processes of the vocal communication of professional attitudes. In other words, the acoustic parameters for displaying professionalism are and the effects of these parameters on listeners' perception of professionalism can be potentially correlated.

3.3.3 Answering the overarching question

After obtaining both the acoustic measures and the attitudinal ratings, the possibility emerges to understand how attitudes are inferred using vocal cues in speech by listeners. This requires the examination of their statistical association. It is possible to use multiple regression analysis with the acoustic measures as the predictors and the attitudinal ratings as the outcome measure. This technique has been established (see Scherer et al. 2003) and has shown to be fruitful in similar research, for example the acoustic profiles of emotional expression (Banse & Scherer 1996). Extending such techniques further to the expression and perception of professionalism in this study can be expected to produce novel findings with important implications.

Results of the statistical analysis after considering all factors investigated can therefore address the overarching research question. If a significant association is established between the prosodic features under investigation and professionalism in communication scale score after controlling performed, even with the effects from speaker attribute factors considered, it will indicate positively that speech prosody does indeed matter in health communication.

Chapter 4 The acoustic study

As discussed in Section 3.5 above, the first stage of the investigation is an acoustic study that contributes to answering Research Questions (1) and (2) raised in Section 2.7 and reproduced below:

(1) Do medical students use speech prosody to convey professional attitudes in medical communication?
(2) Are there differences between native and non-native English students in the use of prosodic features in the simulated clinical interaction?

This chapter reports the data collection, data analysis and results of analysis in the acoustic study. As a pilot to explore the attitudinal role of speech prosody and prosodic differences between native and non-native speakers of English in the clinical context, it is not meant to be a large scale endeavour and may be underpowered.

4.1 Ethics

An application for ethical approval was submitted to the Faculty of Medicine and Health Sciences (at that time, the Faculty of Health) Ethics Committee at the University of East Anglia (UEA) in June 2010, and full ethical approval (see Appendix A) was granted in July 2010. The following ethical considerations were communicated to the Committee in the study protocol:

- All stages of the project are non-invasive and pose no risk whatsoever to any of the participants. However, in the very unlikely case that any of the participants experience distress from the role-play, they are advised to seek support from their personal advisor or the Dean of Students' Office.
- All participants will be recruited on a voluntary basis. Information provided by volunteers who are not selected for the study will be destroyed. Participants can choose not to participate and withdraw without giving a reason. Should a participant withdraw from the study, all his/her data will be

destroyed. Participants will give explicit informed consent to recording and analysing their speech samples.

- All personal information including audio-recorded data will be kept confidential; all paperwork will be stored in locked filing cabinet on campus and audio recordings on password protected UEA server space. If a participant does not consent to the use of his/her data for future research, the data will be archived for a maximum of 5 years and then destroyed. If a participant consents to the use of his/her data for future research by initialling the corresponding box in the consent form, the data will be kept securely with the same level of confidentiality for 5 years after the end of the study and will be destroyed after that.
- Throughout the project the anonymity of the participants will be protected. The real names of the participants will never be revealed; they will be referred to by their participant numbers during the research and in publications.

In addition to the study protocol, the application also included participant information sheet (PIS) for student participants, PIS for the simulated patient, screening questionnaire for participants and consent forms (for student participants and for the simulated patient), task information for student participants and OSCE-style role-play information for the simulated patient (Appendices B-I). These among other supporting documents were submitted together to the Committee for consideration.

4.2 Study design

This study adopted an observational design to explore the potential differences in the speech prosody in medical communication. Two naturally formed groups were of interest in this study, namely native English speakers and non-native English speakers. Such design is sometimes known as a “natural experiment” (Winkelmann 2004), since the two groups were naturally assigned according to their different language exposures. In other words, one group naturally serves as the control of the other group under the study condition. Observations obtained from the two groups are therefore used to examine the between-subjects effect. Section 4.3 below reports the recruitment of the two groups of participants as well as the general characteristics of the participants.

This study also involved a repeated measures design in that each subject performed a series of speech tasks so that the prosodic features in each task could be measured and compared. Such observations are therefore used to examine the with-subjects effect. The detailed procedures in data collection, including the contents of the tasks are reported in Section 4.4 below.

The study design has consequences on the choice of statistical methods. Examination of the between-subjects effect requires the comparison between two independent samples, and of the within-subjects effect the comparison between two related samples or among several related samples. The consequences of the research design on the consideration of statistical analysis, including description of the dependent and independent variables, are discussed in Section 4.5.

4.3 Participant recruitment

Immediately after full ethical approval was obtained, a call for volunteers and the participant information sheet (PIS) were published on the virtual learning environment at UEA, available to all potential participants. Additionally, the researcher was allowed to speak to potentially interested participants at some teaching venues. Interested volunteers were then sent a screening questionnaire which they were instructed to complete and return (either by email or by hand) to ensure that they met the inclusion criteria below. A purposive sample of 20 participants, with 10 in both the native speaker and non-native speaker groups, was originally envisaged. That target was based on the pool of eligible participants according to the Medical School student register, as well as volunteer response rate in previous projects on healthcare student population. The recruitment of non-native speakers proved to be difficult. However, this stage of the investigation required a balanced design with equal distribution in English language background, gender and training level. This resulted in an amendment to the original ethical approval applied for and granted, which authorised contacting all potentially eligible non-native participants via UEA email to increase the chance of recruitment in the group. The opportunity for the participants to win a £50 Amazon voucher in a prize draw was also offered at that time as a more appealing incentive.

Both male and female participants were recruited from students enrolled in the MBBS programme of the Norwich Medical School at UEA into two groups: native speakers and non-native speakers, according to the following criteria:

Inclusion criteria

- UEA MBBS students
- Males and females
- Native speakers of English with a Southern Standard British English (SSBE) accent which does not easily disclose the speaker's geographic origin
- Non-native speakers of English who have met UEA English language requirement and therefore acquired a high level of proficiency
- One year or longer communication skills training on the MBBS course (i.e. Year 2 onwards)

Exclusion criteria

- Self-reported hearing or speech difficulty
- Aged 40 or above (when there is increased likelihood of age-related voice change, especially in females, see Sato 2002, Linville 2004)

The balanced design makes it possible to analyse the effect of the between-subjects factors "English language background" without the interference from the factors "gender" and "training level". Furthermore the criteria above serve as a means to avoid introducing confounding factors into the acoustic analysis of speech prosody which could compromise the validity of the study. The factors in the study design, together with the potential confounders, are expounded below.

a) English language background

The native speaker group sample was taken from a (near) homogeneous speech community, that is, native speakers of SSBE in the UEA MBBS students. There are two main reasons for this inclusion criterion. Firstly, there have been accounts on the attitudinal perception of regional accents in British English, especially on the associated attractiveness and prestige (see Hiraga 2005, Coupland & Bishop 2007), and regional accents can be exhibited in prosodic features (see Cruttenden 1997). However, the current study does not intend to analyse regional variation in prosodic features in British English. Therefore including regional accents in the speech data could potentially confound the analysis.

Secondly, SSBE is commonly used for news reading on national broadcast in the UK and adopted as the pronunciation model (especially in its prestigious form, Received Pronunciation (RP)) to teach English to non-native learners. Extensive documentation of this variety in the literature (e.g. Cruttenden 2008) provides established interpretation of attitudinal meanings of F0 contours, enabling the discussion of appropriateness of some F0 contour types in a professional context. Analysis of the F0 contours in read sentences of native speaker participants can thus be compared to such accounts in the literature. If the participants in the current study used any contextually inappropriate contour types when asked to convey professionally appropriate attitudes for relationship-building, they could be lacking, to some extent, awareness or understanding of the attitudinal function of speech prosody. This could potentially highlight what can be missing from current communication skills teaching.

Volunteers from UEA MBBS students who speak a native language other than English were sampled to the non-native speaker group. Norwich Medical School at UEA has specific English language proficiency requirement for overseas applicants for the MBBS programme, if their first language (L1) is not English. Successful applicants need to demonstrate, prior to enrolment, proficiency in English that is equivalent to an overall score of 7.5 or above in the International English Language Testing System (IELTS) with a minimum of 7.5 in each of the listening, reading, speaking and writing components. English proficiency at this level is to be considered 'acceptable' for 'linguistically demanding academic courses' (*IELTS Guide for Educational Institutions, Governments, Professional Bodies and Commercial Organisations*). An IELTS score of 7.5 or above also well exceeds the threshold of the "proficient user" level within the *Common European Framework of Reference for Languages* (Council of Europe 2001).

All non-native speaker participants recruited to this study, therefore, had acquired a good command of English, and none reported any difficulty in using the language as a medium for communication or learning. A number of these participants had lived in bilingual countries, for example Malaysia and Mauritius, prior to arrival in the UK. None of the non-native participants showed segmental features (i.e. vowels and consonants) of strong foreign accents in their speech, as confirmed by the simulated patient who is a British professional actor role-playing with them. Their accents did not easily disclose their geographic/cultural origin either, as stereotypical foreign accents would. Only one subject revealed to the researcher during the study that she was self-conscious about the potential negative influence that her mild foreign accent might have on her future medical career.

It should be noted here, however, that non-native speakers were recruited regardless of their L1s, because the current study does not intend to investigate how different L1s influence the production of English prosodic features by non-native speakers. For this reason, a plain distinction between native speaker and non-native speaker was maintained in the data. Further details of their L1 backgrounds and age are provided in Appendix J.

b) Gender

Due to the physiological difference in the vocal folds between men and women, the voices of men and women are naturally different, especially in F0 (see Section 2.6.1). Both male and female voices in professional communication are included in the sample. It is important to account for this physiological gender difference in the statistical analyses, which will be discussed in Section 4.5 below. This control is also necessary because in the second stage of the study, the perceptual experiment will explore the potential difference in perceived professionalism between male and female voices in the clinical setting.

c) Training level

The sample in the study contained two groups differing in their level of communication skills training received. One group included MBBS students in the second and the third years whose level of communication skills training was regarded as lower while the other included those in the fifth year whose level of communication skills training was as higher. This control enabled the study to take into account any effect of communication skills training on the prosodic features used by the participants in professional communication.

Communication skills training in the MBBS programme at Norwich Medical School in the fourth and fifth years covers more advanced contents and is more intensive compared to that in the first three years. First year students, however, were not recruited because their training in both medical knowledge and communication skills can still be limited, even for the general consultation task used in the current study (discussed later in Section 4.4). Including those students in the study would also potentially confound the voice analysis, since it cannot be certain yet whether or not medical school training influences the use of prosodic features by the students in professional communication.

A purposive sample of 16 participants whose mean age was 25.81 (*SD* = 5.53) years was therefore achieved, with balanced distribution in English language background, gender and training level, as shown in Table 4.1 below. Additional demographic details of the participants are reported in Appendix J.

Table 4.1 Number and distribution of participants

Level of training	**Native speakers**		**Non-native speakers**	
	Male	Female	Male	Female
Lower	2	2	2	2
Higher	2	2	2	2
M (*SD*) age	**23.50** (4.43)	**23.75** (3.50)	**29.50** (5.98)	**26.50** (7.33)

It can be seen that this study would be underpowered to investigate the effects of all three factors above on prosodic features. Instead, this study chose to focus on the analysis of the effect of English language background alone, and include gender and training level in the balanced design to represent the medical student population under investigation. Therefore the only group difference compared in this study was between the native and the non-native speakers, a consideration reflected in Table 4.1.

4.4 Data collection

4.4.1 Materials for data collection

Materials were prepared in order to collect two sets of speech data, namely sentence reading and consultation role-play. Since speakers are known to use different forms of the same linguistic features according to different styles (Labov 2001), it is important to take both controlled speech and free speech into consideration. In this investigation, these styles were represented by sentence reading and consultation role-playing respectively.

a) Read sentences

In order to ascertain whether medical students manipulate prosodic features to convey professional attitudes in medical communication a list of five clinically relevant sentences were compiled together with two distinct set of attitudinal instructions (see Appendix H). The sentences could be observed in medical communication textbooks or real-life consultations and ranged from greeting the patient to asking close-ended questions, which are reproduced below.

(1) *How are you?*
(2) *How much alcohol do you drink in a week?*
(3) *I'm sorry to hear that.*
(4) *Do you have any thoughts on this?*
(5) *Tell me more about that.*

Since using simple, one-word descriptors as instructions could lead to various interpretations and consequently invalidate the control, concise but clear phrases were used instead. These phrases resemble the style of the Calgary-Cambridge model (Kurtz et al. 2005) in which the participants receive their communication skills training. It was expected that all participants would be familiar with the wording of the instructions and would find little difficulty in following them. In fact none of the participants requested assistance or clarification regarding the instructions even though they had been told that they should ask when they felt necessary.

These instructions represent attitudinal aspects that are appropriate to the professional communication in clinical settings and those unacceptable when interviewing the patient, respectively. The rationale for using them is discussed in Section 3.5.1 above, and their details can be found in Appendix H.

Both professional and unprofessional attitudes, can be effectively conveyed by adjusting prosodic features accordingly. For example, one possibility of characterising the statement '*I'm sorry to hear that*' as an empathetic expression can be by lowing acoustic intensity and/or speaking rate, amongst others. Conversely a patronising '*I'm sorry to hear that*' can take an uptrend F0 contour (see Cruttenden 2008:284) with wide F0 range and normal intensity. Thus this task will allow the examination of the effects of attitudes on prosodic features in the clinical context. When comparing between the native and the non-native groups in reading the same sentences in a professional manner, the effect of the factor "English language background" can also be examined.

Reading style has long been used in sociophonetic research to elicit forms that speakers are consciously aware of (Chambers 2009). Combining sentence reading with attitudinal instructions in data collection enabled the investigation of any adjustment in prosodic features by the speakers according to different attitudinal information that they were to convey.

b) Consultation role-play

In order to collect speech data in a more spontaneous style than sentence-reading, a role-play scenario (see Appendix H) was devised. This scenario involved the subject role-playing as the medical student on primary care placement seeing a patient for hypertension treatment and management, under the GP tutor's instruction. Hypertension treatment and management was chosen because it was a non-challenging, general scenario suitable for all the participants in the study. Even the participants with the earliest stage of training in the current study had already covered the essential medical knowledge about the condition. Therefore all participants in this study would have similar level of confidence to perform the communication tasks in this scenario. While the fifth year medical students were more likely to be able to provide more detailed explanation and advice to the simulated patient, it would only be a difference in the amount of information but not in their voice. In short, this chosen scenario would not allow the different levels of medical knowledge learnt by the medical students to confound the analyses of prosodic features used in consultation. It should be noted, however, that the second year medical students had not formally learnt or practised, in their communication skills training, planning (especially shared decision making) which were more familiar to the other students in this study. Even so, such difference in training was not expected to cause considerable difficulty for the second year students to complete the task which required them to discuss the treatment and management of the condition rather than to negotiate and agree on a plan. With the considerations above taken into consideration, this consultation role-play scenario was still considered suitable to observe prosodic features used by the medical students in spontaneous speech, albeit with some appropriate content control.

A paid actor was employed as the simulated patient to interact with the participants in the consultation role-play. The contract with the actor was through an agency that had worked closely with the Norwich Medical School by supplying role-playing actors for communication skills training. The actor himself had had extensive experience of role-playing in medical communication skills teaching as well as OSCEs in the East Anglia region, and was suitable for the role of the simulated patient in the study. He was briefed about the job and given the scenario material which was produced in the manner familiar to him as an experienced medical role-player. He had prepared his role thoroughly before the sessions. All participants role-played with this same actor in order to avoid the difference in interlocutor which could confound the investigation (see Section 5.2).

Clear task information/instruction was provided in a way familiar to the participants in this study. This particular scenario asked the medical student to perform the following consultation tasks as they were taught in the Calgary-Cambridge model: (a) to initiate the session, (b) to gather information using closed- and open-ended questions, (c) to explain and to plan/suggest plan, and (d) to close the session. The style in which the materials were phrased was similar to that the Objective Structured Clinical Examination (OSCE) communication skills stations, which are extensively used as a form of assessment in UK medical schools. Medical students at

UEA had also familiarised themselves to the structure of consultation through training.

According to the communication process skills guide in the Calgary-Cambridge model, there are opportunities for appropriately displaying professional attitudes (e.g. interest and respect) in performing each task in the structured consultation. In doing so, not only are the necessary contents covered in the consultation, but the essential rapport with the patient is established and developed in the process as well (Silverman et al. 2005).

Using consultation role-play data in this study has two advantages:

(1) The consultation role-play assures a degree of control over the structure and content of the consultation, and eliminates the effects introduced by different interlocutors (the same simulated patient is used), which would not be possible if real consultations were used.
(2) Given that the above control is necessary, consultation role-play provides the most realistic data possible to study communication in a clinical context.

Analysing the acoustic parameters of the prosodic features in this data set can reveal intra-speaker variation as a function of different tasks in the structured consultation, as well as inter-speaker variation due to differences between speaker groups.

4.4.2 Procedure

All participants gave informed, written consent prior to the recording of their speech. The participants had already been informed of the tasks that they would be performing during the data collection. Participants were recorded individually in the sound-treated UEA Telecommunication Studios, with the exception of three participants who were recorded individually in a similarly quiet seminar room, when the building was empty. This was due to the unavailability of the Studios during those particular sessions, and an inability for any of the participants to reschedule.

At the start of each recording session, the participant was supplied with the instruction material and was given time to prepare. All participants completed consultation role-play before sentence reading. The main reason for completing a more spontaneous speech task of role-play before completing a more controlled one of sentence reading is to overcome the “observer’s paradox” (Labov 1970:32 cited in Chambers 2009:19). The observer’s paradox must be taken into consideration in any study of language as it is used in its social context, since the observed behaviour might be influenced by the fact that the participants are conscious of being observed. On the one hand, as the instruction for the role-play was on the content and structure of the simulated consultation (see Section 5.1.2), the participants in this study were expected to approach the task as in an OSCE station or on real placement and to display attitudes in their speech as they would on those occasions. Spontaneous speech thus collected would then indeed be spontaneous sounding. On the other hand, as the instruction for the sentence reading explicitly asked for certain aspects of attitude, the participants would be aware of what attitude was expected in speech when reading that particular sentence. The displayed attitudes in such speech would indeed be controlled. However, if the participants in this study were to complete sentence reading before role-play, the attitudinal instructions for the read sentences might influence their behaviour in the consultation role-play, and as a result make their attitude in the spontaneous speech in the role-play less spontaneous and less representative of real-life situations.

In order to make the recording environment resemble the consultation room as much as possible during the recording, the medical student subject and the simulated patient were sat in a manner recommended in the consultation skills training. The positioning of desk and seats provided the appropriate level of comfort as well as enabled essential eye contact and body language essential to the clinical interaction. This however, limited the choice of microphone mounting. Since only the speech of the medical student participants would be analysed, one channel (mono) recording was deemed appropriate. During the discussion on the recording setup, an in-house media production technician had suggested that a lavaliere (lapel) microphone as the suitable choice for the recording sessions. This was a Sony ECM55B, an omni-directional condenser microphone which was capable of the sensitivity and frequency response required for the acoustic analysis. Mounted on the medical student subject’s clothing at a distance of approximately 20 centimetres from the mouth, the microphone was connected to a Marantz PMD660 solid state recorder set to record in uncompressed PCM format at the 48 KHz sampling rate.

All the recordings were carried out by the researcher who sat at the far end of the desk controlling the equipment during the sessions and was not directly visible to the medical student participants. In this way his presence at the scene would not distract the participants from their role-playing or sentence reading tasks, but would convey a sense that technical help during the recording process would be available to them, should they ever wish. Recordings started with the participants indicating readiness and ended with completion of role-playing or sentence reading.

4.5 Data analysis

4.5.1 Acoustic analysis of speech data

Once collected, the speech data were edited to prepare for analysis using speech analysis software package *Praat* (Version 5.1.31, Boersma & Weenink 2010).This involved removing the actor's voice that was captured in the background from the recordings as well as segmenting the recordings into separate sound files which were the individual read sentences or passages in the consultation role-play. The sentence files contained a short period of silence (approximately 50 milliseconds) at both the beginning and the end, while the role-play files did not. Wherever a complete speaking turn by the participant in the role-play could be isolated, this was preserved entirely in the sound file. However, in some role-play passages there were overlapping voices (of the participant and of the simulated patient), making it not possible to preserve a complete speaking turn containing only the participant's voice. Should this happen in a non-final position in a sentence, efforts would be made to segment the recording at the boundary of an intonation phrase when saving the sound file (this could correspond to the end of a phrase or clause). These files were then labelled with unique participant numbers of the participants as well as content information such as sentence number or text of the passage, and were stored and backed-up in password protected UEA server space. The labels for the role-play files also indicated the purpose of the utterance in the consultation (initiating session, gathering information, explaining/planning and closing session), and such information was recorded alongside the orthographic transcript for each utterance on a spreadsheet.

Initially, the following measurements were taken from each sound file: mean, minimum and maximum of F0 (Hz), mean intensity (dB), duration (seconds) and the number of syllables. F0 range ($F0_{max}$-$F0_{min}$) and speaking rate (number of syllables/duration) were later calculated on the spreadsheet. For the read sentences only, the shapes of the F0 contours were also analysed, which will be discussed separately.

For F0 measurements, the default pitch settings in *Praat* were used. These were setting the floor at 75 Hz and the ceiling at 500 Hz, and take the measurements on the Hertz scale. The voicing threshold (the threshold above which a segment is considered voiced) was normally set at 0.45, except for occasions on which *Praat* reported spurious pitch values. If a pitch value (often with an octave error) was reported in a voiceless segment (based on auditory judgment), it was regarded as spurious with confidence. This could happen when, for example, strong high frequency energy was present in the voiceless fricatives produced by some speakers. If an error was suspected in the pitch value of a segment in which voicing was present, this was checked with calculation by hand. Some spurious errors could be eliminated by raising the default voicing threshold setting when analysing the file. However this would be at the cost of risk missing out some voiced segments in the pitch analysis, especially if the voicing threshold were to be set too high. Therefore if eliminating a spurious pitch value in any segment should require raising the voicing threshold above 0.65, which would be too high for the purpose of this study, that affected segment would be discarded from the analysis. The default voicing threshold setting of 0.45 was used in all unaffected files.

For intensity measurements, the view range was set to 50-100 dB (default) and the averaging method to ‘mean dB’. This averaging method would return values as the mean of the intensity curve in dB in the files.

For duration measurements, the measurement points were taken at the beginning and the end of the files. These coincided with the onset and offset of the acoustic energy in the files, defined as the first and the last zero crossing in the corresponding waveform (see Watson and Hughes 2006, Zhang et al. 2008). Spectrographic and auditory evidence was also used to assist the judgement where necessary, following the established guidelines for acoustic-phonetic analysis (e.g. Ladefoged 2006, Johnson 2003).

Since the role-play data does not control for the content of the consultation, but rather the tasks (see 5.4.1 above), the acoustic measurements of each prosodic feature were then averaged according to the task of consultation. One value of mean F0, F0 range, mean dB intensity and speaking rate in each task was thus obtained for each participant. The results are reported in Section 4.5.1.1 below.

4.5.1.1 Results

Results of the acoustic analysis are reported in this section. The general characteristics of the acoustic data represented by the descriptive statistics are included here. However, the inferential statistics carried out in order to answer the Research Questions are reported in Section 4.5.2.

a) Sentence data

Table 4.2 below shows the F0 mean, F0 range, intensity and speaking rate used when reading the clinically relevant sentences in professional and unprofessional manners.

Table 4.2 Prosodic features according to the attitudes in read sentences

	Male		Female	
	Unprofessional	Professional	Unprofessional	Professional
F0 mean (Hz)	**128.59** (22.96)	**126.34** (18.62)	**236.79** (47.59)	**234.40** (39.80)
F0 range (Hz)	**96.45** (44.13)	**82.30** (39.34)	**199.34** (74.17)	**176.35** (75.99)
Intensity mean (dB)	**64.14** (3.22)	**62.19** (3.24)	**65.74** (5.04)	**62.83** (4.24)
Speaking rate (syllables/second)	**5.62** (0.83)	**5.43** (1.03)	**5.56** (1.09)	**5.34** (1.07)

Note. ***M*** (*SD*); *N* = 16.

From Table 4.2 it seems that the acoustic parameters of prosodic features were adjusted when the participants were instructed to express different attitudinal aspects while reading the sentences. There also seems be a decrease in all four parameters when the sentences read in an unprofessional manner are compared with the same sentences read in a professional manner, regardless of the gender of the participant. However, the question whether the prosodic features are used differently in the two conditions cannot be answered without the inferential statistics reported in Section 4.5.2.

There is another research question whether the two groups of medical students (native vs. non-native English speaking) use the prosodic features differently when reading the same sentences in a professional manner. Table 4.3 reports the descriptive statistics useful for this comparison.

Table 4.3 Prosodic features of native and non-native participants in expressing professional attitudes

	Male		**Female**	
	Native	Non-native	Native	Non-native
F0 mean (Hz)	**131.72** (18.25)	**120.96** (17.83)	**243.05** (44.74)	**226.19** (34.22)
F0 range (Hz)	**95.76** (45.77)	**68.85** (26.48)	**203.88** (87.07)	**150.16** (53.71)
Intensity mean (dB)	**61.68** (2.55)	**62.71** (3.81)	**65.93** (3.29)	**59.89** (2.64)
Speaking rate (syllables/second)	**5.81** (1.04)	**5.06** (0.89)	**5.80** (0.94)	**4.92** (1.02)

Note. ***M*** (*SD*); *N* = 16.

From Table 4.3, in seems that the non-native medical students in this study produced lower values in all four acoustic parameters of prosodic features than their native peers when asked to read the same clinically relevant sentences expressing the same aspects of professional attitudes, with the exception of the male speech intensity. Although these patterns look interesting, answering the question whether or not they differ in using prosodic features to express professional attitudes still needs results of inferential statistics reported in Section 4.5.2.

The F0 contours in the read sentences were analysed using nuclear tone analysis conventions of transcribing intonation patterns. The distribution of different contour types of the sentences read in a professional manner and an unprofessional manner is summarised in Table 4.4 below.

Table 4.4 F0 contours in read sentences

	Unprofessional				Professional			
	Native		Non-Native		Native		Non-Native	
	n	**%**	**n**	**%**	**n**	**%**	**n**	**%**
High fall	21	53.85%	23	57.50%	24	60.00%	27	67.50%
Low fall	2	5.13%	5	12.50%	2	5.00%	2	5.00%
Rise-fall	0	0.00%	1	2.50%	0	0.00%	0	0.00%
High rise	3	7.69%	4	10.00%	6	15.00%	5	12.50%
Low rise	8	20.51%	1	2.50%	6	15.00%	5	12.50%
Fall-rise	5	12.82%	6	15.00%	2	5.00%	1	2.50%
Total	39	100.00%	40	100.00%	40	100.00%	40	100.00%

In both the native and the non-native groups, the most frequently used F0 contour was high fall, and the least frequently used F0 contour was rise-fall. This finding corresponds with previous accounts for British English intonation (see e.g. Cruttenden 1997), especially in that high fall is the default pattern for statements, commands and wh-questions while also contributing to the overtone of being involved. In contrast rise-fall was used on only one occasion in the observation, confirming the account in the literature that this is the least common intonation pattern in British English and is reserved usually for the strongest attitudinal meanings such as impressed or arrogant. In fact none of the instructions for the sentence reading task calls for similar attitudinal meanings.

Interestingly, when asked to read sentences in a professional manner, the non-native group used somewhat fewer rising contours but more falling ones than their native counterpart. Substituting the default falling contours with rising ones can give the impression of softening the tone in conversation (see Collins & Mees 2003). However, the difference in using this strategy in medical communication is very small between the two speaker groups. Indeed the group difference in intonation patterns for the sentence read in a professional manner was not significant ($\chi^2(4) = 0.69$, $p = .95$), nor was it significant for the sentences read in an unprofessional manner ($\chi^2(5) = 8.04$, $p = .14$).It is possible that the non-native speaking students had acquired intonation patterns that are similar to the native patterns, just like their command of vowel and consonant pronunciations. The prosodic differences between the two speaker groups may lie instead in the more gradient acoustic parameters.

b) Role-play data

The consultation role-play task was designed so that the between-subjects effect due to the English language background as well as the within-subjects effect of the consultation task could both be explored. The acoustic analysis was carried out on this data set in a similar way to that on the sentence reading data except that the F0 contour was not analysed. Table 4.5 below reports the parameters of prosodic features used in each of the consultation tasks (initiating the session, gathering information, explaining and planning and closing the session) by the native and the non-native speaking medical students in this study. The males were analysed separately from the females for reasons discussed in Section 4.3.

Table 4.5 Prosodic features in consultation role-play

		INI	GAT	EXP	CLO
	Male				
F0 mean (Hz)	Native	**130.56** (9.97)	**127.43** (8.79)	**109.89** (3.92)	**118.81** (6.06)
	Non-native	**111.73** (9.54)	**114.99** (12.35)	**106.36** (14.79)	**108.62** (9.72)
F0 range (Hz)	Native	**100.66** (21.68)	**86.84** (29.92)	**79.16** (13.92)	**74.90** (19.77)
	Non-native	**73.86** (21.95)	**66.48** (22.44	**71.65** (31.18)	**59.38** (22.03)
Intensity (dB)	Native	**58.70** (4.55)	**58.31** (4.85)	**56.98** (4.82)	**55.29** (3.00)
	Non-native	**61.19** (2.37)	**60.48** (3.90)	**59.34** (5.04)	**59.44** (3.53)
Speaking rate (syllables per second)	Native	**6.15** (0.7)	**5.40** (0.31)	**4.78** (0.36)	**4.94** (0.76)
	Non-native	**4.72** (0.43)	**4.64** (0.26)	**4.53** (0.68)	**5.43** (0.49)
		INI	GAT	EXP	CLO
	Female				
F0 mean (Hz)	Native	**210.51** (18.36)	**209.46** (9.48)	**194.45** (10.67)	**211.71** (12.52)
	Non-native	**217.50** (16.78)	**220.94** (19.07)	**190.40** (24.13)	**203.00** (16.87)
F0 range (Hz)	Native	**189.04** (46.55)	**170.63** (36.59)	**212.41** (29.44)	**181.45** (35.17)
	Non-native	**182.72** (41.4)	**160.48** (61.59)	**167.33** (16.59)	**150.8** (26.62)
Intensity (dB)	Native	**62.20** (2.19)	**61.41** (1.58)	**61.44** (2.57)	**62.09** (2.80)
	Non-native	**59.42** (1.84)	**58.48** (0.66)	**57.59** (1.19)	**55.76** (2.17)
Speaking rate (syllables per second)	Native	**5.33** (0.61)	**5.87** (0.40)	**4.69** (0.53)	**5.82** (0.09)
	Non-native	**5.21** (0.22)	**4.45** (0.78)	**4.27** (0.77)	**4.59** (0.54)

Note. ***M*** (*SD*); *N* = 16; INI = initiating session; GAT = gathering information; EXP = explaining and planning; CLO = closing session.

From Table 4.5, it seems that most of the mean values of the acoustic parameters measured shared the pattern observed in the sentence data when comparing the native and the non-native groups. That is, most values for the non-native group were lower than those for the native group, with the notable exception of the male speech intensity in all consultation tasks, the male speaking rate in closing the session, and the female F0 mean in initiating the session and gathering information. Though interesting, this finding awaits confirmation from the inferential statistics results in order to answer the questions on the effects of English language background and consultation task on prosodic features.

4.5.2 Statistical tests

In order to answer Research Questions 1 and 2 raised in Section 2.7, specific research hypotheses on factors implicated in these questions need to be formulated and tested with inferential statistics, using the sentence reading and the consultation role-play datasets respectively. These are presented as follows.

Hypothesis 1: medical students adjust prosodic features according to the attitudes they are instructed to express.

Hypothesis 2: native and non-native medical students use prosodic features differently to express professional attitudes.

Hypothesis 3: medical students use prosodic features differently according to the consultation task in the simulated consultation.

Hypothesis 4: native and non-native medical students use prosodic features differently in each consultation task.

Hypothesis 5: medical students use prosodic features differently in sentences read in a professional manner and in the simulated consultation.

The dependent variables in the tests were the prosodic features under investigation:

- F0 mean
- F0 range
- Intensity
- Speaking rate

The between-subject factors investigated were: ‘gender’: [male], [female], ‘English language’: [native], [non-native] and ‘training level’: [low], [high].

The within-subject factor in the read sentence dataset was ‘attitudinal instruction’: [unprofessional], [professional].

The within-subject factor in the role-play dataset was ‘consultation task: [initiating consultation], [gathering information], [explaining and planning], [closing consultation].

It can be seen that since each of the four research hypotheses was on one particular factor, univariate analyses could be performed to test the effects of these factors separately. The Kolmogorov-Smirnov test was carried out first to test the normality of the data distribution and it was revealed that not all parts of the acoustic data were normally distributed. In particular the non-normal distribution was found in the sentence reading data set for

- F0 mean of the non-native males for sentences read in a professional manner; D (20) = .25, $p < .01$
- F0 range of the non-native males for sentences read in a professional manner; D (20) = .20, $p < .05$
- speaking rate of the non-native females for sentences read in an unprofessional manner; D (20) = .20, $p < .05$.

A summary of the frequency distribution in the data can be found in Appendix K. Furthermore, log and square root transformations were performed as attempts to transform the distribution to normality, albeit without success, and consequently were not adopted.

Therefore non-parametric tests were regarded appropriate for testing the related hypotheses. Correspondingly non-parametric methods were also chosen for the consultation role-play data set. Given the small sample size and the number of tests needed in this study, choosing non-parametric methods seemed to have good rationale.

It can also be seen that testing Hypotheses 1 and 3 would involve comparing two or more related samples (different speech tasks performed by the same speaker) and that testing Hypotheses 2 and 4 would involve comparing two independent samples (different speakers performing the same speech task). Hypothesis 5 involves the comparison between sentence reading data and role-play data. In order to make the comparison, the five sentences ready by each participant in a professional manner were averaged for each acoustic parameter. The chosen statistical tests are discussed below.

In order to test Hypotheses 1 and 2, Wilcoxon signed ranks tests were performed for the effect of the within-subject factor ‘attitudinal information’ and Mann-Whitney *U* tests were performed for the between-subject factor ‘English language’. It should be noted here that since it is necessary to take into account the physiological difference in the size of vocal folds between men and women, the data file was split by ‘gender’ so that comparisons could be made separately among the males and among the females.

In order to test Hypotheses 3 and 4, Friedman’s ANOVA was performed for the effect of the within-subject factor ‘consultation task’ and Mann-Whitney *U* tests were performed for the effect of the between-subject factor ‘English language’. Again splitting file by ‘gender’ was necessary to analyse the males and the females separately.

Through these tests, the effects of both the within- and between-subjects factors on the prosodic features were revealed. These will be discussed in detail separately.

Additionally, testing Hypothesis 5 can help to understand whether the participants used prosodic features differently in the sentence dataset and in the role-play dataset Comparisons between the two datasets were made by means of Wilcoxon signed ranks tests for each prosodic feature. Results of such comparisons could be interpreted in the light of speakers’ awareness of speech features in when speaking in different styles.

In summary, Hypotheses 1, 3, and 5 were formulated to answer general Research Question 1 whether medical students use prosodic features to convey professional attitudes in medical communication, and Hypotheses 2 and 4 were to answer general Research Question 2 whether native and non-native English speaking medical students differ in the use of prosodic features to communicate in the simulated clinical setting.

4.5.2.1 Results

Table 4.6 below summarises the statistical analysis to test Hypothesis 1. The null hypothesis that medical students used prosodic features in the same way regardless of the attitudes they are instructed to express can be rejected by a significant difference in an acoustic parameter between two conditions (professional vs. unprofessional). However, since this hypothesis involves multiple comparisons in four acoustic parameters of prosodic features, the α-level must be adjusted. The Bonferroni correction was used resulting in the adjusted α-level of .0125.

Table 4.6 Prosodic differences between professional and unprofessional attitudes

Gender		F0 mean	F0 range	Intensity	Speaking rate
Male	z	-0.40	-2.07	-3.93	-1.61
	p [a, b]	.69	.04	< **.01**	.11
Female	z	-0.94	-1.56	-4.35	-1.10
	p [a, b]	.35	.12	< **.01**	.27

a. Wilcoxon Signed Ranks Test
b. 2-tailed

Sentences read in a professional manner were found to exhibit significantly lower acoustic intensity than those same sentences read in an unprofessional manner. Therefore, the null hypothesis can be rejected. In other words, medical students were found to adjust prosodic features according to the attitudes they are instructed to express.

Hypothesis 2 involves testing potential difference in prosodic features between the native group and the non-native group under the same condition, i.e. reading clinically relevant sentences in a professional manner. A significant difference in prosodic features used between the two groups can reject the null hypothesis. Table 4.7 below reports the results of the statistical analysis. Again an adjusted α-level of .0125 by means of the Bonferroni correction was adopted.

Table 4.7 Prosodic differences between native and non-native medical students in expressing professional attitudes

Gender		F0 mean	F0 range	Intensity	Speaking rate
Male	*U*	124.00	131.00	175.00	122.50
	z	-2.06	-1.87	-0.68	-2.10
	p [a,b]	.04	.06	.50	.04
Female	*U*	170.00	121.00	30.00	103.00
	Z	-0.81	-2.14	-4.60	-2.62
	p [a,b]	.42	.03	**< .01**	**.01**

a. Mann-Whitney *U* Test
b. 2-tailed

It can be seen from Table 4.7 that the differences among the male group were not considered significant in this test. However, the differences among the female group in intensity and speaking rate were significant. Therefore although the null hypothesis must be retained for the male medical students in this study, female native and non-native medical students were different in the prosodic features used to express professional attitudes. More specifically, when reading clinical relevant sentences in a professional manner, female non-native medical students used lower intensity and slower speaking rate compared with their native peers.A repeated measures design was used to test Hypothesis 3. Each participant in this study was tested under four conditions (four tasks of consultation) during which the acoustic parameters of prosodic features were measured. A Friedman's ANOVA was performed for each parameter. The results were summarised in Table 4.8 below.

Table 4.8 Comparing prosodic features in the consultation role-play

Gender		F0 mean	F0 range	Intensity	Speaking rate
	χ^2	15.45	8.55	13.65	4.65
Male	*df*	3	3	3	3
	p	**< .01**	.04	**< .01**	.20
	χ^2	10.35	8.70	4.05	14.85
Female	*df*	3	3	3	3
	p	.02	.03	.26	**< .01**

Note. Adjusted $\alpha = .0125$ by Bonferroni correction for multiple comparisons

As Table 4.8 shows, significant difference in prosodic features was found among both the males and the females, rejecting the null hypothesis. More specifically, F0 mean and intensity differed significantly among the males for different consultation tasks, but among the females only speaking rate differed significantly. This suggests an effect of consultation task on the use of prosodic features when the medical students in this study role-played in the consultation with the simulated patient.

Post hoc analysis was performed by choosing the pairwise comparison option for Friedman's ANOVA in the statistical software package SPSS. The following comparisons showed significant differences (with α adjusted by software).

- For the males, F0 mean in explaining/planning was significantly lower than that in initiation ($p = .04$), F0 mean in explaining/planning was significantly lower than that in gathering information ($p = .001$), intensity in closure was significantly lower than that in gathering information ($p = .02$), and intensity in closure was significantly lower than that in initiation ($p = .01$).
- For the females, speaking rate in explaining/planning was significantly lower than that in gathering information ($p = .02$), speaking rate in explaining/planning was significantly lower than that in closure ($p = .02$), and speaking rate in explaining/planning was significantly lower than that in initiation ($p < .01$).

Another between-subjects comparison was needed for testing Hypothesis 4. In doing so prosodic features of non-native participants were compared with those of native participants in each consultation task. Mann-Whitney *U* tests were carried out for each acoustic parameter to test the effect of the between-subjects factor 'English language'. The results of these tests are shown in Table 4.9. As can be seen, none of the results was found to reach the adjusted α-level of .0125 (Bonferroni correction). Therefore the null hypothesis must be retained.

Table 4.9 Results of Mann-Whitney *U* tests

		Male			Female		
		U	*z*	*p*	*U*	*z*	*p*
F0 mean	INI	0.00	-2.31	0.03	10.00	0.58	0.69
	GAT	2.00	-1.73	0.11	12.00	1.16	0.34
	EXP	5.00	-0.87	0.49	8.00	0.00	1.00
	CLO	2.00	-1.73	0.11	4.00	-1.16	0.34
F0 range	INI	2.00	-1.73	0.11	7.00	-0.29	0.89
	GAT	4.00	-1.16	0.34	6.00	-0.58	0.69
	EXP	7.00	-0.29	0.89	1.00	-2.02	0.06
	CLO	6.00	-0.58	0.69	3.00	-1.44	0.20
Intensity	INI	11.00	0.87	0.49	2.00	-1.73	0.11
	GAT	11.00	0.87	0.49	0.00	-2.31	0.03
	EXP	11.00	-0.29	0.19	1.00	-2.02	0.06
	CLO	14.00	1.73	0.11	1.00	-2.02	0.06
Speaking rate	INI	1.00	-2.02	0.06	8.00	0.00	1.00
	GAT	0.00	-2.31	0.03	0.00	-2.31	0.03
	EXP	7.00	-0.29	0.89	6.00	-0.58	0.69
	CLO	11.00	0.87	0.49	0.00	-2.31	0.03

Note. Adjusted α = .0125 by Bonferroni correction for multiple comparisons

However the box plots in Figures 4.1-4.4 below seem to suggest that in some consultation tasks at least, the acoustic properties of speech prosody were quite different.

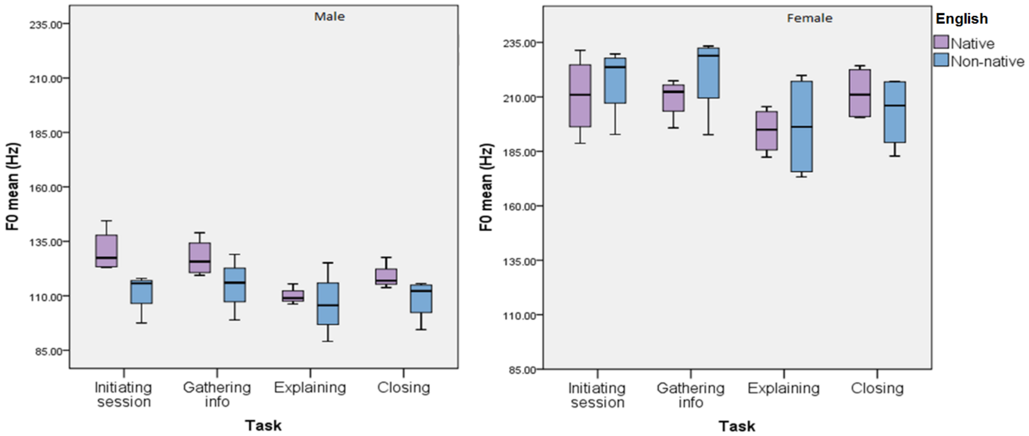


Figure 4.1 F0 mean of native and non-native medical students in the consultation role-play

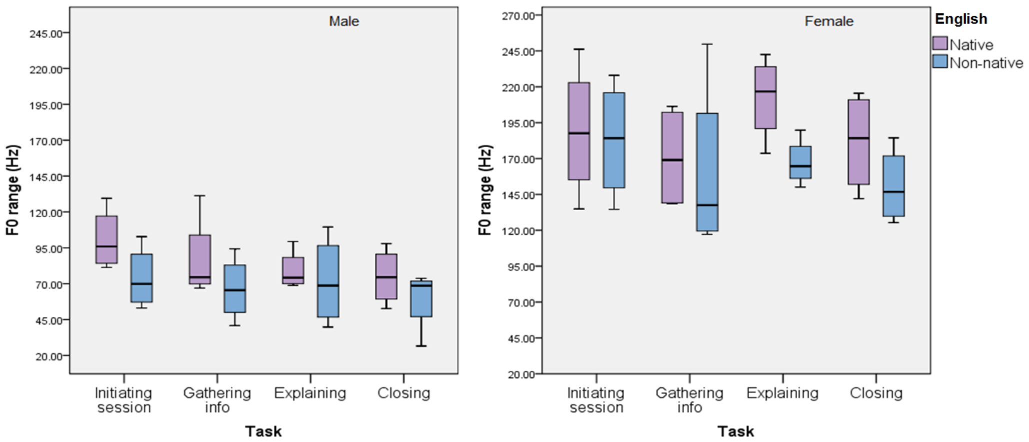


Figure 4.2 F0 range of native and non-native medical students in the consultation role-play

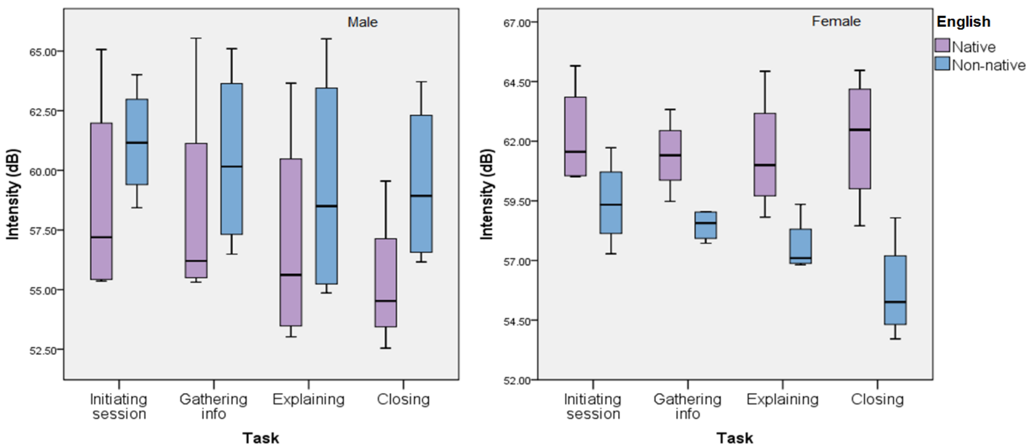


Figure 4.3 Speech intensity of native and non-native medical students in the consultation role-play

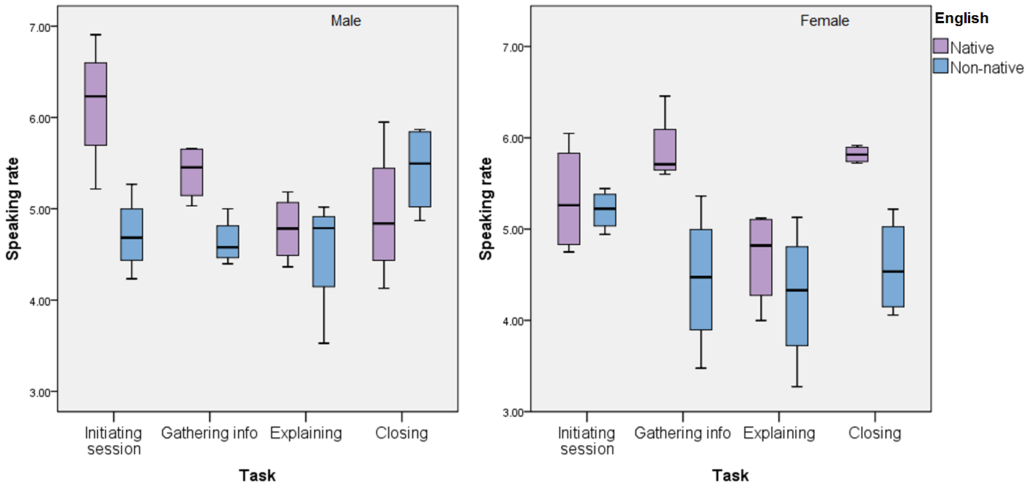


Figure 4.4 Speaking rate of native and non-native medical students in the consultation role-play

The box-plots in Figures 4.1-4.4 above serve as visual aids to the comparison. According to the figures, some prosodic differences seem quite apparent. In particular, in males these are F0 mean in initiating the session (U = 0.00, z = -2.31, p = .03) and speaking rate in gathering information (U = 0.00, z = -2.31, p = .03). In females these are intensity in gathering information (U = 0.00, z = -2.31, p = .03), speaking rate gathering information (U = 0.00, z = -2.31, p = .03) and speaking rate in closing the session (U = 0.00, z = -2.31, p = .03). One possibility could be due to the conservativeness of the Bonferroni correction resulting in too strict an adjusted α-level of .0125. The critical values of the U statistic for a sample of 4 members in each group according to Milton (1964) can only yield significance at the level as low as p = .05. Therefore the U statistic could not show any significant difference at the adjusted α-level in this study, indicating a possible Type II error. However, the available alternative to use t-tests ignoring the assumption on normal distribution is rather poor. Based on the rationale to be on the side of caution in this exploratory study, it was decided that the above results should be accepted as to suggest that though potentially important, the prosodic differences between the two speaker groups in each consultation task were not statistically significant. This conservative decision was considered to be more justifiable than adopting a parametric method while violating the assumption to find “support” for Hypothesis 4. Nevertheless, Figures 4.1-4.4 suggest an explanation that in adjusting prosodic features in the consultation role-play, the consultation tasks might be a more important factor and that the non-native medical students and their native peers might have different patterns of doing this.

Last but not least, Hypothesis 5 was tested by means of Wilcoxon Signed Ranks tests. The null hypothesis that the prosodic features of the medical students in the consultation role-play do not differ from those in the read sentences to express professional attitudes was rejected due to significant (α = .013, Bonferroni correction) differences found in some acoustic parameters of prosodic features of the females. In particular, F0 mean was lower in the read sentences than in the consultation role-play (difference between means = -24.92, 95% CI = -41.30 to -8.54, z = -2.52, p = .01) and intensity was lower in the read sentences than in the consultation role-play (difference between means = -3.17, 95% CI = -4.83 to -1.50, z = -2.52, p = .01). It would be prudent to note that though this finding is potentially important, the technique used in generating the observations is not ideal. The observations used in the comparison are in fact averaged values of the original measurements. More specifically, the five sentences and the four consultation tasks of each participant were averaged respectively to make the comparison possible. Thus some potentially significant differences in the original measurements might have been lost in the process. Nevertheless using these averaged observations was a necessary compromise because there was no other way to control for possible confounders in the comparison (sentence type in read sentences and speech content in role-play, but see Chapter 5 for a way to control for these within a subset of the data). At this stage of the investigation, it was not certain what the prosodic differences between sentences read in a professional manner and the consultation role-play could mean in terms of potential patient experience. This tentative finding must await support from a perceptual study so that interpretation can be made with regards to the listener's perception of professionalism.

4.6 Summary

Speech data collected in the acoustic study enabled examination of prosodic features used in interactions in the clinical context. The design adopted made it possible to examine the effects of attitudes, consultation tasks in the consultation role-play and

English language backgrounds of the medical students on the acoustic property of prosodic features.

It was found that the medical students in this study used prosodic features differently when they were instructed to convey two sets of attitudes (professional vs. unprofessional) in read sentences and when they carried out four tasks (initiating session, gathering information, explaining and planning and closing session) of consultation in the role-play with the simulated patient. When reading the clinically relevant sentences in a professional manner, the native and the non-native English speaking students were found to differ in the prosodic features used. However, similar differences could not be confirmed in each consultation task in the role-play, which might be due to the limitation of the chosen statistical methods. Taken as a whole, results in the acoustic study answered Research Questions 1 and 2 positively. On the one hand, prosodic features are used by medical students to convey attitudinal messages related to professionalism in medical communication. On the other hand, there are differences between native and non-native English speaking medical students in the use of prosodic features in the simulated clinical interaction. Results of the acoustic study in general support findings of previous research on the attitudinal function of prosodic features in speech communication (see Scherer et al. 2003) and earlier reports on the prosodic differences between native and non-native speakers of English (Anderson-Hsieh et al. 1992, Binghadeer 2008, Wennerstrom 1994). Comparisons between findings of the current study and those in previous literature are discussed in Chapter 6.

These results on the acoustic investigation of prosodic features are important on their own merits due to their implication for research, education and practice. However, Research Question 3 whether prosodic features influence the listeners' perception of the speakers' professional attitudes in medical communication remains unanswered, and awaits the second stage of the current study, the perceptual investigation. More importantly, although some prosodic differences were not found to be significant in the acoustic study, they might still have significant effects on perceived attitudes in the clinical interaction. If that is later found to be true, prosodic features of healthcare professionals can be shown to have significant influence on service users' experience of the health communication. The perceptual study is reported in Chapter 5 with general discussion and concluding remarks to follow.

Chapter 5 The perceptual study

So far the acoustic investigation has been able to answer two of the three general Research Questions raised in the current study. The medical students were found to use speech prosody to convey professional attitudes in the simulated clinical interaction. Results of the acoustic analysis also suggest that the native and the non-native English speaking medical students used prosodic features differently in the simulated clinical interaction. Chapter 4 reports the details of that investigation. However, from those findings it is still not certain if differences in prosodic features can influence listeners' perception of attitude in medical communication. Therefore Research Question 3 needs to be answered with evidence from a perceptual investigation. In order to do so, the association between the acoustic properties of the prosodic features used and the level of professionalism perceived by listeners needs to be explored in the medical context. This can be made possible by presenting controlled vocal stimuli (acoustic properties of which are known) to untrained listeners who rate the extent of attitudes they perceive on the specifically developed scale in the experiment environment. By adopting this method it is also possible to understand the contribution of speaker attributes such as gender, language background and training level to the attitudes perceived by listeners.

5.1 Ethics

An application for ethical approval to carry out this perceptual experiment was submitted to Faculty of Medicine and Health Sciences Ethics Committee in September 2012. In particular, the Committee was informed of the following ethical considerations in the protocol.

- The experiment in the study is completely safe and non-invasive and pose no risk whatsoever to any participant.
- All participants will be recruited on a voluntary basis. Information provided by volunteers who are not selected for the study will be destroyed. Participants can choose not to participate and withdraw without giving a reason. Should a participant withdraw from the study, all his/her data will be

destroyed. Participants will give explicit informed consent prior to commencing the experiment.

- Personal information collected in the screening questionnaire is used for selection purposes only, and will constitute no part of the data used in the study. The perceptual data will be collected anonymously, and no participant will be identified. All paperwork will be stored in a locked filing cabinet on campus and digital files on password protected UEA server space. After the data collection, any personal information will be removed.
- In accordance with the Data Protection Act 1998, if a participant does not consent to the use of his/her data for future research, the data will be archived for a maximum of 5 years before being destroyed. If a participant consents to the use of his/her data for future research by initiating the corresponding box in the consent form, the data will be kept securely with the same level of confidentiality for 5 years after the end of the study and will be destroyed after that.

Supporting documents submitted in addition to the protocol included UEA Survey Officer's approval (to send information to UEA staff and students in order to recruit them to the study), recruitment flyer, participant information sheet, recruitment screening questionnaire and consent form (Appendices J-O). Other materials relevant to the experiment, namely, transcripts of the vocal stimuli, participant workbook containing task instruction and rating scale (Appendices P and Q) were also submitted with the application to the Committee. Full ethical approval (Appendix J) was obtained in November 2012.

5.2 Study design

One common method in experimental psychology involves presenting human subjects with carefully manipulated stimuli (audio, visual, etc.) to measure and analyse their responses. The differences in the stimuli serve as the conditions in such experiment (Kantowitz 2009). The design of the current study uses this method.

The stimuli in the current study were recorded speech samples from simulated clinical interactions, and the responses were the ratings of criteria thought to contribute to the level of professionalism in medical communication that the listeners perceived from the speech samples. This experimental design enables the examination of the effect of the prosodic features used in the simulated clinical interaction under investigation on the level of perceived professionalism in communication. More importantly, since the acoustic properties of the prosodic features have already been investigated in the first stage of the study (see Chapter 4), if a significant effect of prosodic features is found, the finding will provide positive answer to the overarching research question whether speech prosody matters in health communication.

In addition to the prosodic features under investigation, the between-subjects effects of a speaker's gender, English language background and training level as well as the within-subjects effects of stimulus type were also considered. Consequently, statistical methods adopted needed to be suitable for examining these factors. These are discussed in Section 5.6.

According to the factors considered in the current study, five specific research questions are formulated as follows.

1. Do listeners give different scores on a professionalism in communication scale (henceforth professionalism scale) to sentences read out in a professional manner and those in an unprofessional manner?
2. Do listeners give different professionalism scale sores to sentences read out in a professional manner and spontaneous utterances from simulated consultation?
3. Are there any differences among professionalism scale scores that listeners give to the utterances from different consultation tasks?
4. Are there any differences in professionalism scale scores between the following groups of medical students in the same consultation task?
   (a) male and female
   (b) native and non-native English speaking
   (c) trained to a low level (Years 2 and 3) and to a high level (Year 5)
5. Are listeners' professionalism judgement affected by the following prosodic features in the utterances for both native and non-native speaking medical

students?

(a) F0 mean
(b) F0 range
(c) mean dB intensity
(d) speaking rate

5.3 Participant recruitment

Several issues were considered before the recruitment of research participants: the research questions, the research design, and the availability of eligible subjects. Since the study did not plan to examine the effect of listener characteristics (in this case gender and employment status) on their perception, such potential confounders needed to be controlled for in the research design while maintaining a representative sample so that the results would be generalisable to the study population. One way of doing this would be to achieve a perfectly balanced sample with equal distribution in these factors. This could also be more realistically and more efficiently achieved by using a Latin square design to assign experiment stimuli and adopting robust techniques in statistical analysis. The current study selected this second method.

A Latin square uses a matrix of equal number of rows and columns in which each different experiment condition appears once only in each row and in each column. The Latin square design ensures that each treatment (experiment condition) will be administered once in every trial (Edward 1951). Therefore a Latin square design is a powerful design that can account for the effects of two other factors at one time in addition to the effect of the treatment (Kirk 2003).

For each specific research question, therefore, the sample size equalled the number of participants recruited in the study multiplied by the number of stimuli used, thanks to the Latin square design. For example, recruiting 32 participants would yield 320 observations to answer Research Question 1 above (32 listeners × 5 sentences read out in an unprofessional manner + 32 listeners × 5 sentences read out in a professional manner). The Latin square used also required at least 16 participants from the whole sample. However, a power calculation could not be carried out for this study. This was mainly because previous literature did not provide information on the likely extent (large, medium or small) of the effects of prosodic features on attitudinal perception. Therefore there were not sufficient grounds to carry out the calculation. Furthermore power calculation was not carried out in similar studies (e.g. Banse & Scherer 1996) reviewed in Chapter 2.

For ethical considerations, the experiment session duration was designed to be within 60 minutes. This was because all sessions took place during daytime and the majority of UEA staff participants could not be expected to be available for longer than that. Consequently the actual duration of a session lasted between approximately 40 minutes and 60 minutes. Most participants recruited for both the test and retest conditions which took place consecutively also used approximately 60 minutes in total. A Latin square design made it possible within this limited time that

- Each participant should hear stimuli from all speakers.
- Each participant should hear all types of stimuli.
- For each participant the combination of speakers and types of stimuli should be random.

The method of assigning stimuli to participants is shown in Figure 5.1. This calls for a minimum of 16 participants so that all 256 stimuli could be heard at least once. Each one participant recruited after this would enable 16 stimuli to be heard more than once, and a multiple of 16 participants recruited would render all 256 stimuli heard once more.

Stimulus	Speaker															
	I	**II**	**III**	**IV**	**V**	**VI**	**VII**	**VIII**	**IX**	**X**	**XI**	**XII**	**XIII**	**XIV**	**XV**	**XVI**
1a	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1b	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	1
2a	3	4	5	6	7	8	9	10	11	12	13	14	15	16	1	2
2b	4	5	6	7	8	9	10	11	12	13	14	15	16	1	2	3
3a	5	6	7	8	9	10	11	12	13	14	15	16	1	2	3	4
3b	6	7	8	9	10	11	12	13	14	15	16	1	2	3	4	5
4a	7	8	9	10	11	12	13	14	15	16	1	2	3	4	5	6
4b	8	9	10	11	12	13	14	15	16	1	2	3	4	5	6	7
5a	9	10	11	12	13	14	15	16	1	2	3	4	5	6	7	8
5b	10	11	12	13	14	15	16	1	2	3	4	5	6	7	8	9
INI	11	12	13	14	15	16	1	2	3	4	5	6	7	8	9	10
GAC	12	13	14	15	16	1	2	3	4	5	6	7	8	9	10	11
GAO	13	14	15	16	1	2	3	4	5	6	7	8	9	10	11	12
EXP	14	15	16	1	2	3	4	5	6	7	8	9	10	11	12	13
ADV	15	16	1	2	3	4	5	6	7	8	9	10	11	12	13	14
CLO	16	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

Note. This square is used to assign stimuli to 16 participants at a time. Participant 1 will be the listener [1] who is presented with the following stimuli to judge: Sentence 1a from Speaker I, Sentence 1b from Speaker XVI, Sentence 2a from Speaker XV etc.. Further participants, 16 at a time, can be assigned in the same way repeating the above procedure.

Figure 5.1 16×16 Latin square used in assigning stimuli to participants

The recruitment information was distributed by means of flyers (Appendix L) on UEA campus, Blackboard announcement and UEA e-bulletin. Subjects were offered a chance to win one of the two £50 Amazon vouchers in a prize draw upon the completion of data collection as an incentive. Interested volunteers were screened for eligibility with a questionnaire (Appendix P) against the following criteria:

Inclusion criteria

- UEA students and employees who are native speakers of English

Exclusion criteria

- Non-native speakers of English
- Students and academic staff in the Faculty of Medicine and Health Sciences (FMH)
- Those with self-reported hearing or speech difficulty

- Those aged over 65 and therefore at increased risk of age-related hearing loss (see Gates & Mills 2005)
- Those with prior training in phonetics or speech sciences

These constraints were necessary due to several reasons.

1. The population studied were potential healthcare service users who are native speakers of English and do not require a hearing-aid or assistive communication device.
2. Students and employees of UEA were a diverse community for the study so that the results from this sample could be generalised.
3. Professional knowledge of healthcare professionals (FMH students and academic staff) could potentially confound the results.
4. Knowledge of phonetics or speech sciences could affect the attitudinal judgments of speech sounds and thus confound the results.

Volunteers who met the criteria were then contacted to confirm their availability for the experiment session, which took place in the School of Allied Health Professions (AHP) Communication Lab. A total of 54 participants, with a mean age of 28.54 (*SD* = 11.31) years were recruited for the study. This mean age should be interpreted with reference to the age ceiling of inclusion, as well as the diversity of the study population. A mean age lower than this was reported in a similar, previous study with undergraduate students as participants (Banse & Scherer 1996). Table 5.1 below summarises their basic demographic information, with additional details reported in Appendix U. Most recruited participants, as expected, speak varieties of Standard English (e.g. SSBE), due to the influence of the university community. In consequence, such research design meant that the compared with the non-native stimuli, the native stimuli (from SSBE speaking medical students) would present an in-group advantage for most participants in the perceptual experiment, due to their familiarity with Standard English accent and prosody. This was a planned effect and would not weaken the interpretation of the results, as the comparison concerned is between native and non-native English prosody. Prior to the start of their experiment sessions, informed written consent was obtained from all participants.

Table 5.1 Number and distribution of perceptual study participants

	Male		Female	
	n	age	*n*	age
Student	18	**21.72** (3.43)	15	**23.67** (8.08)
Staff	8	**33.88** (10.26)	13	**42.45** (10.59)

Note. ***M*** (*SD*); N = 54.

5.4 Data collection

5.4.1 Stimuli materials

Two sets of stimuli consisting of recordings from 16 medical students were used for this experiment, namely read sentences and consultation passages. These were recorded in the previous stage of the study on UEA campus (see Chapter 4).

The first stimuli set consists of the following five clinically relevant sentences:

1. *How are you?*
2. *How much alcohol do you drink in a week?*
3. *I'm sorry to hear that.*
4. *Do you have any thoughts on this?*
5. *Tell me more about that.*

These were read by each participating medical student twice:

(a) once in an unprofessional manner, and
(b) once in a professional manner.

The second stimuli set, namely the consultation stimuli were also selected from the transcribed recordings produced during the previous stage of the study. These were according to the tasks they carried out in the simulated consultation, which are summarised as follows.

1. INI: initiating the session
2. GAC: gathering information with closed questions
3. GAO: gathering information with open-ended questions

4. EXP: providing explanation
5. ADV: providing advice
6. CLO: closing the session

The selection of consultation stimuli firstly involved examining the transcripts of simulated consultations, which became available from the acoustic study. Previously in the acoustic study, the consultations were already segmented into utterances (individual turns), with utterances of the simulated patient and overlapping turns removed, and were labelled according to the consultation tasks (see Section 4.5.1). Secondly, for each consultation task by every medical student, corresponding utterances were aggregated, usually in the order of their occurrence in the consultation, until a total duration of approximately 10 seconds could be obtained. Some aggregated consultation stimuli, however, were shorter in duration (especially the CLO ones) because of the brevity of the representing utterances in the consultation. Furthermore, in some cases, the order in which the utterances appeared in the consultation was altered in order to render the stimuli more natural sounding.

It should be noted that utterances in the GAC and the GAO stimuli were carrying out the task of gathering information according to the Calgary-Cambridge guide. In the textbook (Silverman et al. 2005), these are not regarded as separate stages in a structured consultation, nor were they previously treated separately in the acoustic investigation. However, closed and open-ended questions can have different linguistic structures (e.g. yes-no question vs. wh-question) which can in turn result in different acoustic properties of prosodic features. Therefore in this experiment, they were examined as separate types of stimuli. Similarly both EXP and ADV belong to the stage of explanation and planning in the Calgary-Cambridge model, but were treated as separate types of stimuli due to different sentence types used in the utterances (ADV can take the form of a command). These were necessary because of the need to control for the effects of sentence type and to a lesser extent speech content. In the previous stage of the study, i.e. the acoustic investigation, the control was able to be less strict because the acoustic measures were averaged for all utterances within each consultation task. In the current stage of the study however, the average duration of a stimulus (8.98 seconds for consultation stimuli and 1.27 seconds for sentence stimuli) is much shorter compared to that of all utterances in a consultation task under the acoustic investigation. Therefore the stimuli which were produced this way to represent each consultation task would be comparable across speakers, which is a reasonable level of control over sentence types and speech content (absolute control over verbal content would be neither realistic nor necessary in simultaneous speech for the purpose of the current research, see also Sections 3.2.4 and 3.3.2). Further details are given in Appendix R, with a worked example to illustrate the procedure of stimuli preparation. Appendix R also included the transcript of a complete consultation, following standard conversation analysis conventions which provide some details (e.g. pauses and breathing) beyond the textual level, albeit with minor adaptions (e.g. to ease reading, turns are not broken into single lines and are not numbered).

Anonymity of the speakers was needed to ensure confidentiality in the investigation, so a decision must be made whether the names of the medical students should be removed from the INI stimuli that contained their self-introduction to the simulated patient. These were decided to be retained in the stimuli but manipulation must be implemented so that the names could not be recognisable. This is because the speakers, i.e. UEA medical students were trained to introduce themselves at the initiation stage of the consultation, and removing the names would compromise the integrity and coherence of these utterances. Replacing the concerned segments with non-speech sounds e.g. beeping was also discarded, since this would also corrupt the acoustic properties of the original utterances. A viable solution is to acoustically mask the content of the affected speech, which is a practice that has been long accepted (e.g. Scherer et al. 1984, Ladd et al. 1985, Johnson et al. 1986). Such masking techniques that destroy the speech content while keeping some prosodic features included low-pass filtering, random splicing and reversing (Johnson et al. 1986). Low-pass filtering would keep the fundamental frequency intact, but would mask acoustic energy above the chosen threshold, partly altering the acoustic intensity, and random splicing would partly alter the tempo among others (Johnson et al. 1986), so these techniques would not be suitable for the purpose of the current study. Reversing however, would not alter fundamental frequency, intensity and speaking rate in the stimuli (Johnson et al. 1986), all of which must not be tempered in the current study. Furthermore, Johnson et al. (1986) suggested that reversed speech could affect slightly but would not obstruct the recognition of certain basic emotions (e.g. sadness), which was not an issue of concern for the current study. Therefore reversing was adopted in the current study as a suitable masking technique.

The acoustic signals of the names present in the stimuli were identified using techniques discussed in the previous chapter (boundaries determined with the first and last zero crossings in the waveform patterns concerned, aided by spectrographic cues, where necessary). Then, the identified segments were reversed using the "reverse selection" command in the *Praat* object window. The resulting stimuli rendered the names no longer recognisable, while the reversed segments would still sound like human speech, albeit almost as though from an unknown language.

Having selected and suitably prepared the stimuli, the transcripts of these stimuli were first piloted on a group of 20 judges who were asked to report any attitudinal information that they perceived from the transcript texts to be anything other than neutral. These judges including students and staff at UEA or its affiliated research institute as well as friends outside the University community are all native speakers of English. These native judges, though essentially a convenience sample, were a cross-section of a wide community similar to the study population of the current investigation. Therefore their experience could represent to some extent the general judgment of potential service users. This was a necessary pilot to find out if any underlying attitudinal message in the stimuli texts could potentially confound the experiment which only intended to investigate attitudinal aspects perceived from speech prosody and not the verbal content.

None of the stimuli transcripts were considered to generate perceptions of attitude through the wording. In most of the cases, the judges either did not perceive any attitude from the stimuli transcripts, or they did not agree on attitudinal meanings they perceived from the stimuli transcripts. Just two of the stimuli transcripts led to agreed attitudinal perception but only three people identified these stimuli as having a specific attitude under current investigation. Therefore all selected stimuli were regarded as appropriate for the purpose of the experiment.

5.4.1.1 Acoustic characteristics of the stimuli

The voice stimuli were essentially a subset of the original acoustic data used in the previous stage of the study, and their acoustic properties were analysed in the same way as in the acoustic investigation. The sentence reading data set was kept in its entirety when used as the voice stimuli in this study. Reports of their acoustic properties can be seen in Section 4.5.1.1. While originally embedded within the consultation role-play data set of which the acoustic properties have also been reported in Section 4.5.1.1, characteristics of the chosen simulated consultation stimuli need to be described separately before perceptual responses to them can be examined. Table 5.2 below reports the general acoustic characteristics of the simulated consultation stimuli.

Table 5.2 Simulated consultation stimuli for the perceptual experiment

Prosodic feature	**English**	**Male**	**Female**
F0 mean (Hz)	Native	**121.46** (12.77)	**203.37** (22.77)
	Non-native	**109.63** (13.68)	**209.43** (24.56)
F0 range (Hz)	Native	**113.42** (32.65)	**224.74** (55.08)
	Non-native	**97.16** (41.98)	**216.57** (62.80)
Intensity (dB)	Native	**57.21** (3.91)	**61.56** (2.21)
	Non-native	**59.26** (3.60)	**57.73** (2.19)
Speaking rate (s.p.s.[a])	Native	**5.22** (0.93)	**5.30** (0.81)
	Non-native	**4.80** (0.69)	**4.38** (0.83)

Note. ***M*** (*SD*); *N* = 96.
a. s.p.s: syllables per second

It seems from Table 5.2 that in terms of the prosodic features in the stimulated consultation stimuli, the comparisons between the native and the non-native medical students represented very similar patterns to those in Table 4.5 (of the entire consultation role-play data set). Therefore this subset of data can be said to share common prosodic characteristics as the original acoustic data set. More specifically within this set of stimuli, non-native speakers were found to have lower values for all acoustic measures of prosodic features with the only exception of female F0 mean and male speech intensity.

According to Gussenhoven (2004), higher F0 could be a marker of femininity, although social values of gender differences may be varied, for example, Japanese women use a higher average F0 than American women (Ohara 1992 cited in Gussenhoven 2004). A tentative explanation for the difference in female F0 mean between the native and the non-native medicals students in this stimuli set may be that the non-native females were making more use of F0 to mark femininity than the native females. However previous research did not consider marked femininity / masculinity as part of professionalism in medical communication, so no further discussion on expressing gender in speech prosody is needed for the current study. The interpretation on the difference in male speech intensity between the native and the non-native students in this stimuli set is more difficult, since previous research did not shed much light on such prosodic difference with regards to professionalism. One possibility may be that the non-native medical students in this study used this feature to express, among others, the attitudinal aspect to do with confidence, while their native peers did not. Consequently the non-natives adjusted their intensity setting in professional communication to a higher level than the natives. This interpretation must await support from perceptual evidence.

Theoretically, the structure consultation would proceed in a linear order, with each type of consultation task performed one after another. In reality, some parts of the consultation structure could be very often intertwined. For example, when giving advice, the (student) doctor might ask further questions to check the (simulated) patient's expectation. Prosodic features during the consultation might also be adjusted by the doctor according to the type of task. Since consultation task was a factor for the perceptual experiment, it was considered useful to plot the prosodic features according to the consultation tasks using a line chart (Figure 5.2) to identify any pattern of adjustment during the structured consultation.

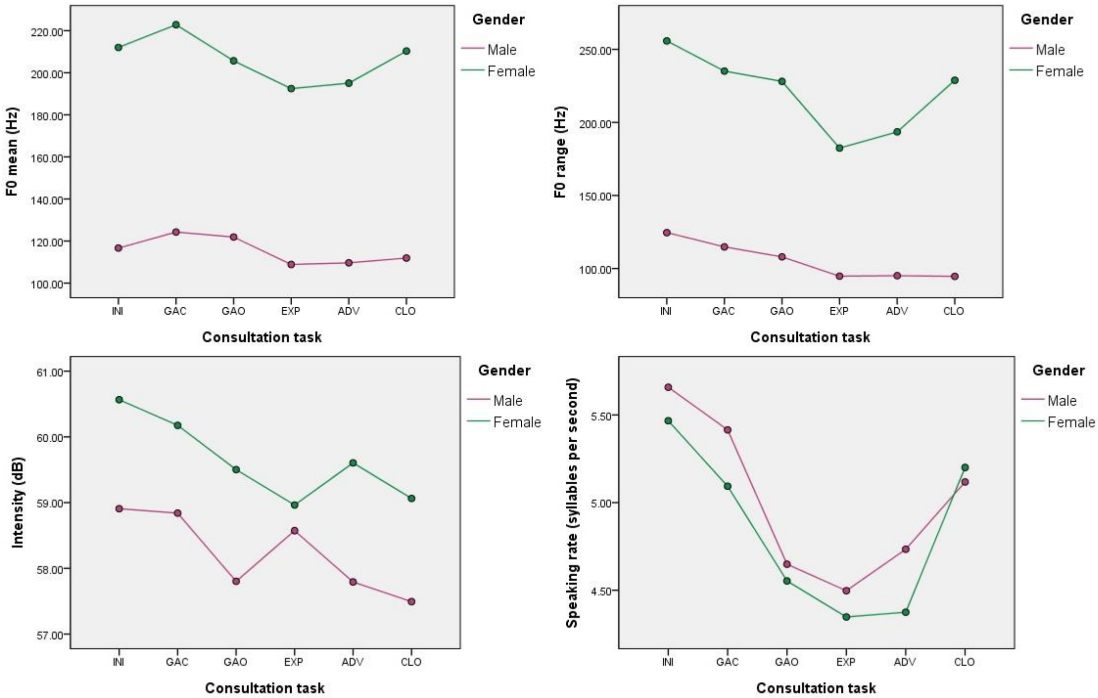


Figure 5.2 Acoustic characteristics of the prosodic features according to the consultation task in the stimuli

In Figure 5.2 it can be seen that both the males and the females showed some similar patterns of prosodic features changing through the consultation tasks. This observation is particularly obvious for F0 mean and speaking rate. Friedman's ANOVA confirmed this observation (the effect on F0 mean was significant for both the males and the females at the adjusted α level of .008 by means of Bonferroni correction for multiple comparisons). This agrees with the finding on the effect of the consultation task on prosodic features in the analysis of the entire consultation role-play data set (see Section 4.5). It will be very interesting if these patterns of prosodic features adjustment are also found to have an effect on perceived professionalism in the perceptual experiment.

Since the perceptual experiment also needs to examine the effects of the gender, the English language background of the medical student in the stimulus on the perceived professionalism with the effect of consultation task adjusted for, it would be useful to know the acoustic characteristics of the prosodic features according to these factors in each consultation task. Figures 5.3-5.7 below show the F0 mean, F0 range, mean dB intensity and speaking rate by consultation type in the stimuli according to gender and English language respectively.

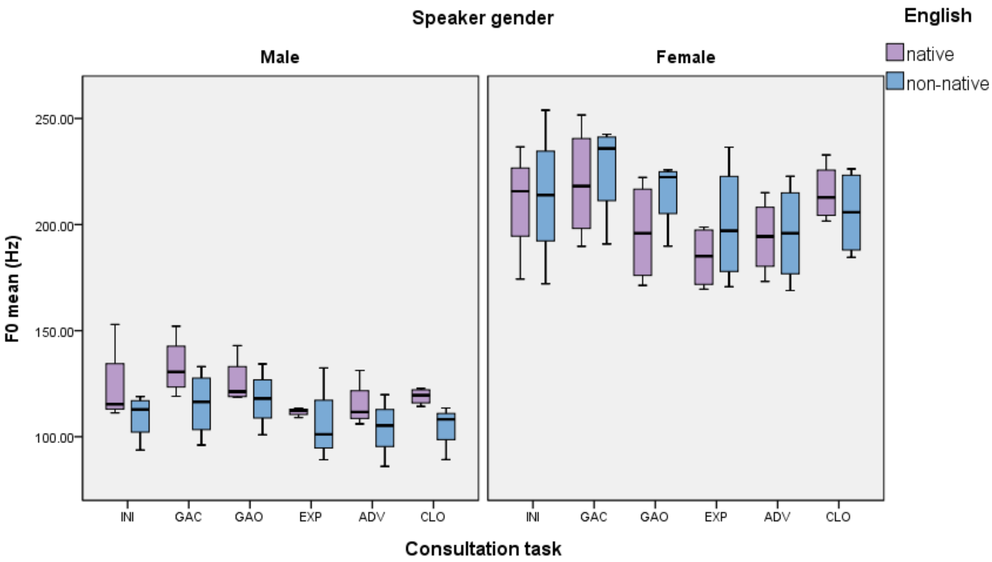


Figure 5.3 F0 mean in the consultation stimuli according to gender and English language background by consultation task.

INI: initiating the session; GAC: gathering information with closed questions; GAO: gathering information with open-ended questions; EXP: providing explanation; ADV: providing advice; CLO: closing the session

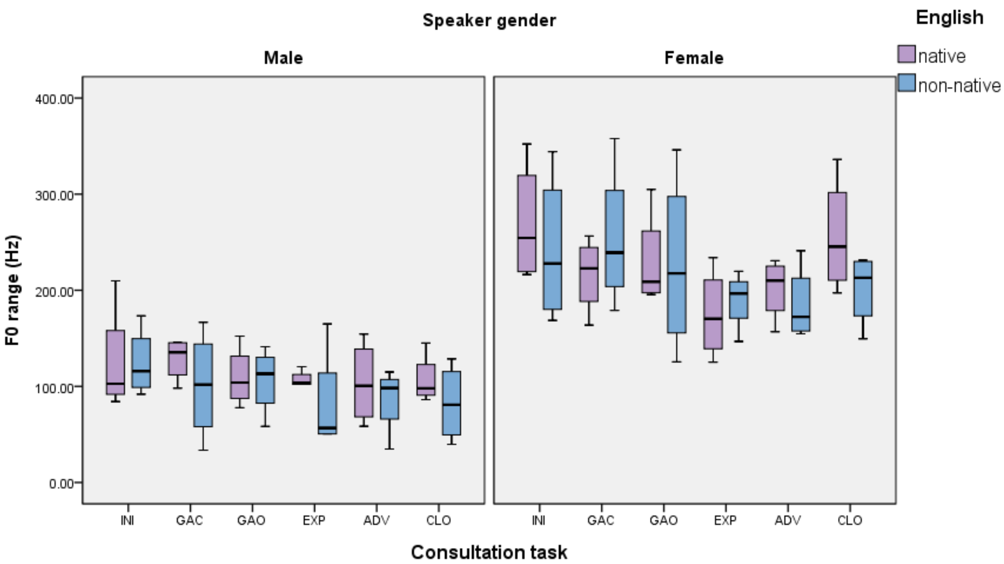


Figure 5.4 F0 range in the consultation stimuli according to gender and English language background by consultation task.

INI: initiating the session; GAC: gathering information with closed questions; GAO: gathering information with open-ended questions; EXP: providing explanation; ADV: providing advice; CLO: closing the session

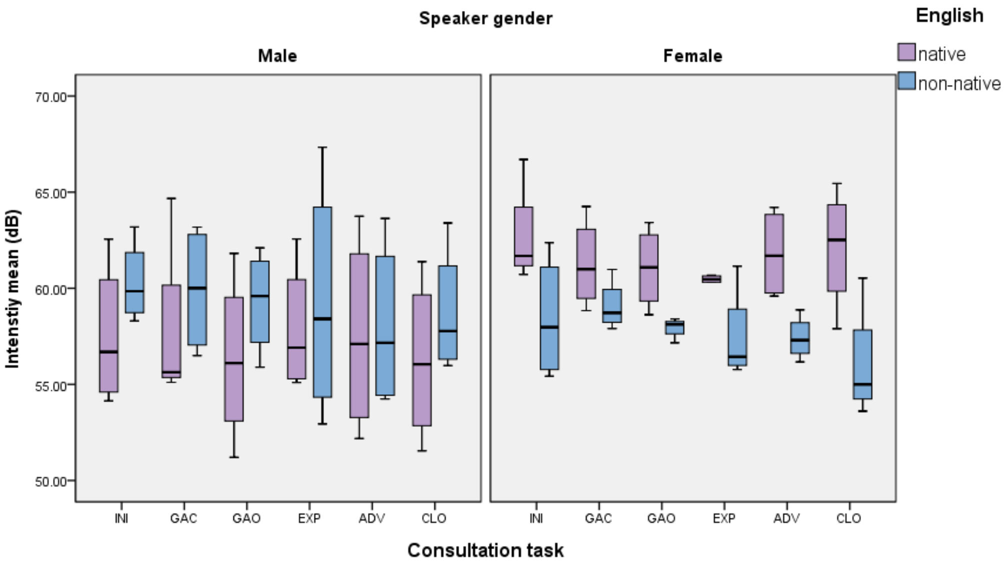


Figure 5.5 Intensity in the consultation stimuli according to gender and English language background by consultation task.

INI: initiating the session; GAC: gathering information with closed questions; GAO: gathering information with open-ended questions; EXP: providing explanation; ADV: providing advice; CLO: closing the session

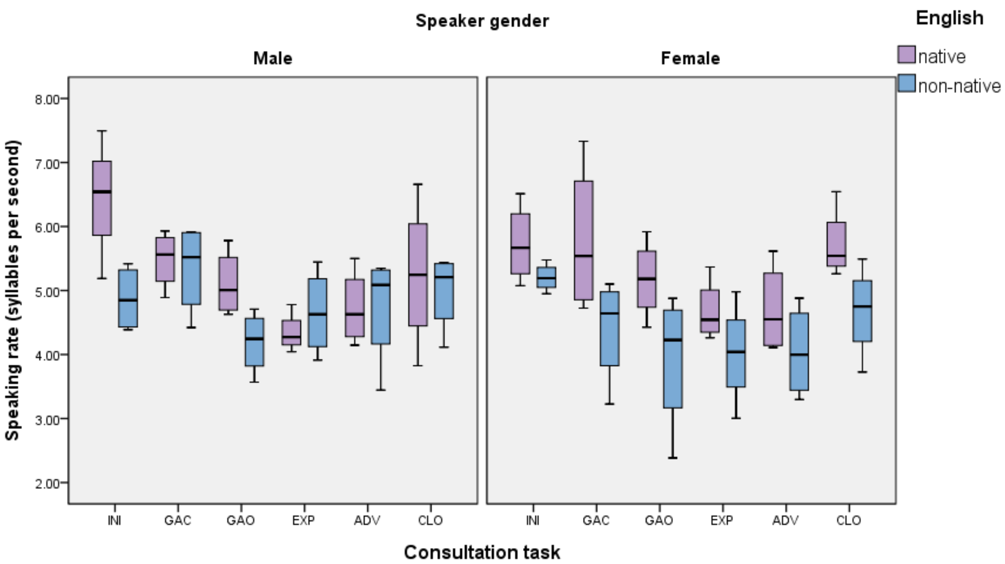


Figure 5.6 Speaking rate in the consultation stimuli according to gender and English language background by consultation task.

INI: initiating the session; GAC: gathering information with closed questions; GAO: gathering information with open-ended questions; EXP: providing explanation; ADV: providing advice; CLO: closing the session

From Figures 5.3-5.6, it seems that the some prosodic differences can be found between the native and the non-native medical students in the stimuli of some consultation tasks after accounting for the influence of gender. However, a series of Mann-Whitney *U* tests did not find any of these differences to be statistically significant (see Section 4.5.2.1for a discussion of a potential limitation of using this method for the comparisons) .Such observation agrees with the findings on the analysis on the native vs. non-native differences in prosodic features within the entire consultation role-play data set (see Section 4.5). If even small prosodic differences are found in the perceptual experiment to have an effect on perceived professionalism in medical communication, then the results will suggest that speech prosody has an important role in medical communication and needs to be addressed more specifically in communication skills training.

In summary, being subsets of the original acoustic study data, the characteristics of prosodic features in the selected stimuli for the perceptual experiment are similar to those in the corresponding sets of the original data. Results of the perceptual experiments using these stimuli can therefore be discussed with reference to the acoustic findings from the first stage of the study.

5.4.2 Instrument

In order to answer the research questions of this stage of the study, the level of professionalism perceived by the listener needs to be measured. An instrument, namely a professionalism scale, was designed for this purpose. The professionalism scale (Appendix Q) took the form of a 100-millimetre 12-item visual analogue scale (VAS) measuring the extent to which the participant perceived the particular attitudinal meanings from the voice in the stimuli. This scale resembled in form the instrument used in other similar studies (e.g. Chen et al. 2004). An example of a scale item is given in Figure 5.7 below.

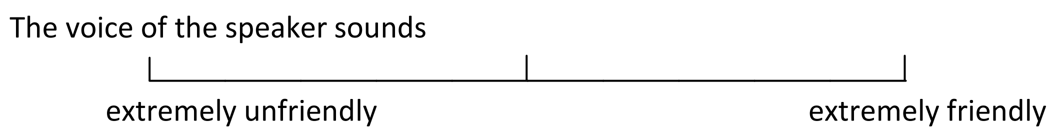


Figure 5.7 An item from the professionalism scale

The first ten of the 12 items referred to particular attitudinal meanings (see Table 5.1), while the last two evaluated the overall impression of the voice used in medical communication. Table 5.3 below is a summary of the professionalism scale items.

Table 5.3 Structure of the professionalism scale

Purpose	Item number	Anchor	
		Negative (left)	Positive (right)
Evaluating aspects of professional attitudes	1	unfriendly	friendly
	2	disrespectful	respectful
	3	uninterested	interested
	4	judgmental	non-judgmental
	5	unconfident	confident
	6	discouraging	encouraging
	7	insensitive	sensitive
	8	unfeeling	empathetic
	9	unpleasant	pleasant
	10	unhelpful	helpful
Evaluating overall satisfaction	11	inappropriate	appropriate
	12	unsatisfactory	satisfactory

While the two ends of the line of each item represent the extreme negative and positive opposing aspect of attitude, the mid-point represents an attitude-neutral message. The positive attitudes included on the scale demonstrate apparent face validity as they are mainly adapted from the Calgary-Cambridge communication process skills guide (Silverman et al. 2005) according to which medical students are trained to demonstrate rapport-building behaviour. These attitudinal messages can be effectively conveyed through speech prosody (see Sections 2.6.6.2 and 2.6.4). Furthermore, the positive attitudes on the scale also reflect service users' expectations as well as GMC (2006, 2009) requirements (see Section 2.1.2). The aspects directly representing a potential patient's satisfaction with the communication experience are evaluated with the final two items on the scale. It would be prudent to note here that, notwithstanding the face validity of the scale, this instrument was largely constructed to reflect one area of professionalism, namely patient-centredness. Of course it cannot be regarded as sufficient to measure the full spectrum of medical professionalism as defined by the GMC (see also RCP 2005 among others discussed in Section 2.1.2), but patient-centredness is the part of professionalism that a patient can easily perceive from the communication experience in the clinical encounter. It is exactly such perceived professionalism that this instrument is designed to measure.

All the 12 items are laid out in the same order, i.e. the lower end representing unprofessional attitudes (0 on the left) and the higher end representing the professional attitudes (100 on the right). This order was also commonly followed by previous attitudinal research using VAS (e.g. Peterson et al., 1999, Chen et al. 2004).

The VAS can be a useful tool to assess attributes that are otherwise difficult to measure directly, for example, pain or mood (Chen et al. 2004). Using the VAS to measure attitudinal perception in this study has several advantages:

- When measuring attitudes, a VAS avoids forcing the participants to put their measurements into categories (Lewis & Stenfert-Kroese 2010), so the nuances are better captured than with categorical scales.
- The reliability of a VAS as a measurement instrument is comparable to that of others (e.g. Likert scale) (McCormack et al.1988, van Laerhoven et al. 2004).
- The two anchor points at both ends of the scale are semantically opposite. This makes it possible to take into account the valence (also called evaluation) aspect of attitude (i.e. positive vs. negative; see Couper-Kuhlen 1986 for the tripartite classification of aspects of attitude).
- The internal consistency reliability of the scale items was tested using pilot data from the simulated patient (actor) when the stimuli were recorded. The instrument showed good consistency reliability (Cronbach's α = 0.98) according to the guidelines of Kline (1999).

The professionalism scale, one for each stimulus, was printed and bound as a workbook containing 16 scales (Appendix Q) for the participants to use during the experiment sessions.

In order to establish the test-retest reliability for the study, 22 of the 54 participants heard and rated each stimulus assigned to them a second time (in a different order) during the same sessions. It took most of these participants just under 60 minutes to complete both their experiment sessions, although a few of them spent as long as approximately 65 minutes. The result of the test-retest reliability analysis is reported in Section 5.6.

5.4.3 Procedure

The experiment stimuli were presented using sound files embedded within a PowerPoint presentation. Compared to other methods of stimuli presentation (e.g. via dedicated software package), this was thought to require minimal technical support, and be less prone to unexpected software malfunction during the experiment. Since the presentation was be slide by slide, with each slide containing a single stimulus sound file at the participant's own pace, it could help to reduce the participant's stress, which could in turn reduce potential error in response.

Sixteen PowerPoint presentations were created, one for each participant with the allocated stimuli as per the Latin square (Figure 5.1), and packaged onto CDs prior to the experiment. These presentations were identifiable with their CD numbers matched to the participation numbers. The order of stimuli presentation was the same for all participants in terms of stimuli types. Using CD media also ensured trouble-free loading of the presentations onto the AHP Communication Lab workstations without added burden on the server space, and kept the participants' workflow straightforward.

After confirmation that they had signed the consent form (Appendix O), the participants were given the rating scale workbooks (Appendix Q), their corresponding stimuli CDs and access to the AHP Communication Lab workstations (using their own UEA accounts) equipped with headphones. They were given time to read the written instructions, and to have a practice in completing the rating task. In particular, they were instructed to adjust their headphones' volume to a comfortable yet constant level, and to listen to each stimulus only once. The participants who were assigned to establish the test-retest reliability had their CDs and workbooks specially prepared, so that at the second time they heard and rated the stimuli in a different order without this being made explicit first. Most participants completed the experiment sessions within 60 minutes in total.

5.5 Data analysis

Completed workbooks were collected at the end of experiment the session. For each item on the professionalism scale, scores were measured with a ruler to the nearest integer, and entered onto a spreadsheet by the researcher for statistical analysis. The 12 individual item scores for each stimulus were then summed to create the outcome variable, i.e. the professionalism scale score.

Out of these 14,592 responses there were 14 missing. These were the result of participants having omitted to complete those individual scale items. In the case of a scale affected with item(s) containing missing value, a mean of the completed items was used to replace the missing value. Although this traditionally used imputation method can be criticised to be likely to result in a biased estimate (Donders et al. 2006), this would not be of a grave concern for the current study, as they represented less than 0.1% of the data.

The independent variables (and the units of measurement) were:

- F0 mean (Hz)
- F0 range (Hz)
- Mean dB intensity (dB)
- Speaking rate (syllables per second)
- Stimulus (1a-5a: sentences read in an unprofessional manner; 1b-5b: sentences read in a professional manner; INI; GAC; GAO; EXP; ADV; CLO)
- Speaker gender (male; female)
- Speaker English language (native; non-native)
- Speaker training level (low; high)

The research hypotheses to be tested in the analysis were:

Hypothesis 1: listeners give different professionalism scale scores to sentences read in a professional manner and those in an unprofessional manner.

Hypothesis 2: listeners give different professionalism scale sores to sentences read in a professional manner and spontaneous utterances from the simulated consultation.

Hypothesis 3: there are differences among professionalism scale scores that listeners give to the utterances from different consultation tasks.

Hypothesis 4: professionalism scale scores are different for utterances between the following groups of medical students in the same consultation task

(a) male and female
(b) native and non-native English speaking
(c) trained to a low level and to a high level

Hypothesis 5: listeners' professionalism judgement is influenced by the following prosodic features in the utterances for both native and non-native speaking medical students

(a) F0 mean
(b) F0 range
(c) mean dB intensity
(d) speaking rate

The following statistical methods were adopted and all tests were carried out using SPSS (Version 18).

- Cronbach's α for internal consistency and intraclass correlation coefficient (ICC) for test-retest reliability of the professionalism scale.
- Independent samples *t*-tests for Hypotheses 1-2, with alpha level adjusted for multiple comparisons where necessary.
- General linear models (GLMs) for Hypotheses 3-5. For Hypothesis 3, the GLM was used as a mixed analysis of variance (ANOVA) and for Hypotheses 4 and 5, the GLMs were multiple regression analyses.

5.5.1 Results

a) Internal and test-retest reliability of the professionalism scale

The internal consistency of the professional scale was assessed with all 54 participants using Cronbach's α (further details can be seen in Appendix V). The reliability was found to be good (Cronbach's α = .97) which was similar to that found in pilot data (see Section 5.4.2) and, if any one of the items was deleted, there was only a small consistent change (α ranging from .963 to .974).

Twenty-two listeners out of all 54 participants contributed to the test-retest reliability analysis by listening to and judging their allocated stimuli for a second time (presented in an order different from the original one) during their experiment sessions. The test-retest reliability of the instrument was analysed using intraclass correlation coefficient (ICC) in which some listeners reached a nearly 90% agreement between two runs. The details of the analysis are reported in Table 5.4 below.

Table 5.4 Calculating the test-retest reliability of the professionalism scale

Listener [a]	*r*	95% CI lower bound	95% CI upper bound	*F*	*df*1	*df*2	*p*
1	.80	.52	.93	9.02	15	15	< .001
2	.69	.32	.88	5.53	15	15	.001
3	.64	.22	.86	4.52	15	15	< .01
4	.72	.37	.89	6.20	15	15	0.001
5	.58	.14	.83	3.80	15	15	< .01
6	.76	.43	.91	7.17	15	15	< .001
7	.29	-.23	.67	1.80	15	15	.13
8	.79	.49	.92	8.43	15	15	< .001
9	-.02	-.50	.47	0.96	15	15	.53
10	.61	.18	.84	4.08	15	15	< .01
11	.89	.71	.96	16.68	15	15	< .001
12	.57	.12	.83	3.65	15	15	< .01
13	.55	.07	.82	3.43	14	14	< .05
14	.71	.34	.89	5.82	15	15	.001
15	.79	.51	.92	8.69	15	15	< .001
16	.81	.53	.93	9.42	15	15	< .001
17	.35	-.87	.77	1.53	15	15	.21
18	.72	.36	.89	6.01	15	15	.001
19	.71	.34	.89	5.81	15	15	.001
20	.79	.50	.92	8.62	15	15	< .001
21	.59	.15	.84	3.89	15	15	< .01
22	.87	.66	.96	14.77	14	14	< .001

a. These are not their originally allocated participant numbers on file but are given for the sake of presentation in this table.

A very low ICC can indicate a lack of agreement between the judgments under the test and the retest conditions. As can be seen in Table 5.2, this was reported for three of the 22 participants, although in those cases the correlation was found to be non-significant, suggesting that the lack of agreement might be due to chance. For most listeners who measured the test-retest reliability, the ICC ranged from moderate to high as an indication of the reliability of the instrument measuring perceived professionalism on separate occasions. The mean ICC excluding the non-significant correlations was .72, which was satisfactory. In fact even with the non-significant correlations included, the mean ICC would still be .65, suggesting a reliability well above chance. Thus when a listener is asked to rate the perceived level of professionalism in a voice of the doctor’s communication with the patient, this rating will be reliable and will agree with a later rating of the same event.

Although as shown in Table 5.4 above that perceived attitudes can be reliably measured, different individuals cannot be expected to show strong agreement when asked to make subjective judgments. Since a similar previous study (Scherer et al. 1984) did not assess the inter-rater reliability for exactly the same reason, the inter-rater reliability of the instrument in the current study was not established either.

b) Hypothesis 1

Hypothesis 1 states that "listeners give different professionalism scale scores to sentences read in a professional manner and those in an unprofessional manner". Thus it was expected that the acoustic differences between the sentences read in an unprofessional manner and those read are interpreted by listeners to reflect different levels of professionalism displayed by the speaker.

It was expected that listeners would give higher scores on the professionalism scale to the sentences read in a professional manner than those read in an unprofessional number. This was observed and reported in Table 5.5 below, which summarises the means (and standard deviations) of the professionalism scale scores for these sentence stimuli.

Table 5.5 Professionalism scale scores for sentences read in unprofessional and professional manners

Stimulus	Professionalism scale score
Unprofessional [a]	**565.94** (230.14)
Professional [b]	**720.36** (231.46)

Note. ***M*** (*SD*).

a. $n = 266$

b. $n = 270$

An independent samples *t*-test was carried out to test the differences between these means. In both samples, normal distribution was found by means of Kolmogorov-Smirnov tests (D (265) = .05, p =.20 for unprofessional manner and D (269) = .04, p = .20 for professional manner stimuli respectively) and therefore the test assumption was considered to be met.

Participants gave significantly higher professionalism scale scores to sentences read in a professional manner than to sentences read with in an unprofessional manner (t (534) = -3.181, p < .01). The mean difference was -63.423 with a 95% confidence interval ranging from -102.59 to -24.255.

It can therefore be concluded that listeners gave higher scores on the professionalism scale to the sentences read in a professional manner than to the same sentences read in an unprofessional manner by the medical students.

c) Hypothesis 2

Hypothesis 2 states that "listeners give different professionalism scale sores to sentences read in a professional manner and spontaneous utterances from the simulated consultation". It explores the difference in perceived level of professionalism from more consciously produced speech (read sentences) and more spontaneously produced speech (simulated consultation) in the clinical context where professional attitudes are expected. The hypothesis aims to ascertain if the acoustic properties in these two types of speech stimuli (performing the same communication goal) are perceived by listeners to convey the same attitudes. Since medical students are trained to appropriately display professionalism in the consultation with patients (whether real or simulated), one would expect to find similar levels of perceived professionalism between the more spontaneous utterances in consultation and the more consciously read sentences expressing professional attitudes.

Comparisons were made between sentence 2b (read in professional manner) and gathering information with closed-ended questions, between sentence 4b (read in professional manner) and gathering information with open-ended questions, and between sentence 5b (read in professional manner) and gathering information with open-ended question on the professionalism scale (see Section 5.4.1 for full sentences). The means (and standard deviations) of these stimuli are summarised in Table 5.6 below.

Table 5.6 Professionalism scale score for sentence reading in a professional manner and the simulated consultation

Stimulus	Professionalism scale score
Sentence 2b	**617.67** (208.84)
Gathering information with closed-ended question	**674.28** (195.68)
Sentence 4b	**722.52** (220.18)
Sentence 5b	**766.11** (214.25)
Gathering information with open-ended question	**696.41** (228.30)

Note. ***M*** (*SD*); *n* = 54.

These comparisons were subsequently subjected to a series of three *t*-tests, with an adjusted $\alpha = .017$ by means of Bonferroni correction for multiple comparisons. The data were found to be normally distributed in all samples and therefore to have met the assumption of the tests. No significant differences were discovered in these tests, as shown in Table 5.7 below.

Table 5.7 Results of comparing corresponding stimuli on the professionalism scale

Comparison	*t*	*df*	*p*	Mean difference	95% CI	
2b vs. GAC	-1.454	106	.15	-56.611	-133.824	20.602
4b vs. GAO	0.605	106	.55	26.111	-59.460	111.683
5b vs. GAO	1.636	106	.11	69.704	-14.766	154.173

Note.
INI: initiating the session; GAC: gathering information with closed questions; GAO: gathering information with open-ended questions; EXP: providing explanation; ADV: providing advice; CLO: closing the session

These test results indicate that listeners gave similar professionalism scale scores to sentences read out in a professional manner and corresponding spontaneous utterances in the simulated consultation. This conclusion is also in the same direction as normal expectations about student doctors' professional behaviour.

d) Hypothesis 3

Hypothesis 3 states that "there are differences among professionalism scale scores that listeners give to the utterances from different consultation tasks." It compares within-speaker differences using a repeated measures design. The means and standard deviations of the professionalism scale scores for the different consultation tasks are summarised in Table 5.8 below.

Table 5.8 Professionalism scale score by consultation task

Consultation task	**Professionalism score**
Initiating the session	**811.69** (147.40)
Gathering information with closed-ended questions	**674.28** (195.68)
Gathering information with open-ended questions	**696.41** (228.30)
Providing explanation	**727.30** (190.01)
Providing advice	**705.24** (200.78)
Closing the session	**734.11** (248.34)

Note. ***M*** (*SD*).

In order to test the differences in professionalism scale scores between the consultation tasks, a general linear mixed model was used with the different consultation tasks (i.e. initiation, etc.) as the fixed factor and speaker ID (to distinguish one medical student from another) as a random factor into the model. Both factors were found to have significant main effects: listeners gave different professionalism scale scores to different consultation tasks (F (5,77.32) = 3.91, p = .003), and also to different medical students (F (15,76.11) = 5.41, p < .001). There was no interaction between consultation tasks and speaker ID (F (75,228) = 1.05, p = 0.38). *Post hoc* analysis using the Tukey HSD test only found significant differences between initiation and closed-ended question (p < .01), between initiation and open-ended question (p < .05), between initiation and advice (p < .05), but not between any of the others.

The analysis supported the research hypothesis and showed that listeners reported significantly different levels of professionalism when the medical students carried out different tasks in the simulated consultation. More specifically the professionalism scale scores for initiating the session were significantly higher than those for gathering information with closed-ended questions, for gathering information with open-questions and for providing advice. Significantly different levels of professionalism for different individual medical students were also reported by listeners.

In an attempt to aid the interpretation of the different professionalism scale scores for stimuli representing different consultation tasks (INI, GAC, GAO, EXP, ADV and CLO), a count of empathetic utterances in the simulated consultations was additionally carried out. It was revealed that just 30% of all utterances of the medical students could be regarded empathetic (e.g. attentive listening, clarifying patient's agenda, or encouraging the expression of thoughts and feelings; see Silverman et al. 2005).

When the numbers of empathetic utterances in the simulated consultations were compared across the different consultation tasks, these were found to be significantly related to the task that such utterances carry out in the consultation ($\chi^2(3) = 69.38$, $p < .001$). In particular, more empathetic utterances occurred in gathering information than in imitating session, in explanation and planning, and in closing session (GAC and GAO both fell into "gathering information", and EXP and ADV both fell into "explanation and planning" according to the Calgary-Cambridge model). However, it should be pointed out that there were much more utterances for gathering information and explaining/planning than for initiating and closing session. The percentages of empathetic utterances within the consultation tasks were more revealing. For example, approximately 51% of the initiating session utterances were empathetic, but only approximately 33% of the gathering information utterances were empathetic. Full details of the distribution are visually presented in the bar chart below.

It can be seen in Figure 5.8 below that it would be more likely for empathetic utterances to occur in the initiation stage than, for example, in gathering information. The sampling of the GAC, GAO, EXP and ADV utterances was were constrained by the content of information exchange characterising these consultation tasks and, in effect, was excluding the empathetic utterances to be selected as the GAC, GAO, EXP and ADV stimuli. Although empathy was only one item of the professionalism scale, the count of empathetic utterances in the simulated consultations shed some light on why listeners perceived higher level of professionalism (patient-centredness in particular) when hearing the INI stimuli than the GAC, GAO and ADV stimuli.

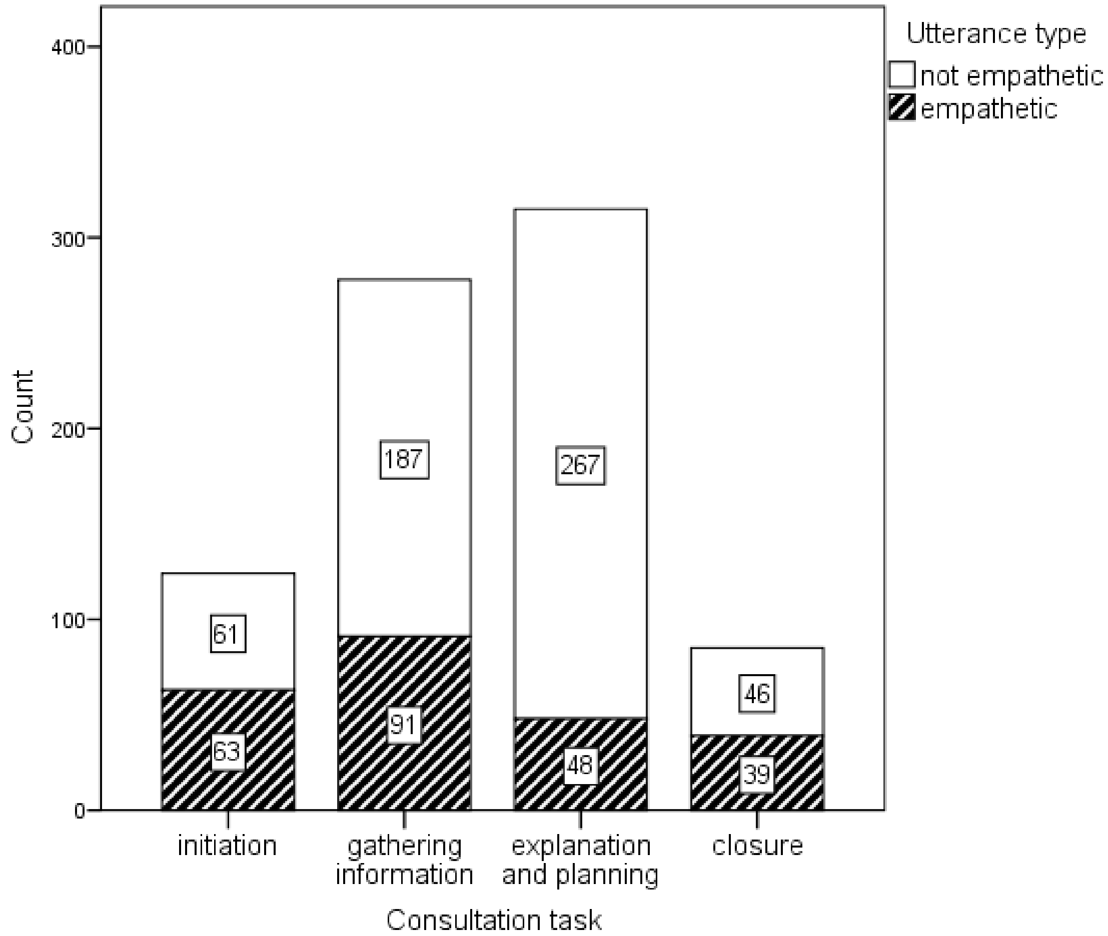


Figure 5.8 Number of bridging empathetic utterances by consultation task

e) Hypothesis 4

Hypothesis 4 states that "professionalism scale scores are different for utterances between the following groups of medical students in the same consultation task

(a) male and female
(b) native and non-native English speaking
(c) trained to a low level and to a high level".

The hypothesis involves examining effects of both the speaker attributes, i.e. gender, English language, and training level, as well as consultation tasks. Based on the results of testing Hypothesis 3, a significant effect of consultation task on perceived level of professionalism can be expected, so this was entered into the model so that it would be adjusted for. Furthermore Hypothesis 4 aims to reveal any effects of the speaker attributes regardless of the consultation task. Similar to testing Hypothesis 3, the investigation to test Hypothesis 4 also utilised a general linear model.

Using the method of Type III sums of squares, the model examined the main effects of the factors speaker gender (Speaker_gender), speaker English background (Speaker_English), speaker training level (Speaker_training) and consultation task (Stimulus_type), as well as the two way interactions between them. The model can therefore be formulated as:

Intercept + Stimulus_type + Speaker_gender + Speaker_English + Speaker_training + (Speaker_gender * Speaker_English) + (Speaker_English * Speaker_training) + (Stimulus_type * Speaker_English) + (Speaker_gender * Speaker_training) + (Stimulus_type * Speaker_gender) + (Stimulus_type * Speaker_training)

Significant main effects were found in the consultation task (F (5,297) = 3.053, p = .01), speaker English background (F (1,297) = 6.788, p = .01), and speaker training level (F (1,297) = 12.265, p = .001), with native English speaking and medical students trained to a high level scoring higher on the professionalism scale. Effect of speaker gender was not significant (F (1,297) = 0.87, p = .35). These differences are summarised in Table 5.9 below.

Table 5.9 Professionalism scale score by speaker gender, English language and training level.

Factor	**Professionalism scale score**	
Gender	Male **735.55** (229.65)	Female **713.01** (179.36)
English language background**	Native **753.35** (217.28)	Non-native **695.61** (192.80)
Training level***	Low **685.48** (207.98)	High **763.24** (199.83)

Note. ***M*** (*SD*)**.**

** p < .01; *** p < .001.

None of the interactions between the investigated factors was found to be significant. The model itself however, was not significant (F (26,297) = 1.47, p = .07), nor did it show a good fit (R^2= .11, *adjusted* R^2= .04), despite that the standardised residuals were normally distributed. This is perhaps not unreasonable given the poor contribution of the non-significant gender effect and the non-significant interactions to the model.

The above results (see Appendix W for further details) indicated that listeners perceived higher level of professionalism from native speaking medical students than from their non-native speaking peers, and that the increased level of communication skills training for the medical students had a positive effect on their professionalism scale scores. The gender of the medical student had no significant effect on the level of professionalism perceived by listeners.

f) Hypothesis 5

Hypothesis 5 states that "listeners' professionalism judgement is influenced by the following prosodic features in the utterances for both native and non-native speaking medical students

(a) F0 mean
(b) F0 range
(c) mean dB intensity
(d) speaking rate".

This is the step to find an answer to the overarching research question whether speech prosody matters in health communication. Since the previous model to test Hypothesis 4 already showed that native and non-native speaking medical students scored differently on the professionalism scale, the next step would be to ascertain if any prosodic feature investigated in Stage 1 (F0 mean, F0 range, mean dB intensity and speaking rate) could have an effect on the professionalism scale scores for both groups. In order to do so, the general linear model with the factors F0 mean (F0meanHz), F0 range (F0rangeHz), mean dB intensity (dB) and speaking rate (speaking_rate) entered as covariates was tested with both the native group and the non-native group separately. For both groups, the method Type III sums of squares was used in the model, which can be formulated as:

Intercept + F0meanHz + F0rangeHz + dB + speaking_rate

A summary of the variables is provided in Table 5.10 below.

Table 5.10 Professionalism scale scores with prosodic features in native and non-native English speaking medical students

English language background	Native	Non-native
F0 mean (Hz)	**157.09** (44.55)	**161.12** (54.95)
F0 range (Hz)	**161.76** (70.28)**	**158.72** (83.23)
Mean intensity (dB)	**59.12** (3.90)***	**58.47** (2.98)
Speaking rate (syllables/second)	**5.25** (0.87)	**4.61** (0.77)**
Professional scale score	**753.35** (217.29)	**695.61** (192.80)

Note. ***M*** (*SD*)**.**
** $p < .01$; *** $p < .001$.

Not all the prosodic features, as can be seen in Table 5.10, were found to have significant effects, and the model revealed different patterns between stimuli from the native group and the non-native group.

For native speaking medical students, the prosodic features F0 range and mean dB intensity were found to have significant effects on the professionalism scale scores listeners gave, with utterances with higher F0 range (F (1,159) = 7.158, $p < .01$, *partial* η^2 = 0.04) and lower intensity (F (1,159) = 23.398, $p < .001$, *partial* η^2= 0.13) scoring better on the scale. Effects of the other prosodic features were not significant.

For non-native speaking medical students, speaking rate was the only prosodic feature found to have a significant effect on the professionalism scale scores listeners gave to non-native speaking medical students, with faster speaking rate in the utterance (F (1,155) = 8.274, $p < .01$, *partial* η^2= .05) contributing to higher level of perceived professionalism. F0 mean, F0 range and intensity had no significant effect on professionalism scale scores for the non-native group.

Regression diagnostics were also examined. The distribution of the standardised residuals in both groups was normal, suggesting that the assumptions of the model were met. The model fit better in the native group (R^2= .15, *adjusted* R^2= .13, F (4,159) = 6.95, $p < .001$) than in the non-native group (R^2= .06, *adjusted* R^2= .04, F (4,155) = 2.57, $p < .05$). Further details of the analysis can be found in Appendix X.

The factor "training level" was previously found to have significant positive correlation with perceived professionalism when testing Hypothesis 4, therefore could be an important one to consider in the model. The alternative model that included this factor was tested, which yielded the same results; that is, the significant prosodic features for the natives were still greater F0 range and lower intensity, and the significant prosodic feature for the non-natives was still faster speaking rate. However, including training level could improve the model fit a little, for both the natives ($R^2 = .17$, adjusted $R^2 = .14$, $F(5,158) = 6.35$, $p < .001$) and the non-natives ($R^2 = .09$, adjusted $R^2 = .06$, $F(5,154) = 6.12$, $p = .01$). This suggested that prosodic features could themselves play a part in affecting perceived professionalism, and such effect could be considered to be in addition to the effect of training level.

In conclusion, it can be said that for both native and non-native speaking medical students, prosodic features in their simulated consultation had significant effects on the level of professionalism that listeners perceived. However these effects were stronger in the native speaking group.

5.6 Summary

The perceptual study addressed a series of research questions and carried out statistical analyses on the collected data attempting to answer these questions. The questions aimed to explore the link between the prosodic features used and the perceived professionalism in medical communication. More specifically the effects of several factors related to the attitudinal intentions, speaker attributes and acoustic properties of the voice stimuli on the level of perceived professionalism were investigated. The investigation was implemented by means of examining the statistical associations between the outcome measure on the study instrument, a specially designed professionalism scale and the factors under investigation.

Most of the results from the statistical analysis were expected findings. The following findings emerged from the analysis as expected.

- Attitudinal intentions as encoded in the voice stimuli (in the form of two manners of reading clinically relevant sentences) had an effect on the perception of listeners who were ask to rate the level of professionalism from the voice in the stimuli.

- There was no significant difference in the level of perceived professionalism between clinically relevant sentences read in a professional manner and utterances representing similar tasks in the simulated consultation conducted by the medical students.
- The levels of professionalism perceived from stimuli representing the various tasks in the simulation consultation were different. Listeners reported a higher level of professionalism from the stimuli representing the initiation stage of the simulated consultation compared with those representing gathering information with closed or open-ended questions or providing advice.
- After the effects from the consultation task and idiosyncrasy between each individual medical student being accounted for, the English language background and the training level of them still influenced the level of professionalism perceived from the voice stimuli contributed by these medical students. Being a native speaker of English and being trained to a high level both had positive influence on the level of professionalism in the voice of the medical student perceived by the listeners.
- Prosodic features in brief voice stimuli had a significant effect on the level of professionalism that the listeners perceived from hearing them. However, the exact pattern of the effect was dependent on whether or not a stimulus was produced by a native English speaking medical student. For stimuli from native speaking medical students, a wider F0 range (perceived as wider ranging pitch) or a lower mean dB intensity (perceived as softer voice) were associated with higher level of perceived professionalism. For stimuli from non-native medical students, only a higher speaking rate (perceived as faster speech delivery) contributed to a higher level of perceived professionalism.

These findings suggest that in medical communication, prosodic contribute to the listeners' perception of the speakers' professional attitudes and they support the findings in the acoustic study that the native and the non-native English speaking medical students differ in the use of prosodic features in the clinical setting. In fact, some of the perceptual findings have been predicted by the acoustic findings. The perceptual study showed that different professionalism scale scores could result from different attitudinal intentions in the utterances which also exhibited different acoustic properties of speech prosody. Furthermore consultation task in the stimuli had an effect on the level of perceived professionalism as well as on the prosodic features. Finally, different professionalism scale scores were given to simulated consultation stimuli from the native and the non-native English speaking medical students who were found in the acoustic study to differ in prosodic features as well. To investigate the effect of speech prosody on perceived professionalism alone, an alternative research design which can utilise stimuli varying in prosodic features from the same speaker. Such design would be a strong one for that purpose. However, the current study had a rather different purpose. It aimed to link the perceptual evidence to the prosodic differences between the native and non-native speaking medical students observed in the acoustic study. Consequently, the design adopted in the current study was able to contribute evidence needed to answer the overarching question.

In summary, prosodic features in medical communication can influence the perception of professional attitudes which have been shown to influence patient experience. So far all three Research Questions in the study have been supported by empirical evidence. The findings of the first two stages of the study jointly contribute to the overarching question whether speech prosody matters in health communication and hence calls for dedicated training. The answer to that question can be more convincing if the joint findings are discussed in the context of the emerging themes in relevant literature. Such general discussion which is presented in Chapter 6 serves as the final step in answering that overarching question.

Chapter 6 Discussion of findings

In order to answer the overarching question as to whether or not speech prosody matters in health communication, two stages of investigation were carried out using evidence from native and non-native English speaking medical students. Three Research Questions were addressed and the acoustic properties of prosodic features used in a simulated clinical interaction, the level of professionalism perceived from listening to voice stimuli recorded during the acoustic investigation and the association between prosodic features and perceived professionalism were examined. Findings in answer to these Research Questions are discussed in the current chapter with reference to the identified themes in the body of literature reviewed in Chapter 2. The significance of the findings in this study must be clarified, and the acoustic findings will be considered alongside the perceptual findings so that the answer to the overarching question can be found and interpreted in context. Thus this Chapter begins with a summary of the findings of each stage respectively. The overarching question is then answered and discussed in the context of previous research in the areas of speech prosody and health communication. The methodology adopted in this study is also critiqued in the current chapter.

6.1 The research findings

6.1.1 The acoustic findings

Firstly, the medical students in this study were found to use prosodic features when conveying professional attitudes that were appropriate to the clinical setting. The attitudinal function of prosodic features in health communication was confirmed by the differences in prosodic features used in read sentences read first in an unprofessional manner and then in a professional manner. It was revealed that the medical students changed their prosodic features to communicate vocally the intended professional attitudes, and the change was found in both genders. It should also be pointed out that the observed prosodic differences in the read sentences could only partly be accounted for by the attitudinal instructions – the function of interest to this study. It could even be said that the "unprofessional" instructions were highly hypothetical, as the unprofessional attitudes complained about by patients could often differ (see Hunt & Glucksman 1991, Wofford et al. 2004). Therefore the attitudinal instructions were not meant to represent the full range of attitudes that are appropriate or inappropriate to medical communication, and were only examples of what can be regarded as professional and unprofessional attitudes. Nonetheless, the acoustic evidence still suggested the attitudinal role of prosodic features in the clinical context.

More specifically, when instructed to convey aspects of professionalism, it was found that both male and female medical students used lower values of the acoustic parameters under investigation, that is, mean and range of F0, mean dB intensity and speaking rate. Interestingly, not all these parameter adjustments reached the corrected significance level in the statistical analysis. This was not entirely unexpected, because these parameters were not considered to have equally strong association with a wide range of speaker affects in previous literature (Scherer 2003, Scherer et al. 2003). Despite the general role of speech prosody in conveying attitudes and emotions, some particular emotional messages such as interest/disinterest and cold anger did not have strong or significant correlation with any of the F0, intensity or temporal parameters studied by Banse & Scherer (1996). In the current study, intensity emerged as the only significant acoustic parameter that the medical students manipulated while reading the clinically relevant sentences after correction for multiple comparisons in hypothesis testing: lowered intensity was significantly associated with the expression of professional attitudes. Even with the conservative Bonferroni method, the eveidence still suggested that one prosodic feature was likely to be adjusted to convey the attitudes as instructed. Medical students might be aware that certain prosodic configuration (e.g. speaking too loudly) should not be used in the clinical communication context. Claiming, based on this evidence alone, that prosodic features are used to convey professional attitudes in health communication would be overly bold at this stage, although it should certainly encourage further research to extend such finding of the attitudinal functions of prosodic features in this very context.

Secondly, the medical students, both male and female, were found to vary their prosodic features when conducting the simulated consultation. In other words the values of the prosodic features differed when they performed different tasks in the simulated consultation. Like in the read sentences, not all acoustic parameters were responsible for the prosodic differences in simulated consultations. Among the males, their F0 mean and intensity differed significantly for different consultation tasks, but among the females, only speaking rate differed significantly. This was further supported by *post hoc* analysis which showed, for example, that the acoustic parameters (F0 mean for the males and speaking rate for the females) were significantly lowered in explaining/planning compared to those in gathering information.

This does not mean that the males and females were completely different in prosodic adjustments for the simulated consultation, as the males and the females showed similar patterns of changing prosodic features according to the consultation tasks. For both genders F0 mean remained at relatively stable between initiation and gathering information but dropped during giving explanation before rising again at closure, their speech intensity showed a continual drop throughout the consultation tasks, and their speaking rate adjustment resembled a V-shape with the dip at giving explanation. This last pattern of speaking rate was particularly interesting, which suggested that the medical students in the study, male and female, tended to speak faster at the initiation and the closure of the session but more slowly when giving explanation to the simulated patient. One possible reason for this is that different tasks in the consultation may have varying affective load according to the medical communication training model used (Silverman et al. 2005). For example, in explaining to patients there may be fewer opportunities to convey affective information than in the initial phase of the consultation where it is natural to display friendliness and willingness to help. The varying affective load can have an effect of the prosodic features used by medical students (see Banse & Scherer 1996). Alternatively, it is also possible that the adjustment of speaking rate during the simulated consultation might be influenced by the content load of the different tasks. Given that there was a heavier load of factual information when giving explanation on the part of the medical students in the simulated consultation, it is would not be unreasonable for them to lower their rate of speech (including using more pauses) in order to be understood better by the simulated patient. However, these possibilities are not to be viewed as conflicting rather than complementary. As discussed in

Chapter 2, the attitudinal and the information functions (and indeed other functions such as linguistic and discourse) of prosodic features are not mutually exclusive but rather act jointly (with necessary interactions) in speech communication. Therefore prosodic adjustment according to the affective (and/or content) load in different tasks may genuinely be functioning in the simulated consultation.

Thirdly, the English language background of the medical students had a significant effect on their prosodic features used in the clinical context. Like the analyses of other factors, this was found when the males and the females were examined separately. In the clinically relevant sentences read in a professional manner, speech intensity and speaking rate were significantly lower in the non-native females than in the native females, although none of the prosodic differences between the native and the non-native males reached the statistical significance level required. However, as before, this could reflect the method of correction for multiple comparisons that resulted in a conservative significance level.

Fourthly, the statistical analysis did not reveal any significant prosodic difference between the native and the non-native medical students in any particular consultation task, despite that visually plotting the differences suggested otherwise (see Figures 4.1-4.4). A tentative interpretation is that compared to the consultation task, the English language background of a medical student may not be as predominant a factor in influencing prosodic features in the simulated consultation. This finding also agreed with the observation that in the clinically relevant sentences read in a professional manner, the native and the non-native speakers did not have very different choice of intonation patterns (See Table 4.4).

Finally when the two acoustic data sets were compared, a significant within-subject effect was found among the female medical students in the study. This finding showed that they had lower F0 mean and speech intensity values in simulated consultation than for the sentences read in a professional manner in which the values were in turn lower than those for the sentences read in an unprofessional manner. Though not possible to be confirmed by this study, the female medical students could be very aware of the attitudinal continuum from the unprofessionally read sentences to the professionally read sentences to the simulated consultation, which might have been reflected in their prosodic features.

In summary, prosodic features were found to be important in the simulated clinical interaction in this study, while a range of within- and between-subjects factors were influencing their acoustic values. The importance of prosodic features lies in particular in their attitudinal function in health communication medical students seemed to be aware of. The perceptual experiment carried out in the study examined the effects of prosodic features together with several other attributes of medical students on the level of professionalism communicated in the simulated clinical interaction.

6.1.2 The perceptual findings

Five research questions were raised in the perceptual study to understand the effects of the intended attitudes, the types of stimuli, the consultation tasks, the attributes of the speakers, and the prosodic features on the level of professionalism perceived by untrained listeners. Examining the effects was a first step to explore the influence of these factors on potential patient experience. The effects of these factors answered the research questions and are summarised and interpreted as follows.

Firstly, the clinically relevant sentences read out by the medical students in a professional manner were rated by the listeners as to reflect a higher level of professionalism than the same sentences read out in an unprofessional manner. Therefore after controlling for the speech content, the attitudes encoded in the voice stimuli were still perceived by the listeners to reflect different levels of professionalism. This clearly showed that the attitudinal message can affect the perception of professionalism in medical communication regardless of the verbal content.

Secondly, the listeners rated the sentences read out in a professional manner and the corresponding spontaneous utterances in the simulated consultation as reflecting similar levels of professionalism. This important finding is encouraging because it shows that the medical students can maintain a comparable level of professionalism in the simulated clinical interaction even without explicit attitudinal instructions. That is to say, the medical students can project aspects of professional attitudes spontaneously when interviewing the simulated patient, and they can do this at a similar level when asked to do so when reading out clinically relevant sentences with professional attitudes. If they were interviewing a real patient this way, the patient experience with the clinical communication would be likely to be positive due to the level of professionalism conveyed. Viewed together, the above findings suggest the importance of displaying professional attitudes in the communication behaviour during clinical encounters. Other things being equal, the professional attitudes present in medical communication contribute to the communication being perceived as to reflect professionalism.

Thirdly, listeners reported different levels of perceived professionalism for different tasks of the simulated consultation. In particular, initiating the session was perceived as to reflect a higher level of professionalism than asking closed-ended questions, asking open-ended questions and giving advice. The fact that such important tasks were rated lower than the initiation, although not entirely unexpected, can suggest room for improvement. It showed that the students sounded professional at the beginning of the consultation, but when the tasks of taking history and advising patients were performed, they did not sound as professional. The vocal display of professionalism should not be limited to the “ice-breaking” stage, but should be applied to every consultation task so that the therapeutic professional relationship can be maintained. That said, it is still natural for some consultation tasks (e.g. initiating the session) to have more affective colouring than others: it is natural for a patient to feel the warmth when greeted by name rather than when being suggested the treatment plan for his or her condition. Thus the different professionalism scale scores may reflect the nature of each consultation task rather than whether professional attitudes were displayed. However, given that attitudes can be cognitive as well as affective, there are still opportunities for displaying sensitiveness and empathy vocally when gathering information with open and closed-ended questions. Therefore medical professionalism based on a range of attitudes (and reflected in behaviour) can still be consistently perceived in the consultation.

Fourthly, some speaker attributes were associated with higher level of professionalism perceived by the listeners. These were native (compared to non-native) speaking medical students and medical students with a high level of communication skills training i.e. those in the final year of medical school (compared to Year 2 and Year 3 medical students).

Gender of the medical students did not have a significant effect on the level of professionalism the listeners perceived from the voice. However, as hypothesised, the significant positive effects of native English language background and high level of training on the perceived level of professionalism were both found. The finding on English language background was backed up by the evidence from the research literature on the prosodic differences between native and non-native speakers of English (Binghadeer 2008, Wennerstrom 1994). This finding also agreed with previous studies on the attitudinal perception from prosodic features (Scherer et al. 1984, Banse & Scherer 1996, Gobl & N í Chasaide 2003, Scherer et al. 2003, Spackman et al. 2009). The finding of this study extends the understanding of non-native English prosody and the vocal expression of attitudes to a communication context previously not examined in the corresponding disciplines.

The effect of training level calls for more careful consideration in the interpretation. Due to the pattern of the training programme for the medical student samples, it was not possible to entirely separate the effect of communication skills training from the effect of other trainings happening in the medical school. However, the design of the perceptual experiment attempted to minimise the confounding effect of increasing medical knowledge of the students. Neither the simulated consolation scenario nor format was unfamiliar for any of the medical students, though the final year students would understandably tend to be more confident than the Year 2 and Year 3 students. While this study cannot assert that the effect of training level was free from aspects of medical education which may have influenced their attitudinal development, it can nevertheless be assumed that the different level of communication skills training influenced the level of perceived professionalism in their communication.

In terms of communication skills training, the finding of this study highlights the potential additional needs of non-native speaking medical students as well as the importance of the teaching and learning of medical communication skills. The highlight on the effect of English language background on the perceived level of professionalism also suggests that in answering the question on the effect of prosodic features, a model needs to be built while tailoring to the two different speaker groups.

Fifthly, prosodic features were found to be significant predictors of the perceived level of professionalism after accounting for the English language background. However, the effect sizes were not big, and the effect was more obvious in native speaking students than in non-native speaking students. Recalling the acoustic study findings, the two stages of research can paint a fuller image of prosodic features used in medical communication, that is, the English language background of medical students has a significant effect on the level of professionalism perceived from their clinically relevant utterances in addition to their prosodic features.

The overarching aim of the study was to explore and ascertain the role of speech prosody in health communication. The final analysis of the perceptual data was an attempt to serve that aim. Prosodic features as a whole do have a significant role in terms of conveying professionalism in clinical encounters. The evidence for this can be found in both the native and the non-native groups. The significance of this finding lies in that regardless of the English language background of a medical student, and perhaps of a doctor as well, the speech prosody in his or her consultation conducted in English with a patient is important for the perceived professionalism. In other words, everything else being equal, adjusting the prosodic features in a medical student's or a doctor's speech can improve the level of professionalism perceived by a potential patient/service user during the consultation.

More importantly, the effect of prosodic features on perceived professionalism in clinical consultation is different between the native group and the non-native group. Such difference is two-fold. On the one hand, the effect of prosodic features on perceived professionalism appeared larger for the native group than for the non-native group. Approximately 15% of the variation in the perceived professionalism in the native group data could be explained by prosodic features alone, while in the non-native group data, only 6% of the variation could be explained by prosodic features alone. On the other hand, the effect of prosodic features on perceived professionalism showed different patterns from the native group and the non-native group. For the native group increasing F0 range and decreasing intensity in particular, perceived as wider pitch excursions and lower voices in their speech, lead to a higher level of perceived professionalism. For the non-native group, however, only increasing speaking rate, resulted in a perceived faster speech will in turn be associated with a higher level of perceived professionalism. It might be the case that when native listeners judge the professionalism conveyed through the speech prosody of the medical students, they can have different expectations for the native speakers and for the non-native speakers. Furthermore, in this kind of judgment, the native listeners might (unconsciously) use other characteristics in the speech voice of non-native medical students even if they are told not to. As a consequence, the unconsciously used factors for judgment might have prevailed against prosodic features in the judgment. For example, although none of the non-native speaking medical students in this study showed strong foreign accents (represented by, particularly, non-standard vowels and consonants, see Section 4.3), it was not able to rule out the effect of even mild foreign accents on perceived professionalism judgment. Since slower speaking rate has generally been accepted as a feature of non-native English speech (Yuan et al. 2006), it could not be certain, at this stage, whether native listeners were biased against this feature and perhaps consequently biased in their judgment of professionalism perceived from non-native utterances.

Nevertheless, the effect on perceived professionalism was unlikely to be confounded by verbal content of the speech in this study. This was because that when selecting the perceptual experiment stimuli from the recorded simulated consultations, care was taken so that the verbal content of the selected stimuli could be comparable across the groups (see transcripts in Appendix S). Although the consultation task did not serve as a strict control over the wording (otherwise the resulted speech would be less spontaneous), the verbal content in the stimuli representing the same consultation task from different medical students was nonetheless very similar. In giving explanation, for example, all medical students explained the same clinical condition (hypertension) or treatment, and any variation in the actual words used was not deemed to lead to very different attitudinal perceptions. Therefore there can be a fair degree of confidence that the effects on perceived professionalism could be attributed to prosodic features in the stimuli.

It is also necessary to point out that the effect of prosodic features on professionalism in clinical consultation, although significant, was not a large one. This is not at all unexpected. The reason can potentially be that perceived professionalism, although measurable in the form of its attitudinal aspects (communication behaviour in this current study, see discussion in Chapter 2) using the purposely developed instrument, can be a highly subjective matter. It is only natural for some observers to be more sensitive and more able to pick up even minute prosodic cues in speech communication. Such observers are likely to report much higher or much lower than average values compared to other observers. The instrument, however, showed good internal consistency reliability as well as good test-retest reliability. The format of the instrument, i.e. the VAS has been accepted as a sensitive tool for measuring properties that are otherwise difficult to quantify (van Laerhoven et al. 2004, Yiu & Ng 004). Therefore the measure of perceived professionalism in the current study was considered reliable and valid.

6.1.3 Answering the overarching question

The overarching question in this study is whether speech prosody matters in health communication. To this end the acoustic investigation and the perceptual investigation in this study were carried out in order to explore the effects on prosodic features used by medical students and the effects on perceived professionalism respectively. Factors affecting prosodic features included the attitudinal intentions, and the English language backgrounds of the medical students. Factors affecting perceived professionalism included the attitudinal intentions, the English language backgrounds and the training levels of the medical students and the prosodic features that they used in medical communication. Interestingly not all prosodic features measured in the study were significantly affected by the attitudinal and speaker attributional factors, as revealed in the acoustic investigation. However, in the perceptual experiment it was found that prosodic features affect perceived professionalism significantly, and this effect was significant regardless of the attributes of the medical students. These findings put together showed that prosodic features play an important role in health communication, since even small differences in prosodic features can have a significant effect on the professionalism in communication. Even more importantly, this effect applies to both native and non-native English speaking medical students. Therefore the overarching question raised in the current study can be answered with confidence that speech prosody matters in health communication. Limited by samples used, these findings cannot be directly generalisable the communication between qualified doctors and real patients, though the current study still intends to deliver a message for healthcare professionals as well as students. That is, although the content of the exchange during clinical encounters and the amount as well as structure of the conveyed information are important, the way healthcare professionals speak to patients during the encounters are equally important. Previously healthcare professionals and students have been told that the non-verbal channels of communication are particularly responsible for emotions and attitudes expressed. There has also been suggestion to them in the communication skills textbooks that the prosodic features of speech (also called vocal cues by some authors) can play a role. Until the current study, the role of prosodic features to display professionalism in health communication has not been clarified. Since this study has demonstrated that the role of prosodic features is associated with perceived professionalism in communication with patients, healthcare students and professionals alike should be encouraged to actively explore the prosody as well as the content of their speech. An active, appropriate use of prosodic features to enhance health communication will not only make healthcare professionals and students better communicators, but will ultimately improve patient experience with the clinical encounter.

Even so, as a note of caution, one needs to accept that speech prosody is by no means the only communication channel in which patients perceive professionalism in clinical encounters. Given the complex construct of professionalism, a wide range of communication channels including the verbal content, the information and discourse structure, the non-vocal, non-verbal features, and even environmental factors can be responsible for the conveyance and perception of medical professionalism. Furthermore, it is very likely that many of these channels function collaboratively in good health communication that is characterised by a high level of perceived professionalism. This concept is visualised in Figure 6.1 (Section 6.2.1).

6.2 Interpreting the findings in context

Most of the findings, both acoustic and perceptual of this study summarised above supported the experimental hypotheses and had their foundation in the existing literature reviewed in Chapter 2. This section compares the findings of this study with those reported in previous research.

6.2.1 Findings supported by existing literature

The acoustic study findings confirmed that prosodic adjustments were exploited by the medical students as a means to express professional attitudes as well as to structure consultations. Furthermore, prosodic differences between the native and the non-native speaking medical students were found in both reading clinically relevant sentences and conducting simulated consultations in which professional attitudes were required.

The attitudinal function of prosodic features is well documented in linguistic and psychological research (see e.g. Nolan 2006, Scherer 2003, Scherer et al. 2003). Such function of prosodic features in conveying attitudinal information in speech and modifying the overtone of the verbal content is supported by this study. Through the comparison in the acoustic properties of prosodic features between content-controlled utterances from the simulated clinical interaction, findings in the study showed that prosodic features serve as important communicative resources in clinical interactions to display professional attitudes. This finding applied to both male and female medical students regardless of their English language background. It is true that the information conveyed in the verbal content of the consultation is important and so is the structure of the content (Silverman et al. 2005). However it is also important that prosodic features are used appropriately in the consultation, otherwise the intended attitudinal message might not be conveyed and even the content of the conversation might be compromised, leading to potential confusion if not misunderstanding. This was found in the sentence reading data in which the medical students used decreased intensity in speech, resulting in the perception of a softer voice when trying to convey aspects of professional attitudes. Conversely the same content spoken with increased intensity in speech can be a means of displaying attitudes inappropriate to the clinical communication context and should therefore be avoided.

The acoustic parameters of prosodic features can also be changed according to the tasks of the simulated consultation, which was observed among the medical students in this study. The prosodic differences between some consultation tasks were particular observable, for example between initiating the session and explaining/planning, and between gathering information and explaining/planning. This manipulation of prosodic features can be significant in practice: it might be intended as a gesture to build and maintain the therapeutic relationship. Previous research in vocal expression of emotions (see Scherer et al. 2003) showed evidence that faster speaking rate can be linked to a greater arousal state of the speaker. It can be said that the medical students might be using increased speaking rate at the initial stage of the consultation to show that they were highly involved in speaking to the patient while keeping the mood of conversation reasonably upbeat (the speech samples analysed in this study did not include any particularly sad verbal content). It would be interesting to know why such gesture was not used in stage of explanation and planning. A tentative interpretation may be that the information load was higher when explaining and planning with the patient than at the beginning of the consultation when mainly greetings and introductory messages were exchanged. The use of speaking rate to indicate involvement and interest might have given way to the complexity of the verbal content requiring the speech to slow down. It is necessary for professionalism to be present at all stages of communication with patients (Arnold & Stern 2006, Kirk 2007). Therefore medical students and professionals need to be aware of the interaction between the attitudinal function and prosodic features and content of speech so that they can use other communicative resources (e.g. other prosodic features, non-verbal behaviour and choice of words) to demonstrate professionalism.

There is a considerable body of work on the association between the acoustic properties of prosodic features and the affective information communicated in speech (e.g. Basne & Scherer 1996, Schröder 2001, Pfizinger & Kaernbach 2008, Spackman et al. 2009, see also Scherer 1986, Scherer et al. 2003for summary of evidence). Most of the previous work focused on the vocal expression of emotions, especially the basic ones. The current study adds to the collection of evidence from the perspective of professional attitudes communicated through speech prosody in particular. The reason for basic emotions to be included in existing studies seems apparent: there has been great advancement in the theories of vocal emotional expression and thus solid grounds for investigative studies being carried out. While direct reports specifically on the vocal attitudinal expression are rarer in the literature, the current study has demonstrated that the mechanisms and resources for attitudinal and emotional expression in speech are likely to be similar. The adjustment in a speaker's prosodic features can be used to convey attitudes, including professional attitudes in healthcare settings. Therefore just as the way that prosodic features are important to express emotions, an active exploration by the healthcare professional in prosodic resources in the clinical interaction is important to display attitudes central to professionalism in communication.

There is also a multitude of reports on the differences in prosodic features between native and non-native speakers of English in existing literature, many of which included acoustic measures of the prosodic features under investigation (e.g. Zhang et al. 2008, Wennerstrom 1994, Verdugo 2005). Some of the previously studied distinctive non-native prosodic patterns were reported to affect intelligibility (Munro & Derwing 1998, Hahn 2004) or to contribute to the impression of foreign accents (Anderson-Hsieh et al. 1992, Munro & Derwin 1998, Zhang et al. 2008). While these previous findings were all relevant to the current study, fewer studies focused on the effects of non-native prosodic features on the expression of attitudinal information (Verdugo 2005), and they were not conducted in the healthcare context. Evidence from the current study on the difference between native and non-native speakers in English prosodic features enriches the existing research literature by observation from the viewpoint of the expression of professional attitudes in the healthcare context.

The current findings on the prosodic differences between native and non-native speaking medical students in the simulated clinical interaction agree with previous research evidence on non-native English prosody outside the healthcare context. This shows that while the non-native medical students in this study achieved good command of the English language as per the requirements for entering the medical school, and did not exhibit strong foreign accents in their vowels and consonants during the simulated consultation, their English prosody could still make them “sound different” from the native peers. That said, due to the high level of English proficiency of the non-native students in this study, it would still be difficult, if possible at all, for an untrained listener to guess correctly the L1 backgrounds of these speakers. Furthermore, the focus of this study was not the linguistic function of English prosody (e.g. marking word classes or sentence types) but the paralinguistic function (i.e. conveying attitudinal and emotional messages). In other words, this study did not investigate whether or not non-native English speaking medical students accomplished the same linguistic use of the prosodic features as their native speaking peers. Instead, this study revealed that the non-native students differed from the native peers in using the English prosodic features as a means of conveying professionalism during the clinical encounter. The potential reason for this and the implication for future medical communication skills training are discussed in the sections to follow.

The findings on the prosodic differences between the two speaker groups in the current study are pertinent to both ELT and medical communication skills teaching research. The current study examined the intonation patterns used by the two speaker groups when reading the same clinically relevant sentences in a professional manner. It was found that in intonation patterns the non-native English speaking medical students were not significantly different from their native peers, reflecting the good English command of this non-native group. While it is true that in ELT, it is desirable for foreign learners to achieve native-like intonation among other speech features, this teaching objective is at the most only limited to contextual appropriateness, which can relate more to the linguistic rather than paralinguistic function. Attitudinal expression using prosodic resources might be regarded as personal trait rather than part of the language skills. Even if expressing attitude in speech prosody were a commonly aimed for objective in teaching, it would not be easily assessed in the major English language tests (see Brown 2006, 2007 for a discussion on pronunciation rating for the IELTS speaking module). Moreover, one native English speaker might configure particular acoustic parameters differently from another to vocally communicate an intended attitude, but they would generally agree on a set of prosodic configurations for a specific attitude which would be commonly recognisable (Gobl & N íChasaide 2003, Spackman et al. 2009, see also the Brunswikian lens model introduced in Chapter 2). Some configurations for a particular attitudinal expression might be more efficient and might be preferred at the receiver's end. Idiosyncratic variation of course exists, but when the general recognisable prosodic configuration for a specific attitude is violated, confusion and misunderstanding can happen. This is similar to acting on the theatrical or cinematic stage: actors may have their own performing styles, but when the good actors portray an emotion or attitude, their vocal acting may be better appreciated by the audience than that of poor actors. Therefore the effective use of prosodic features to convey professionalism in medical communication, which is not totally dissimilar to emotional portrayal, is an objective beyond usual ELT outcomes. The improvement of the current teaching and assessment to benefit the needs of the non-native English speaking doctors (and medical students) in the UK and consequently patient experience needs evidence from research in a context directly related to healthcare services. The current study used speech data from communication in the clinical context, and as a result the evidence from the study has direct implication for medical communication training especially for non-native English speaking healthcare professionals (and trainees).

Furthermore, the influence of prosodic features on the listeners' perception of the speakers' professional attitudes in medical communication was found to be significant in this study, as previous perceptual studies (Scherer et al. 1984, Banse & Scherer 1996, Gobl & N íChasaide 2003) suggested. Not only were the different attitudes encoded by the medical students through prosodic differences in the sentence reading stimuli perceived by the listeners to convey significantly different levels of professionalism (lacking vs. present), but prosodic features were found to be able to predict the level of perceived professionalism too. Messages in these findings are twofold. On the one hand, the medical students need to have a core set of attitudes that reflect professionalism in the clinical interaction. On the second hand, they must have the skills to convey the professional attitudes through speech prosody in the communication with patients.

The finding of the significant effect of intended attitudes through speech prosody signifies the role of prosodic features in health communication. As the experiment showed, the same speech content with different attitudinal intentions encoded with prosodic features can result in different levels of perceived professionalism. This is clearly a case of "it's not what you say, but how you say it". That is to say, part of patient experience of the communication with healthcare professionals is beyond the content of speech, and extends to attitudinal information communicated through the prosody of speech. It is the "extra bit of information" of the professional attitudes conveyed through speech prosody, perhaps also with non-verbal features such as eye contact etc., that makes health communication more than just the exchange of information. In other words, the prosodic expression of professional attitudes in the communication with patients is capable of marking good quality of care in terms of patient communication experience.

Current medical communication training curricula, especially those using the Calgary-Cambridge model, tend to treat attitudes as a secondary outcome in favour of focusing largely on a set of skills for exchanging information and structuring the consultation. Although non-verbal skills for building doctor-patient relationship are introduced, prosodic features are often mentioned without further discussion. For example, students are not encouraged in the training model to explore the link between prosodic features and the attitudinal information that they can convey. The finding that clinically relevant sentences read with different prosodically encoded attitudes can lead to different levels of perceived attitudes therefore signals a gap in current medical communication training which currently emphasises cognitive rather than affective skills.

While the communication of professional attitudes might be a cognitive process as well as an affective process, the medical students seem to be able to encode the set of attitudes spontaneously in unscripted speech through the use of prosodic features. This was observed when the listeners gave similar professionalism scale scores to the stimuli from sentences ready by the medical students in a professional manner and those from spontaneous utterances in the simulated consultation. In fact the students used greater adjustments in the acoustic parameters of prosodic features while displaying professionalism in the simulated consultation, which was shown in the acoustic study (Section 4.5.2.1). On the one hand, it is true that the communication of attitudes differs from the display of emotions in the level of cognitive control, as some kinds of vocal expression of high-arousal emotions such as unsuppressed hot anger might be involuntary (Scherer et al. 2003). On the other hand, emotional expression and attitudinal expression may also be similar, naturally occurring processes. In both processes, there are some shared physiological and psychological bases (Scherer 2003, Clore & Schnall 2005) as well as shared communicative devices, i.e. prosodic features (Nooteboom 1997). Though attitudinal expression can be natural, the importance of cognitive skills to express attitudes effectively cannot be denied. Without developing and possessing the necessary skills to have the use of prosodic features under cognitive control, the communication of attitudes might not be effective, just like an actor with poor performing skills might not vocally portray emotions as appropriately as a skilful actor. Encouragingly, the inference can be made that that if medical students or professionals have the skills of using prosodic features to express attitudes, it is likely that they will do so when communicating with patients. However, they also need to have developed a set of professional attitudes appropriate to the communication with patients (Kirk 2007, Martin et al. 2002, Cushing et al. 2005, Woloschuk 2004). Lacking either the prosodic skills or the professional attitudes might contribute to a low level (or a lack) of professionalism perceived by patients.

The findings on prosodic features as predictors of perceived professionalism in medical communication were expected given the related evidence in existing literature on the emotional perception through prosodic features. However, compared attitudinal perception, evidence in previous research is more available on the perception through prosodic features of basic emotions than on that of attitudes, with just a few exceptions reviewed in the current study. Notable among these were attitudinal judgments on aspects such as “reproachful”, “friendly”, “insecure” or “understanding” included in a study on the perception of these through vocal cues by Scherer et al. (1984:1348). Similarly attitudes such as “friendly/hostile” and “timid/confident” were studied together with emotions such as “content/angry” and “sad/happy” in Gobl & Ní Chasaide (2003). Such evidence on prosodic features as perceptual cues to speaker attitudes had importance implication for the current study, however, since none of these studies used healthcare-related speech data, their findings cannot have direct implication for medical communication skills training or practice. The current study, by adapting their methodologies to use clinically relevant speech, produced findings that are possible to be generalised to a range of educational and professional settings.

The significant effect of the English language background of the medical students on the perceived professionalism found in the perceptual study confirmed the effect of the same factor on their prosodic features found in the acoustic study. Previously research only confirmed that English language background could lead to observable prosodic differences (e.g. Wennerstrom 1994) or that prosodic cues could be used for the perception of speaker affects (e.g. Scherer et al. 2003). This study contributed evidence demonstrating that the English language backgrounds of the medical students were associated with differences in prosodic features which influenced the perception of professionalism. Figure 6.1 below is a model that illustrates the relationship between these variables

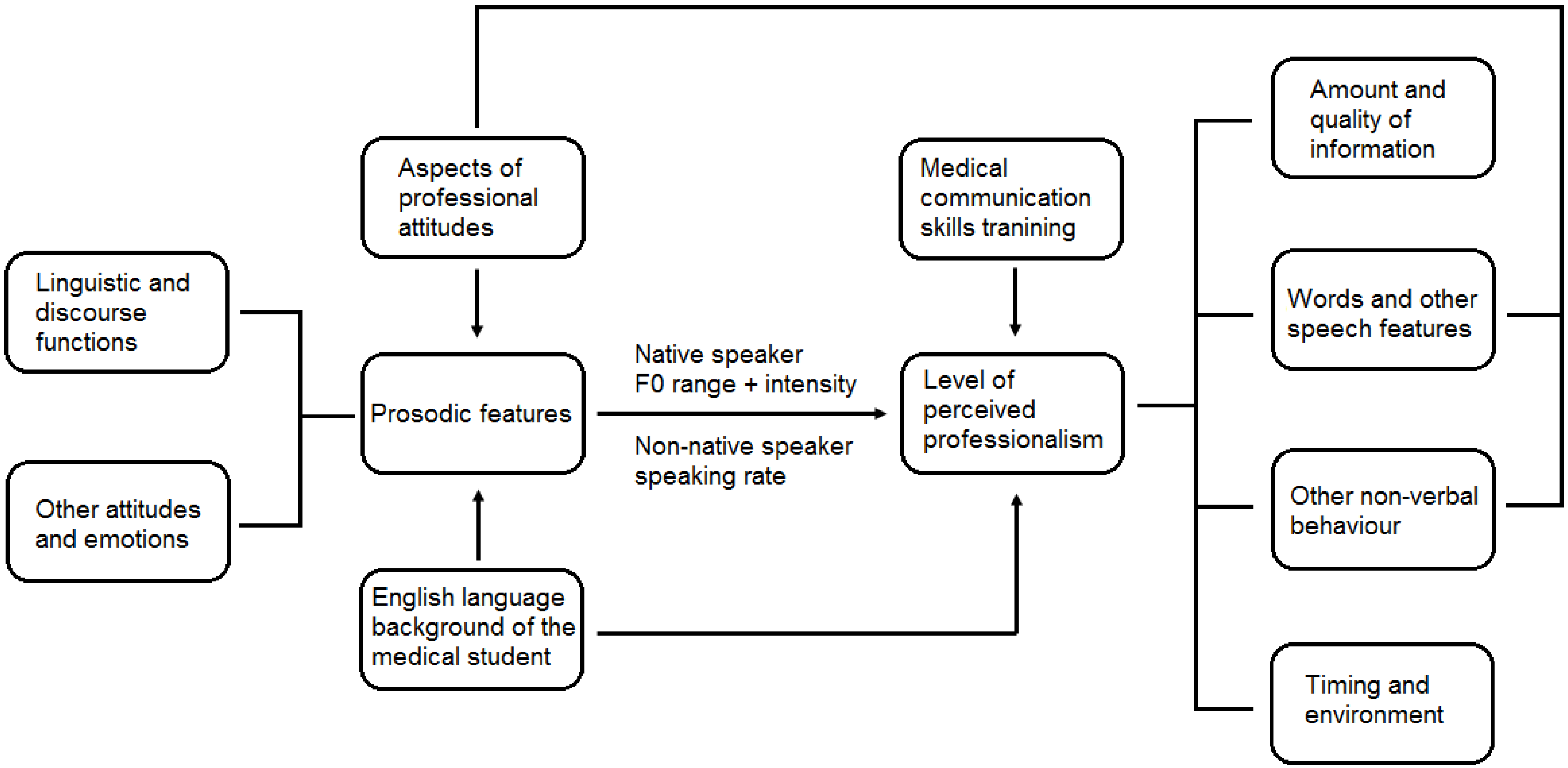


Figure 6.1 Prosodic features, perceived professionalism and associated factors in medical communication

The lines with arrows in Figure 6.1 represent the effects of the factors that have been ascertained in the study, while the plain lines represent potential relationships suggested in the literature but not investigated in this study. Similar models attempting to account for the relationship between these variables have not been reported in the existing literature to the author's knowledge. Yet the importance of this account is reflected in at least two aspects of reviewed literature. Firstly, both the large number of foreign doctors in the UK partly thanks to the increased mobility of doctors within the EU and disproportional number of foreign doctors facing complaints, referrals to or disciplinary procedures with the GMC were reported (Allen 2009, Devlin 2010). Secondly, challenges for doctors practising in a non-native language reach beyond linguistic knowledge and extend to language use (including attitudinal expression) (Van del Poel & Brunfaut 2010). Thirdly, although prosodic features have been known to convey attitudes and emotions in speech (Scherer et al. 2003), how prosodic features influence the perception of professional attitudes appropriate to clinical encounters remain uncertain in existing literature. The findings in the current study can serve as an important exploratory step in an enhanced understanding of the potential challenges faced by the large number of non-native English speaking doctors (and medical students) in the UK in terms of display professionalism in the communication with patients. Helping them with any potential communication difficulty can mean improving patient experience in the longer term and this will be addressed in the concluding Chapter 7.

6.2.2 Findings contrary to existing literature

Although most of the findings in the current study were anticipated from the findings of the existing literature, a few findings did not reflect the findings of previous studies. These are discussed below with possible explanations proposed.

Not all the examined acoustic parameters were significantly adjusted when the medical students were reading the clinically relevant sentences to convey aspects of professionalism. Lowered intensity was the only significant acoustic parameter associated with this attitudinal function. There are several potential explanations for this. One possibility is that while prosodic features have an attitudinal role in speech communication including that between doctors and patients, there might be more significant acoustic parameters than the ones examined in this study. F0 mean and range, dB intensity and speaking rate in this study can be considered as global prosodic features in a speech sequence. Although previous research demonstrated how these parameters could characterise speaker affective states (Banse & Scherer 1996, Scherer et al. 2003), it might be true that these parameters are more closely connected to arousal-based emotions than cognitive attitudes (Gobl & Ní Chasaide 2003). It is possible that the acoustic parameters of speech prosody themselves do not determine some professional attitudes which were cognitive, but work in coordination with other features of communication during the simulated clinical encounter. Though important, their effects are not independent of the communication context. As a first-step, exploratory investigation in the attitudinal function of prosodic features with evidence from native and non-native speakers in a simulated clinical setting, it is prudent to have chosen the global prosodic features as a starting point. These global prosodic features are likely to serve only a paralinguistic indexing function in speech. That is to say, they are likely to contribute to configuring a set of attitudinal aspects in speech communication. Other prosodic features such as F0 contour (perceived as the intonation pattern) can act as markers for cognitive attitudes in speech communication (see e.g. Ladd et al. 1985). The cognitive attitudes might be more effectively reflected in speech by prosodic features that are linguistic (e.g. intonation patterns distinguishing sentence types) rather than paralinguistic in nature, as well as verbal content. Such prosodic features are often restricted to specific units in the linguistic structures, for example, the F0 mean and range in the phrasal accents, and can therefore be called local prosodic features. It might be the case that some local prosodic features have a stronger indexing function or even a marking function (similar to intonation) than the global prosodic features investigated in this study. Chen et al. (2004) is an example of a study that investigated the effect of prosodic feature in which the researchers experimentally manipulated the F0 in the nuclear tones (among other features, by means of speech resynthesis) for listeners to rate the perceived meanings of friendly, confident, surprised and emphatic. The current study did not adopt a similar design due to the lack of theoretical prediction on the attitudinal meanings associated with professionalism in health communication conveyed through local prosodic feature adjustment. When the theoretical support becomes available, future studies planning to investigate the attitudinal function of local prosodic features can benefit from testing hypotheses about relationships between those features and specific attitudes (e.g. rude or condescending, rather than a set of attitudinal aspects) motivated by linguistic and psychological theories.

Another possibility is that the measures (i.e. Hz for F0 measurements and dB mean for intensity) of the prosodic features in this study reflected the acoustic properties more than the perceptual equivalents. Psycho-acoustic measures (e.g. ERB for pitch) of the features might produce different results in the statistical analysis. However, the perceptual experiment following the acoustic investigation confirmed that the acoustic measures (especially F0 range and mean dB intensity for native speaking medical students) also accounted for the attitudinal information communicated in perceptual terms. The adopted statistical methods might also have consequences for the findings. Due to the non-normal distribution of some sets of the acoustic data, a non-parametric test was carried out, which might not be the most powerful technique to answer the question. Furthermore, since the question concerns four prosodic features, Bonferroni correction was carried out to account for the multiple comparisons. This technique resulted in a conservative significance level which might explain why not all the prosodic differences were found significant. Nevertheless, after the correction, intensity was still found to show differences between two sets of attitudinal expressions for both genders, and perhaps suggested a role even more significant than the other prosodic features studied.

Similarly not all the prosodic features of the native and the non-native speaking medical students were found to be significantly different in the sentences read in a professional manner. In particular only speech intensity was found to be significantly lower in the non-native females than in their native peers (with α adjusted to .0125 by Bonferroni correction) in this condition. Previous literature, however, also reported differences between native and non-native speakers of English in F0 range in attitudinal expressions (e.g. certainty, Verdugo 2005) as well as in sentence reading with no attitudinal control (e.g. Binghadeer 2008). One possible explanation is that the English proficiency of the non-native participants in the current study was higher compared to previous studies (Verdugo 2005, Binghadeer 2008). Such finding, though not predicted by existing literature, is particularly exciting, because it shows that the non-native English speaking medical students can, through sufficient training, become similar to their native peers in at least some prosodic features in medical communication that requires vocal display of professional attitudes. It also reaffirms the importance of communication skills training especially for those who communicate in English as a non-native language.

More challenges might lie in a possible explanation for the fact that none of the native vs. non-native prosodic differences in any consultation task reached the set significance level. That is to say the prosodic differences between the native and the non-native speaking students in the simulated consultation could not be located to a particular consultation task, despite visual plotting that suggested otherwise (Figures 4.1-4.4). This finding could mean that the non-native medical students would approximate native norms of the prosodic adjustment in all stages of the consultation in order to make their voice sound native like. Looking back at the sentence data, the two groups did show similar choices for intonation patters when reading the clinically relevant sentences in a professional manner. This suggested that the non-native medical students in this study might have acquired the sufficient phonological knowledge of using intonation patterns for attitudinal expression just like the native ones (c.f. Wennerstrom 1994). However, this might be more complex because it was later found that after the consultation tasks being accounted for, their English language backgrounds still affected the level of professionalism perceived from the voice stimuli contributed by the medical students. Though not possible to be entirely free from listener biases, it appeared possible after linking the findings of the acoustic investigation and the perceptual experiment, that even small prosodic differences may be associated with significant differences in the level of perceived professionalism in speech. The finding that prosodic features as predictors of perceived professionalism exhibited different patterns (Figure 6.1) for the native and the non-native speaking medical students could therefore be explained.

In a model (Figure 6.1) built from the findings in this study, it was suggested that although prosodic features used in medical communication could predict the level of professionalism perceived, the model worked in different patterns for the native and the non-native speaking medical students. In particular, the patterns differed in the prosodic features as significant predictors. For the native students, the significant effects were from the F0 range and speech intensity in their simulated clinical interaction, with higher F0 range and lower intensity predicting higher level of professionalism perceived. For the non-native students, the significant effect was from their speaking rate only, with higher speaking rate predicting higher level of professionalism perceived. Although evidence from the existing literature (e.g. Banse & Scherer 1996, Scherer et al, 2003) suggested these significant effects, such previous evidence could not account for the different patterns between native and non-native speakers for the effects. The current study was able to take previous research findings in the vocal communication of speaker affects further to the area of conveyed professionalism in doctor-patient communication with evidence from both native and non-native speaker groups. Although a research design incorporating both the vocal encoding and the decoding through vocal parameters of speaker affects was advocated (Banse & Scherer 1996, Scherer 2003), past research rarely used vocal stimuli representing the speech of non-native speakers of English. Placing the findings of the current study in the context of multicultural/multilingual characteristics of UK health care today, more similarly collected evidence is needed so that perhaps meta-analysis can be conducted for this very topic in the future.

6.3 Significance of the findings and the methodological implications

It is clear that more work in the future is needed on this identified gap of research to consolidate the understanding of the use of prosodic features in health communication especially as a means to convey professionalism. Detailed recommendations are made in Chapter 7. Nevertheless the significance of the current study findings can be commented on with regard to several issues.

The study indicates that without prosodic features appropriately used in the communication with patients, there is the risk that attitudinal information might not be effectively conveyed. Attitudinal information should not be viewed as of secondary importance, i.e. after verbal content in clinical encounters. As reviewed in Chapter 2, doctor-patient communication is more than the simple exchange of information, but is a means of building the therapeutic relationship itself. Thus the display of professional attitudes in clinical encounters is directly linked to patient experience which in turn is an element of the quality of care. The findings on the attitudinal function of prosodic features in medical communication can therefore be interpreted as to suggest that prosodic features used in clinical encounters can influence patient experience of professionalism. The quality of care needs the display of professional (e.g. caring, compassionate) attitudes (Francis 2013), so prosodic features as the communicative resource for attitudinal expression must be actively explored.

Another issue of importance is that the level of perceived professionalism examined in this study is also directly linked to potential service users' experience of professional attitudes communicated in the clinical encounters. It conforms to the attitudinal objectives for medical students and qualified doctors set by the GMC and with the learning outcomes of the Calgary-Cambridge model of communication skills training. The higher the level of perceived professionalism is, the better the expectation of patient-centred care (see also Chapter 2) is met. Therefore examining the level of perceived professionalism (as this study has done) is not only of research importance, but also of educational and practical importance. The implications for education and practice are detailed in Chapter 7.

By investigating the prosodic features used for the intended attitudinal communication together with the perceived attitudes communicated through the prosodic features, this study has been able to establish the link between the two ends of the attitudinal communication, i.e. the doctor and the patient, through the channel of speech prosody. This is a response to the call for research combining both encoding and decoding studies of vocal expression of speaker affects, ideally from an interdisciplinary perspective (see Banse & Scherer 1986, Scherer 2003). Despite this, medical communication studies have not benefited enough from investigation using phonetic and psychological research methods, as reported in Chapter 2. Evidence of the role of speech prosody in attitudinal communication, especially in a professional context is more comprehensive from integrated research such as the current study than that from separate encoding and decoding studies, thanks to the advantages in a more complex design. Such evidence contributes to enhanced understanding of the features of the communication of professionalism, including prosodic features, in clinical encounters so that medical communication training and practice can be informed and ultimately patient experience can be improved.

Though interdisciplinary in nature, the current study has made contribution the knowledge of the related individual disciplines. There was scarcely any previous research in prosodic features that used data collected from medical communication. Studies on the vocal expression of speaker affects often examined spoken portrayals by actors (see Scherer 2003: Section 2.1.3). Meanwhile, studies comparing non-native prosodic features with native norms often collected speech data in classroom conditions. Furthermore, studies on intonation patterns in native English varieties often analysed utterances that were not from real life situations.

Nonetheless, acoustic findings of this study confirmed many of those from previous research. In particular, it was confirmed that differences in prosodic features existed between native and non-native speakers of English, regardless of the communication context. Acoustic analysis also confirmed that prosodic features varied according to the intended attitudes. This effect is similar to that of speaker emotions encoded in speech. This similarity in effect is very interesting because it reflects the commonality between attitudes and emotions not only in the internal appraisal process but also in the external communication behaviour (see also Scherer 2003: Table 4). Because of these aspects that attitudes and emotions have in common, attitudinal communication and emotional communication can be seen processes that are very similar, if not the same. While there are many valid reasons for doctors to control their own feelings when communicating with patients, the suppression of emotional expression might have a negative side effect which is the lack of effective attitudinal display. As reviewed in Chapter 2, the absence of displayed professional attitudes in medical communication can have a negative impact on patient experience. Therefore it seems desirable for communication skills training to actively encourage the vocal expression of professional attitudes while practising the control of personal emotions. Unfortunately, the exploration of attitudinal aspects in communication skills training might not fit in predominantly skills-oriented training models that exclude the active training of attitudes.

Furthermore, as previous research in prosodic features regardless of the design generally used speech data not collected in medical communication, the findings could only have limited generalisability to the clinical context. This is because doctor-patient interaction has a communication goal beyond the simple exchange of information in patient-centred care, which is building the therapeutic relationship. Since the communication with patients is part of the therapeutic relationship, it is quite unique for its attitudinal intentions which must reflect essential aspects of professionalism. Failing to do so can result in negative consequences, including compromised patient experience and in some cases formal complaints. Evidence from this study however, is from speech data in the clinical context and applies directly to communication skills of medical students and the messages are also useful for qualified professionals interacting closely with patients.

More importantly, evidence from this study suggests that for a better understanding of English prosodic features in professional communication, including that in the clinical setting, research can lead to more comprehensive findings when the investigating is not limited to a single discipline. On the one hand, as mentioned above, linguistic research on prosodic features and psychological research on speaker affects communication through speech prosody seldom collected data from doctor-patient communication, which limited the generalisability of their findings. On the other hand, medical communication research and training texts could have been informed more thoroughly if evidence from linguistic and psychological studies on the expression of attitudinal information were incorporated. Previous medical communication research rarely employed linguistic investigation methods, with Roberts et al. (2003) and Van de Poel & Brunfaut (2010) among the very few exceptions. Even so these more linguistic studies did not focus on attitudinal expression, and nor did they focus on the prosodic features as a channel of attitudinal expression. This study clearly shows that such linguistic investigation methods can be fruitful in medical communication research, and should be used by future studies with aims of understanding the display of professional attitudes in clinical interactions and above all improving patient experience in high quality health care.

Adopting a complex research design for a first step exploratory investigation such as the current study has some consequences to bear. As discussed above, the global prosodic features i.e. F0 mean and range, mean dB intensity and speaking rate were studied which represent the general characteristics of the speech prosody of an utterance. This were chosen as a starting point due to the lack of existing literature suggesting specific prosodic features as to be associated with distinct attitudinal meanings (Banse & Scherer 1996, Scherer 2003). Though these features were known to express speaker attitudes, their effects on the perception of listeners might be subtle. It would be interesting to know if such subtle effects could be detected in the perceptual study stage.

In the perceptual study, the factors of interest, (i.e. intended attitudes, speaker attributes, communication modes, consultation tasks and prosodic features) had already been established and measured as the independent variables for the analysis, thanks to the investigation of a previous stage. The verbal contents of the speech stimuli were also controlled so as not to confound the effects of the factors of interest in the perceptual study. In the sentence reading stimuli set, the sentences read by all the medical students were the same, so the effect of the text could be accounted for. In the simulated consultation stimuli set, the attitudinal information in the transcript of the stimuli were judged by a group of 20 native speakers of English to be neutral/not overlapping with attitudinal aspects of interest in the perceptual study. Furthermore the simulated consultation stimuli were from the same scenario with the same simulated patient performing the same consultation tasks. Thus the verbal content of these stimuli was considered unlikely to interfere with the perception of professional attitudes of interest conditioned by the factors under investigation.

Another issue considered for the perceptual study methodology was that the listeners participating in the experiment might base their attitudinal judgments on the English (including foreign) accents of the medical students instead of the prosodic features in their speech. The design of the perceptual study took this into consideration. To start with, it would not be possible to entirely divorce prosodic features from foreign English accents because the research literature reviewed suggests that the former were part of the latter. However, the prosodic features of interest in this study characterise speech events globally rather than locally, as discussed in Chapter 3. That is to say, the features under investigation affect local speech segments such as vowels and consonants so little that they can hardly be associated with any particular foreign English accent. The instruction of the perceptual experiment also asked the listeners very clearly to judge from the voices rather than the accents of the medical students as means to minimise such listener bias, which was not possible to avoid completely without compromising the validity of this study. In addition, because the non-native English speaking medical students in this study all had prolonged immersion in an English speaking country, their accents (in vowel and consonant pronunciation) could only be judged as mildly foreign at the most, and would be unlikely to be subject to a strong listener bias. In fact the accents of the non-native English speaking medical students in this study would not be reliably identified without training in linguistics or phonetics which none of the listeners had received. Due to the absence of any strong foreign accent in the non-native group, listener judgments were not likely to be affected by this. In other words, it was quite certain that the effects on listener judgments were from prosodic features rather than foreign accented vowels and consonants in speech. That said, findings from this study should not be generalised to a non-native population with lower English proficiency/shorter immersion without due caution exactly because of the sample used. Finally, although the verbal content and consequently accents of the speech stimuli can be masked by common techniques such as low-pass filtering, it is not ideal for this study because these can affect some of the prosodic features under investigation (for example low-pass filtering removes acoustic energy in the high frequencies and therefore alters intensity). More importantly speech stimuli free of verbal content can hardly represent medical communication and can thus compromise the validity of the study. Future studies wishing to similarly study non-native speakers in expressing attitudes through speech prosody will also have to take all of the above into consideration.

A third issue to consider regarding the perceptual study methodology was the instrument used. The professionalism scale used in this study was purposely designed using items chosen according to the Calgary-Cambridge guide which represented the demonstrated professionalism in communication with patients. The face validity of the instrument was therefore apparent, and the internal consistency reliability of the instrument also appeared good when tested using pilot data. However, it would be unrealistic to expect a high inter-rater reliability of this instrument mainly due to the nature of attitudinal perception. Unlike observable behaviours and unlike self-reported attitudes (what one says he or she believes, intends and values), perceived attitudes can naturally vary from one observer to the next. Since observable behaviours are often judged against a set of outward displays or markers accordingly, which can be compared to computers performing automatic scans, the inter-rater reliability will be naturally high. However, when the outward behavioural displays or markers are used to infer the inward attitudes, a complex cognitive process will be involved on the part of the observers. Understandably, variability will arise due to the ability of the observers, the observation environment and so on. This is not to say that any measure of attitude is inherently flawed, but to acknowledge the changing and subjective nature of attitudinal judgment. Similar comment on judgment of speaker emotions was given in Scherer et al. (1984) as a justification for not to establish the inter-rater reliability using scales of this nature. Test-retest reliability was not established in Scherer et al. (1984) for the same reason, but the test-retest reliability was assessed for the instrument, i.e. the professionalism scale used in the current study, and the result was encouraging (average of significant ICC = .72). Since attitude is temporally unconstrained compared with emotion (Clore & Schnall 2005), perhaps it is not surprising to find that the two ratings for the same stimuli by the participants within a short period of time agreed moderately well. Given all the limitations, perceived attitudes could still be regarded as more objective than self-reported attitudes, and would therefore suitably represent the “real” attitudes that matter in a professional communication context. Thus perceived attitudes, at least the aspects relevant to professional communication, can be reliably measured after taken the limitations into consideration. Future studies of professional attitudes in medical communication are likely to benefit from using this or a similar instrument to measure perceived attitudes instead of using self-reported attitude measures. Perceived attitudes can also be more important measures than observed behaviours if the study aim is focused on the communication of professionalism in healthcare settings, because the rated quality of the conveyed professionalism matters more than the amount of the related communication behaviour.

It is prudent to point out that the professionalism scale, being designed for the current study, has not been validated previously. In fact the perceptual study itself can be regarded as a pilot for investigating the effects of prosodic features (possibly also other communication features in future studies) on the perceived level of professionalism.

Two issues emerging from applying the statistical modelling should be noted. One of these was identified in the model to predict the effects of the gender, English language background and training level of the medical student together with the interactions between these factors on the level of perceived professionalism. Although the model was able to confirm the significant positive contribution of the native English language background and the high training level after controlling for consultation task, the model itself was found to be non-significant (F (26,297) = 1.47, p = .07). Examining the other statistics shed some light on the possibility for this. A substantial shrinkage was suggested by the goodness of fit measures (R^2 = .11, *adjusted* R^2 = .04) indicating that some explanatory factors might be missing from the model. This was not unexpected, because the prosodic features were not yet included in this model. The design of the perceptual study intended to solve the jigsaw puzzle one step at a time, that is, the understanding of the effects of speaker attributes should precede the examination of the effects of prosodic features. Furthermore, the assumption of homogeneity of variance in the ANOVA was not met (p < .01, Levene's Test of Equality of Error variances), restricting the generalisability of this model beyond the sample of this study. However, this statistical model did suggest it would be more appropriate to treating the native and the non-native groups separately when testing a second statistical model of the effects of prosodic features.

This second statistical model only examined the potential effects of any of the prosodic features under investigation (F0 mean and range, intensity and speaking rate) on the level of perceived professionalism after understanding the effects that the attributes of the medical students had on the outcome. The model however, showed different effects among the native and the non-native speaking medical students, as discussed earlier. Although it was expected that the model might fit the native group and the non-native group differently, the fact that different prosodic features showed significant effects for the two groups was in some unexpected. A possible explanation might need to go back to the acoustic differences in speech prosody between the two groups. Because of the acoustic differences, some prosodic features that can be used by native medicals students might not work for non-native medical students. Since it could not be certain that the results were completely free from bias, recommendation for improvement cannot be made except for within the two groups. Goodness of fit measures for the non-native group (R^2= .06, *adjusted* R^2= .04) could suggest that there seemed to be a lot more than speech prosody that might also matter in conveying professional attitudes in the communication with patients. Compared to the native medical students who can rely a bit more on speech prosody to display professionalism (R^2= .15, *adjusted* R^2= .13), the non-native students must overcome more difficulty to achieve the same goal. While they need a strategy for self-motivated communication skills learning as the GMC (2009) required, supplementary solutions might also incorporate the development of (cross-) cultural awareness as well as professional attitudes.

6.4 Further reflections

The current study through its stages of investigation and synthesis of findings has shown the contribution of speech prosody in medical communication in terms of conveying the professionalism essential to building the therapeutic relationship with patients. In particular, it has demonstrated the attitudinal function of prosodic features in a simulated clinical interaction and speaker attributes that affect their acoustic properties. Interestingly the attitudinal effects as well as speaker attribute effects are also significant for the listener perception of professionalism in medical communication, and so are the effects of the prosodic features. As a first step exploratory study, it has answered several of the “whether” questions, but a few “how” or “what” questions have not been answered due to its nature, scope and design.

One message this study has for future research is that prosodic features can and should be studied with realistic speech data, especially from a perspective of their attitudinal function. Through acoustic analysis and perceptual experiment, the study has ascertained the link between particular prosodic features and aspects of professionalism conveyed in a realistic simulated clinical interaction. Previous studies using acted, classroom style, or resynthesised speech limited their findings to either vocally expressed basic emotions or aspects of attitudes not directly relevant to the healthcare context. The questions for future studies in this area therefore are: "what other prosodic features should be studied with clinically relevant speech data" and "how should these features be investigated". Work in the future motivated by linguistic and psychological theories can test hypothesis regarding more local prosodic features (e.g. F0, intensity or spectral parameters in the phrasal accent), which might be additionally associated with cognitive attitudes. Future work on the attitudinal perception through prosodic features can, again based on strong theoretical prediction, select a smaller, more specific set of attitudes to test in the experiment. Resynthesis as a way to control for speech content can still be useful, but when used in future research, care must be used so that the resynthesised speech stimuli will not be of undermined validity. Future studies in this area should particularly bear in mind the purpose of informing medical education and practice, since the communication with patients can be a therapeutic tool itself.

One other message from this study for future work is that professional attitudes and their expression should be studied more in medical communication research. What has been included in the current study can be regarded as the essence of professionalism in communication, but there are other aspects of professional attitudes that might be particularly linked to a certain consultation task (e.g. gathering information) or a certain patient group (e.g. palliative care). A variety of research methods can be beneficial to this area of work in the future. For example, the comparison between self-reported attitudinal measures and observed attitudinal measures can be interesting in quantitative studies, and qualitative methods such as discourse analysis and focus group can be useful in developing a deeper understanding of certain professional attitudes (e.g. empathy). However, when using speech data from real clinical settings, care should be used to minimise any confounding from the speech content or communication environment, due to the subjective and adaptable nature of attitudinal judgement.

The current study has benefited from adapting the established methodology in emotional communication research for the exploration of professional attitudes. It will be interesting to see the development of theoretical frameworks that can incorporate the display and perception of attitude in the professional communication settings. Existing literature has already suggested the overlap between attitude and emotion, especially in terms of their appraisal and evaluative object (Scherer 2003, Clore & Schnall 2005). Emerging empirical evidence from the likes of the current research can be expected to encourage the work of theorists in the future.

This chapter has thus served to interpret and reflect on the findings of the current study, especially in terms of its contribution to knowledge. Implications and limitations in this study are presented in the next, final chapter in which the concluding remarks are also given.

Chapter 7 Conclusion

7.1 General summary

The overarching aim of this study was to ascertain the role of speech prosody in medical communication in terms of conveying professionalism. To this end, investigations were carried out on the acoustic properties of the prosodic features found in utterances to convey attitudinal intentions during the simulated clinical encounter and on the perceptual properties of the prosodic features with encoded attitudes in the communication. In summary, this study has reached three main conclusions below.

The first main conclusion is that medical students adapt their speech and use prosodic features to convey professional attitudes in medical communication when asked. Differences in the acoustic values of prosodic features were found to be significantly different when medical students read out the same clinically relevant sentences in a professional manner and in an unprofessional manner. Similarly, significant differences in the acoustic value of prosodic features were also found in different tasks within a simulated consultation during which medical students were expected to convey the aspects of professionalism in communication that are essential to the building of the partnership with the (simulated) patient.

The second main conclusion is that native and non-native English speaking medical students differed significantly in the use of prosodic features to communicate in the simulated clinical setting. The difference in the use of prosodic features between native and non-native medical students was found when reading out the same clinically relevant sentences. Native and non-native medical students were also found to differ in the use of prosodic features when interviewing a simulated patient. However, such prosodic difference could not be located to any specific task in the simulated consultation.

The third main conclusion is that prosodic features influence listeners' perceptions of professional attitudes in the medical communication. This effect was found for utterances from both native students and non-native students. However the effect was greater in the native group than in the non-native group. Furthermore, prosodic features were shown to be able to account for some of the variance in the level of perceived professionalism. Tasks in the simulated consultation and the attributes of the medical students also had a significant effect on this.

Evidence from this research confirmed the effects of the intended attitudes, the consultation tasks and the medical student's English language background on the acoustic properties of prosodic features. Findings also confirmed the effects of the intended attitudes, the communication modes, the consultation tasks, the medical student's English language background, the medical student's training level and the acoustic properties of prosodic features on the level of perceived professionalism. Interestingly, when examining the two sets of effects together, that is, by establishing links among medical student English language background, prosodic features and perceived professionalism, the effects emerged to be sophisticated: the effects of prosodic features on perceived professionalism seemed to be dependent on both the English language background and the significant acoustic parameters. There are strong indications that speech prosody in medical communication contribute to conveying professionalism in the interaction with patients, and this contribution is even more important if the medical student or professional is a native speaker of the English language. Furthermore, although prosodic features can have an effect on the professionalism communicated regardless of the English language background of the medical student or professional, the acoustic parameters that are effective can be different. In the perceptual experiment carried out in this study, the effective parameters in speech prosody to convey a higher level of professionalism for the native speaking medical students were found to be higher F0 range and lower mean dB intensity. That is to say, for a native medical student, a wider range of pitch excursions and a quieter voice were associated with higher level of professionalism perceived by a potential patient. However for a non-native medical student, only an increased speaking rate was associated with the communication of professionalism. On a potential patient's part, they might perceive a faster speaking medical student whose mother tongue is not English as displaying a higher level of professionalism.

The overarching question raised in the current study whether speech prosody matters in health communication has therefore received an affirmative answer for both native and non-native English speaking medical students. Such evidence has been gathered from the analysis of acoustic and perceptual data of a simulated clinical interaction. It has been suggested in Chapter 6 that future research in this area is likely to benefit from the analysis of data of real clinical interactions. Findings of this suggest have also confirmed a place for speech prosody as a means of contributing to the conveying of professionalism in medical communication skills training. Sections 7.2 and 7.3 below provide the implications of this study for research, education and practice.

7.2 Implication for research

The findings in this study highlight the need for interdisciplinary research on the use of prosodic features in professional, especially medical communication. By collecting clinically relevant speech data and applying established methods in linguistics and psychology, this study has demonstrated the role of speech prosody in expressing professional attitudes when communicating with patients for both native and non-native speaking medical students and professionals. This study has also shown the attitudinal role of prosodic features in the clinical context is not dissimilar to that found in other communication contexts. However, verifying the attitudinal role of prosodic features in medical communication is particularly important given the substantial proportions of patient complaints related to staff attitude both in the NHS and in the healthcare service elsewhere in the world. The evidence of the attitudinal role of prosodic features lies not only in the acoustic differences due to different attitudinal loadings but also in the differences in perceived professionalism due to different acoustic properties. This latter perceptual finding echoes the accumulated evidence of the role of speech prosody in the communication of speaker affects (see Scherer 1986, Scherer et al. 2003)

Furthermore, although differences between native and non-native speakers of English in prosodic features were frequently reported in linguistic research literature (Chapter 2), the nature of the speech data used in previous studies limited the generalisability to the healthcare context. To overcome this limitation, this study has utilised clinically relevant speech data collected from both native and non-native medical students. Since the speech materials in this study are topic- (and to a lesser degree content-) controlled, simulated clinical interactions, the evidence generated can be used to inform future training and, to some extent, practice.

Researchers in linguistics should be encouraged to test their theories with data that are more relevant to real life communication, including that between doctors and patients, as well as develop theories that account for not only the communicative functions, but also the communication experience in these real life situations. Only by doing so can research in medical communication be better informed of the features in communication leading to outcomes of patient experience. As pointed out earlier, the goal of medical communication should not be exclusively the exchange of information but should be part of the experience that maintains the therapeutic relationship. Therefore it will be of interest to see how such communication experience can be accounted for by research of the involved disciplines.

Likewise, psychologists should advance their theories so that the communication of speaker attitudes through prosodic features can also be accounted for in a systematic fashion. Currently there is a level of understanding of the relationship between the expression of basic emotions such as joy and fear and a range of prosodic features. It will be also interesting to have theoretical backing for speaker attitudes expressed in speech prosody from a psychological perspective. This is because although linguists and speech scientists now have theories that explain the characteristics of prosodic features in terms of their attitudinal function, to theoretically explain the perceptual realities of the prosodic expression of speaker attitudes remains a challenge for psychologists.

This study is among the first steps into informing medical communication research and training with empirical evidence, both acoustic and perceptual, on the attitudinal function of prosodic features used in the clinical context. It has also a focus on the non-native speakers who communicate in this professional context. The purpose of including a non-native speaking sample in this study is not only to ascertain any variation from native prosodic norms, but to understand what impact such variation might have on patient experience in terms of perceived attitudes as well. In addition, this study is an endeavour to bridge the gap in communication skills research which lies in evidence on the non-native English speakers in medical education and practice.

In this study, although the non-native medical students have native-like proficiency in the English language, they still differ from their native peers in both the acoustic properties of prosodic features and professionalism scale scores. If they are judged on the content of speech in the simulated interaction, they might not differ from the native medical students in the verbal attitudinal expressions (see Section 5.4.1 on the validation of voice stimuli). Evidence in this study therefore suggests that medical communication skills training can benefit from taking a step beyond teaching only the content (and structure) of information exchange and taking into account communication features (including speech prosody) that are responsible for expressing attitudes alongside other features of non-verbal behaviour. Furthermore attitudinal topics in the training cannot be only on the sterile cognitive entity divorced from any affective information. Although it is advisable for medical students and professionals to control personal emotions in the interaction with patients, measures should be taken to account for the affective needs from patients. Because of this, students should be encouraged in the training to explore all of the affective, cognitive and behavioural aspects of attitudes that are appropriate to the clinical context. After all, as mentioned earlier, if one possesses all aspects of professional attitudes beneficial to patient experience but do not express those in communication behaviour, patients will still perceive a lack of professional attitudes in the encounter.

7.3 Implication for education and practice

Although communication skills training in UK medical schools is evidence-based, the presumption of the training model can accordingly bias the scope and adoption of evidence. In a predominantly skills-oriented training model, attitudinal expression, including that through speech prosody, might either be suppressed or be merged with the so-called relationship skills in which features of speech prosody and non-verbal communication proper might be introduced without in-depth consideration. This together with even a healthy level of personal emotional control on the part of medical students can risk leading to a potential unemotional, disconnected perception of the communication as an outcome. However the exclusion of a dedicated topic on prosodic features for attitudinal expression can be a disadvantageous omission, since prosodic features are an organic part of an utterance and, as reviewed in Chapter 2, are responsible for a number of linguistic and paralinguistic functions in speech communication. Unlike the features of non-verbal communication proper (gesture, posture etc.) which function *parallel to* the speech units, prosodic features function *intertwined with* the verbal content in speech to enhance or modify the information conveyed. As this study shows, inappropriate use of prosodic features (so as to express attitudes that are not professional in the clinical context) lowers the level of professionalism conveyed and non-use of this paralinguistic function reduces the expressiveness in speech, both compromising the effectiveness of communication which should display a high level of professionalism. Since professional attitudes lie at the centre of the therapeutic relationship between clinicians and patients, they deserve their place in the medical communication skills training programme. Additional topics on the expression of professional attitudes should therefore be included in the communication skills training in which different channels of attitudinal expression, including speech prosody, should be discussed in depth. Without targeted training, there cannot be confidence in the presence of professional attitudes in the communication with patients, nor can medical students and professionals understand the wide range of effective communicate devices for the display of professionalism in the relationship with patients. Given the unmet patient expectations on professionalism in medical communication reflected in formal complaints as well as recent public enquiries, the urgency and necessity of target training can no longer be ignored. Furthermore, the differences of English prosodic features by native and non-native speaking medical students found in this study can also lead to differences in the level of professionalism conveyed in the communication with patients. Since the number of non-native English speaking medical students and professionals in the UK is not expected to fall significantly, enhanced communication skills training with topics on attitudinal expression in speech prosody and other channels can be particularly beneficial.

Evidence from this study supports the emphasis on the importance of effective medical communication in patient-centred care, especially with regard to demonstrating professional attitudes appropriate to the clinical context. The study shows that the same content of communication with different intended attitudes results in different level of perceived professionalism, and that speech prosody is a significant factor in the effective communication of professionalism. In the wake of the call for re-establishing a caring culture in the NHS (Francis 2013), the improvement in doctor-patient communication must go beyond the amount or content of information and extend to the enhancement of patient experience. This includes the display of professional attitudes in the doctor-patient relationship, which can be demonstrated in multiple channels including speech prosody as discussed in this study. Doctors (both in practice and in training) must possess and demonstrate the appropriate set of professional attitudes which can be expressed in speech prosody as well as other features in communication. In this way, patients are likely to perceive a high level of professionalism and therefore experience a higher quality of health care, when other factors in the service such as clinical treatment and service administration also work together towards patient satisfaction. In the current healthcare climate, patients expect more than high quality clinical treatment; they expect the treatment to be delivered with compassion, dignity and respect. These values lie at the core of the set of professional attitudes that should be communicated to the patients. Therefore patient expectation on medical communication is fulfilled not only by the information exchanged, but also by the way the information is exchanged. In the speech communication between doctors and patients, the information includes that on the attitudes that are present, and such information can only be effectively communicated if the channel of prosodic features is activated together with the speech content and the non-verbal features proper. In short, patients expect the appropriate level of information exchanged with apparent professionalism in a multimodal communication experience during the clinical encounter. This study serves as a step into better understanding of the modes and the professionalism of medical communication.

Furthermore, the international mobility in the medical profession (and education) has increased and many medical students and professionals in the UK are non-native speakers of the English language. Their effectiveness of communication with patients in English has drawn great attention. This study shows that English proficiency is not the only factor in this issue. Even medical students with a high level of English proficiency can vary from native patterns of using English prosodic features and convey a level of professionalism that is undermined. Therefore non-native English speaking medical students and professionals must actively explore all modes of speech communication to display their professional attitudes when interacting with native English speaking patients. Patient experience must not be compromised and highly effective and professionalism in the communication with patients is not an unachievable goal. With targeted training, native and non-native speaking medical students and professionals alike can give communication experience that is compassionate, dignified and respectful to the patients. In a multicultural society, non-native speaking medical students can, with a good understanding of the communication features contributing to the conveying of professionalism and their own cross-cultural awareness, excel as excellent communicators in the delivery high quality healthcare service.

7.4 Limitations and recommended future work

As a first study in its field, the current study is not free from limitations. These are due to either limited foundations from limited theoretical frameworks or from methodological decisions that are made before carrying out the investigation but which, in hindsight, are able to be improved.

A caveat needs to be noted regarding the small sample size of 16 medical students in the acoustic investigation, although in order to mitigate influence from potential confounders, a sample equally distributed in gender, English language background and training level was achieved. However, investigations of similar topics in disciplines such as linguistics has often used a comparable or even smaller sample size. Furthermore, all non-native speaking medical students were of a high level of English proficiency. Because of this, the findings are of limited generalisability to healthcare professionals of lower English proficiency, or a similarly wider population. Future research studying a singular prosodic feature in medical communication can potentially benefit from a larger, more representative sample without having to overcome the difficulty involved in an enormous amount of acoustic analysis. From hindsight, the evidence of the effect of prosodic differences (rather than foreign accent in vowels and consonants) on perceived professionalism could also be stronger, if analysis involving categorical judgment on accents (native vs. non-native) by listeners had been carried out. This lesson learnt should also inform future research design.

A related limitation lies in the prosodic features investigated in this study. As discussed in Chapter 6, it still cannot be certain if other prosodic features local to linguistic units also contribute to the expression of professional attitudes alongside the ones in the current study that represent the general characteristics of the prosodic of a longer speech event. This study chose the general, global features of speech prosody as a starting point because of the lack of provision in existing research in terms of predicting specific prosodic features likely to be capable of the expression of professional attitudes appropriate to clinical settings. For this reason, further research in this area is urgently needed.

Regarding linguistic research that influenced the understanding of the attitudinal role of prosodic features, especially intonation patterns, the current study has made little theoretical contribution. Compared with this, the current study has more theoretical implication for research on vocal communication of speaker affects by extending the scope of investigation to the aspects of attitude related to professionalism in health communication. Several researchers (e.g. Scherer 2003, Spackman et al. 2009) advocated the adoption of a version of Brunswikian lens model that explains the variable transmission of speaker affects between the distal cues at the encoder's end and the proximal percepts at the decoder's end. This study adds to the support of the Brunswikian lens theory with evidence on a similar process of attitudinal expression and perception through the cues of speech prosody. Furthermore, linguists and psychologists use appraisal theory in studying attitudinal and emotional expression, even though appraisal theory can refer to different realms of knowledge in their separate disciplines. Despite their sharing basic theoretical assumptions, researchers in these separate disciplines have not extensively collaborated in their work on the communication of attitude and emotion, and consequently might have limited the advancement of understanding in this area. This study, although not taking a stance in either version of the appraisal theory, does show promises for future work to bridge the theoretical gap currently hindering interdisciplinary research in attitudinal and emotional expression. In the future, empirical evidence such as the findings in this study can be used to test theoretical accounts, be they on the affect, judgment and appreciation in attitude or on the appraisal dimensions in emotion, as long as the aspects in the theoretical framework can be clearly defined when raising research questions.

7.5 Concluding remarks

Though a first-step exploration with limited theoretical contribution, this study attempts to deliver messages pertinent to both medical education and practice. For medical educators, this study highlights that communication skills training should go beyond the content and structure of information exchange, and should extend to the display of professionalism in the communication with patients as a means of maintaining the therapeutic relationship. In particular, communication skills training should make provision for in-depth discussion on features for communicating professional attitudes. These not only include features of non-verbal communication proper, but also include the prosodic features in speech. Prosodic features are especially powerful for this purpose since they enhance and modify the verbal information in speech as demonstrated by the current study findings. This study also highlights special training needs for medical students who are not native-speakers of English. The needs do not lie in the English language proficiency which is high and often native-like; the needs are in the effective use of English prosodic features for communicating professional attitudes in the clinical context. It is true that prosodic features matter in medical communication for both native and non-native speaking communicators. Without explicit training, the subtlety of English prosodic features especially with regards to their attitudinal meanings can be unnecessarily challenging for non-native medical students. In addition, including topics on using prosodic features for more effective communication of professional attitudes can benefit native and non-native students alike. As this study has shown with its findings, prosodic features used in medical communication have a significant effect on the level of professionalism perceived by the potential patient.

For medical practice, this study further highlights the benefits of doctors as good communicators to patient experience. Good communication is not only defined by what information is conveyed to patients but also how. This study suggests that attitudinal information in communication with patients may be as important as the "hard" factual information. While this study does not question the importance of putting doctors' personal emotion aside in medical communication, it does show that the therapeutic relationship should allow room for attitudinal display. Such attitudinal information incorporates both the affective and the cognitive domain, and represents the core values of the NHS and of the medical profession, that is quality care delivered with respect, dignity and compassion. A doctor's voice can convey the colour of warmth and caring through prosodic features to a high level that is difficult for printed patient information booklets to achieve. Recently there has been increasing volume of the public call for restoring the caring culture in the NHS. This can be responded to by not only professional attitudes towards patient-centred care but also good communication which is in turn fulfilled by the right manners of saying the right things to patients. Doing so, as this study suggests, can result in a better perception of professionalism. The excellence in medical communication contributes to excellence in care where patients are satisfied, compliant to treatment and with improved outcome. Furthermore if this can be achieved, it can be expected that complaints related to communication and attitudes will be minimised, thus helping to maximise the efficiency of service. In short, good communication with appropriate professional attitudes conveyed through speech prosody and other communication features benefits patients, the medical profession and the healthcare service.

In addition, for non-native English speaking doctors practising in the UK, this study suggests a potential hindrance to effective communication with patients may not necessarily lie in the English language proficiency, provided that they are competent users of the language. The potential may be the speech prosody responsible for expressing attitudes and emotions, a same challenge faced by native speaking doctors as well. While many native doctors might rely on their implicit understanding of the English language and manage without dedicated training, this challenge might appear to be more substantial for an untrained non-native speaker of the language. The GMC asks all UK doctors to be lifelong learners, non-native speaking doctors should be encouraged to learn and master English prosody as an essential channel of communication with patients. This together with great cultural awareness will effectively contribute to positive patient experience with the communication.

Finally, in clinical encounters as well as other sensitive situations in life, what matters in communication is not only the message itself, but also the way in which the message is conveyed (Ong et al. 1995). This significance is quite neatly captured in a line from the animation film *Cars* (Anderson & Lasseter 2006)

> "Was it just a 'Yeah, OK', or 'Yeah … OK', or 'Yea-yeah, OK'?"
>
> – Sally to "Lightning" McQueen

References

Ajzen, I., & Fishbein, M. (2005). The influence of attitudes on behavior. In D. Albarrac ń, B. T. Johnson & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 173-221). Mahwah, NJ: Erlbaum.

Albarrac ń, D., Zanna, M. P., Johnson, B. T., & Kumkale, G. T. (2005). Attitudes: Introduction and scope. In D. Albarrac ń, B. T. Johnson & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 3-19). Mahwah, NJ: Erlbaum.

Allen, I. (2009). Doctors crossing borders: Europe's new reality. *Canadian Medical Association Journal, 180*(2), 158-161. doi: 10.1503/cmaj.081817

Ambady, N., LaPlante, D., Nguyen, T., Rosenthal, R., Chaumeton, N., & Levinson, W. (2002). Surgeons' tone of voice: A clue to malpractice history. *Surgery, 132*(1), 5-9.

Anderson, D. K. (Producer), & Lasseter, J. (Director). (2006). *Cars* [Motion Picture]. United States: Walt Disney Pictures.

Anderson, K., Allan, D., & Finucane, P. (2001). A 30-month study of patient complaints at a major Australian hospital. *Journal of Quality in Clinical Practice, 21*(4), 109-111.

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning, 42*(4), 529-555.

Anvik, T., Grimstad, H., Baerheim, A., Fasmer, O. B., Gude, T., Hjortdahl, P., . . . Vaglum, P. (2008). Medical students' cognitive and affective attitudes towards learning and using communication skills – a nationwide cross-sectional study. *Medical Teacher, 30*(3), 272-279. doi: 10.1080/01421590701784356

Arnold, L., & Stern, D. T. (2006). What is medical professionalism? In D. T. Stern (Ed.), *Measuring medical professionalism* (pp. 15-38). New York: Oxford University Press.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70*(3), 614-636.

Baumann, M., Baumann, C., Le Bihan, E., & Chau, N. (2008). How patients perceive the therapeutic communications skills of their general practitioners, and how that perception affects adherence: use of the TCom-skill GP scale in a specific geographical area. *BMC Health Services Research, 8*(1), 244.

Beisecker, A. E., & Beisecker, T. D. (1993). Using Metaphors to Characterize Doctor-Patient Relationships: Paternalism Versus Consumerism. *Health Communication, 5*(1), 41-58. doi: 10.1207/s15327027hc0501_3

Binghadeer, N. (2008). An acoustic analysis of pitch range in the production of native and nonnative speakers of English. *Asian EFL Journal, 10*(4), 96-113.

Boersma, P., & Weenink, D. (2010). Praat: doing phonetics by computer (Version 5.1.31) [Computer Programme]. Retrieved from http://www.praat.org/

Brinkman, W. B., Geraghty, S. R., Lanphear, B. P., Khoury, J. C., Gonzalez Del Rey, J. A., DeWitt, T. G., & Britto, M. T. (2006). Evaluation of resident communication skills and professionalism: a matter of perspective? *Pediatrics, 118*(4), 1371-1379.

Brown, A. (2006). An examination of the rating process in the revised IELTS Speaking Test. *IELTS Research Reports, 3*, 41-70.

Brown, A. (2007). An investigation of the rating process in the IELTS oral interview. *IELTS collected papers: Research in speaking and writing assessment*, 98-139.

Brown, J. (2008). How clinical communication has become a core part of medical education in the UK. *Medical Education, 42*(3), 271-278. doi: 10.1111/j.1365-2923.2007.02955.x

Brown, J. B., Stewart, M., & Ryan, B. L. (2003). Outcomes of Patient-Provider Interaction. In T. L. Thompson, A. M. Dorsey, K. I. Miller & R. Parrott (Eds.), *Handbook of health communication* (pp. 141-161). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Chambers, J. K. (2009). *Sociolinguistic theory* (Revised ed.). Chichester: Wiley-Blackwell.

Chavan, R., Porter, C., & Sandramouli, S. (2007). Formal complaints at an eye hospital: a three-year analysis. *Clinical Governance: An International Journal, 12*(2), 85-92.

Chen, A., Gussenhoven, C., & Rietveld, T. (2004). Language-specificity in the perception of paralinguistic intonational meaning. *Language and Speech, 47*(4), 311-349.

Chin, J. J. (2002). Doctor-patient relationship: from medical paternalism to enhanced autonomy. *Singapore Medical Journal, 43*(3), 152-155.

Clore, G. L., & Schnall, S. (2005). The influence of affect on attitude. In D. Albarrac ń, B. T. Johnson & M. P. Zanna (Eds.), *Handbook of attitudes* (pp. 437-489). Mahwah: Erlbaum.

Cohen, J. J. (2006). Professionalism in medical education, an American perspective: from evidence to accountability. *Medical Education, 40*(7), 607-617. doi: 10.1111/j.1365-2929.2006.02512.x

Collins, B., & Mees, I. M. (2003). *Practical phonetics and phonology: A resource book for students*. London: Routledge.

Council of Europe. (2001). *The common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Couper-Kuhlen, E. (1986). *An introduction to English prosody*. London: Edward Arnold.

Coupland, N., & Bishop, H. (2007). Ideologised values for British accents. *Journal of Sociolinguistics, 11*(1), 74-93. doi: 10.1111/j.1467-9841.2007.00311.x

Crellin, J. K. (2005). *Public expectations and physician's responsibilities: voices of medical humanities*. Oxford: Radcliffe.

Cruttenden, A. (1997). *Intonation* (2nd ed.). Cambridge: Cambridge University Press.

Cruttenden, A. (2008). *Gimson's Pronunciation of English* (7th ed.). London: Hodder Education.

Cushing, A., Evans, D., & Hall, A. (2005). Medical students' attitudes and behaviour towards sexual health interviewing: short- and long-term evaluation of designated workshops. *Medical Teacher, 27*(5), 422-428. doi: 10.1080/01421590500046502

Deladisma, A. M., Cohen, M., Stevens, A., Wagner, P., Lok, B., Bernard, T., . . . Lind, D. S. (2007). Do medical students respond empathetically to a virtual patient? *American Journal of Surgery, 193*(6), 756-760. doi: 10.1016/j.amjsurg.2007.01.021

Department of Health. (2013a). *The NHS Constitution*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/170656/NHS_Constitution.pdf.

Department of Health. (2013b). *The Handbook to the NHS Constitution*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/170649/Handbook_to_the_NHS_Constitution.pdf.

DePaulo, B. M., & Friedman, H. S. (1998). Nonverbal communication. In D. T. Gilbert, S. T. Fiske & G. Lindzey (Eds.), *The handbook of social psychology* (4 ed., Vol. 2, pp. 3-40). Boston, MA: McGraw-Hill.

Deveugele, M., Derese, A., & De Maeseneer, J. (2002). Is GP-patient communication related to their perceptions of illness severity, coping and social support? *Social Science & Medicine, 55*(7), 1245-1253.

Devlin, K. (2010, 22 May). More foreign doctors struck off than British-trained medics, *The Daily Telegraph*.

DiMatteo, M. R., Reiter, R. C., & Gambone, J. C. (1994). Enhancing medication adherence through communication and informed collaborative choice. *Health Communication, 6*(4), 253.

Donders, A. R., van der Heijden, G. J., Stijnen, T., & Moons, K. G. (2006). Review: a gentle introduction to imputation of missing values. *Journal Clinical Epidemiology, 59*(10), 1087-1091. doi: 10.1016/j.jclinepi.2006.01.014

Eagly, A. H., & Chaiken., S. (1998). Attitude structure and function. In D. T. Gilbert, S. T. Fiske & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 269-322). Boston, MA: McGraw-Hill.

Edwards, A. L. (1951). Balanced Latin-square designs in psychological research. *The American Journal of Psychology, 64*(4), 598-603. doi: 10.2307/1418200

Ekman, P. (1992). Are there basic emotions? *Psychological Review, 99*(3), 550-553. doi: 10.1037/0033-295X.99.3.550

Epstein, R. M., Franks, P., Fiscella, K., Shields, C. G., Meldrum, S. C., Kravitz, R. L., & Duberstein, P. R. (2005). Measuring patient-centered communication in Patient–Physician consultations: Theoretical and practical issues. *Social Science & Medicine, 61*(7), 1516-1528.

Fatt, J. P. T. (1998). Nonverbal communication and business success. *Management Research News, 21*(4/5), 1-10.

Fox, A. (2000). *Prosodic features and prosodic structure: the phonology of suprasegmentals*. Oxford: Oxford University Press.

Francis, R. (2013). *Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry*. London: The Stationery Office.

Gandour, J., Tong, Y., Wong, D., Talavage, T., Dzemidzic, M., Xu, Y., . . . Lowe, M. (2004). Hemispheric roles in the perception of speech prosody. *Neuroimage, 23*(1), 344-357.

Gates, G. A., & Mills, J. H. (2005). Presbycusis. *The Lancet, 366*(9491), 1111-1120. doi: 10.1016/s0140-6736(05)67423-5

General Medical Council. (2006). *Good medical practice*. London: GMC.

General Medical Council. (2009). *Tomorrow's Doctors.* London: GMC.

Gobl, C., & N íChasaide, A. (2003). The role of voice quality in communicating emotion, mood and attitude. *Speech Communication, 40*(1-2), 189-212.

Grebelsky-Lichtman, T. (2010). The relationship of verbal and nonverbal behavior to political stature: the political interviews of Israel's prime minister Ariel Sharon. *Journal of Political Marketing, 9*(4), 229-253. doi: 10.1080/15377857.2010.518057

Griffith, I. C. H., Wilson, J. F., Langer, S., & Haist, S. A. (2003). House staff nonverbal communication skills and standardized patient satisfaction. *Journal of General Internal Medicine, 18*(3), 170-174.

Gussenhoven, C. (2002). *Intonation and interpretation: phonetics and phonology*. Paper presented at the Speech Prosody 2002, Aix-en-Provence, France.

Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

Gysels, M., Richardson, A., & Higginson, I. J. (2004). Communication training for health professionals who care for patients with cancer: a systematic review of effectiveness. *Supportive Care in Cancer, 12*(10), 692-700.

Hahn, L. D. (2004). Primary stress and intelligibility: research to motivate the teaching of suprasegmentals. *TESOL Quarterly, 38*(2), 201-223.

Hall, J. A., Harrigan, J. A., & Rosenthal, R. (1995). Nonverbal behavior in clinician-patient interaction. *Applied and Preventive Psychology, 4*(1), 21-37. doi: 10.1016/s0962-1849(05)80049-6

Halliday, M. A. K., & Greaves, W. S. (2008). *Intonation in the grammar of English*. London: Equinox Publishing.

Hargie, O., Dickson, D., Boohan, M., & Hughes, K. (1998). A survey of communication skills training in UK Schools of Medicine: present practices and prospective proposals. *Medical Education, 32*(1), 25-34.

Health and Social Care Information Centre. (2009). *Data on Written Complaints in the NHS 2008-09*. Retrieved from http://www.hscic.gov.uk/catalogue/PUB00566/data-writ-comp-nhs-2008-2009-rep.pdf.

Health and Social Care Information Centre. (2012). *Data on Written Complaints in the NHS 2011-12*. Retrieved from http://www.hscic.gov.uk/catalogue/PUB07197/data-writ-comp-nhs-2011-2012-rep.pdf.

Heather, M. D., Alexandra, L. B., Phyllis, N. B., & Jonathan, J. D. (2007). A one-item question with a Likert or Visual Analog Scale adequately measured current anxiety. *Journal of clinical epidemiology, 60*(4), 356-360.

Hewings, M. (1995). Tone choice in the English intonation of non-native speakers. *International Review of Applied Linguistics in Language Teaching, 33*(3), 251-265.

Hewlett, N., & Beck, J. (2006). *An introduction to the science of phonetics*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Hilton, S., & Southgate, L. (2007). Professionalism in medical education. *Teaching and Teacher Education, 23*(3), 265-279.

Hiraga, Y. (2005). British attitudes towards six varieties of English in the USA and Britain. *World Englishes, 24*(3), 289-308. doi: 10.1111/j.0883-2919.2005.00411.x

Hirst, D., & Di Cristo, A. (1998). A survey of intonation systems. In D. Hirst & A. D. Cristo (Eds.), *Intonation systems: A survey of twenty languages* (pp. 1-44). Cambridge: Cambridge University Press.

Humphrey, C., Hickman, S., & Gulliford, M. C. (2011). Place of medical qualification and outcomes of UK General Medical Council "fitness to practise" process: cohort study. *British Medical Journal, 342*. doi: 10.1136/bmj.d1817

Hunt, M. T., & Glucksman, M. E. (1991). A review of 7 years of complaints in an inner-city accident and emergency department. *Archives of Emergency Medicine*, *8*(1), 17-23. doi: 10.1136/emj.8.1.17

IELTS Guide for Educational Institutions, Governments, Professional Bodies and Commercial Organisations. (2011). Retrieved from http://takeielts.britishcouncil.org/sites/default/files/Guide%20for%20Institutions%20and%20Organisations%202011.pdf.

Institute for Healthcare Communication. (2011). Impact of Communication in Healthcare. Retrieved from http://healthcarecomm.org/about-us/impact-of-communication-in-healthcare/

Jangland, E., Gunningberg, L., & Carlsson, M. (2009). Patients' and relatives' complaints about encounters and communication in health care: evidence for quality improvement. *Patient Education and Counseling, 75*(2), 199-204. doi: 10.1016/j.pec.2008.10.007

Jimenez-Corona, M. E., Ponce-De-Leon-Rosales, S., Rangel-Frausto, S., & Mohar-Betancourt, A. (2006). Epidemiology of medical complaints in Mexico: identifying a general profile. *International Journal for Quality in Health Care*, *18*(3), 220-223. doi: 10.1093/intqhc/mzl004

Jinks, C., Ong, B. N., & Paton, C. (2000). Mobile medics? The mobility of doctors in the European Economic Area. *Health Policy*, *54*(1), 45-64. doi: 10.1016/s0168-8510(00)00097-x

Johnson, K. (2003). *Acoustic and auditory phonetics* (2nd ed.). Malden, Oxford and Carlton: Blackwell.

Johnson, W. F., Emde, R. N., Scherer, K. R., & Klinnert, M. D. (1986). Recognition of emotion from vocal cues. *Archives of General Psychiatry*, *43*, 280-283.

Johnstone, T., & Scherer, K. R. (2000). Vocal communication of emotion. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (pp. 220-235): Guilford Publication.

Kachru, B. B. (1992). World Englishes: approaches, issues and resources. *Language Teaching, 25*(01), 1-14. doi: doi:10.1017/S0261444800006583

Kantowitz, B. H., Roediger III, H. L., & Elmes, D. G. (2009). *Experimental psychology* (9th ed.). Belmont: Wadsworth.

Kirk, L. M. (2007). Professionalism in medicine: Definitions and considerations for teaching. *Baylor University Medical Center Proceedings, 20*(1), 13-16.

Kirk, R. E. (2003). Experimental design. In J. A. Schinka & W. F. Velicer (Eds.), *Handbook of psychology* (Vol. 2, pp. 3-32). New Jersey: Wiley.

Klamen, D., & Williams, R. (2006). Using standardized clinical encounters to assess physician communication. In D. T. Stern (Ed.), *Measuring medical professionalism* (pp. 53-74). New York: Oxford University Press.

Kline, P. (1999). *The handbook of psychological testing* (2nd ed.). London: Routledge.

Krumhansl, C. L. (2002). Music: A link between cognition and emotion. *Current Directions in Psychological Science, 11*(2), 45-50. doi: 10.1111/1467-8721.00165

Kurtz, S., Silverman, J., Benson, J., & Draper, J. (2003). Marrying content and process in clinical method teaching: enhancing the Calgary-Cambridge guides. *Academic Medicine*, *78*(8), 802-809.

Kurtz, S. M. (2002). Doctor-patient communication: principles and practices. *Canadian Journal of Neurological Sciences, 29 Suppl 2,* S23-29.

Kurtz, S. M., Silverman, J., & Draper, J. (2005). *Teaching and learning communication skills in medicine* (2nd ed.). Abingdon: Radcliffe Publishing.

Labov, W. (2001). *Principles of linguistic change* (Vol. 2: Social factors). Oxford: Blackwell.

Ladd, D. R., Silverman, K. E. A., Tolkmitt, F., Bergmann, G., & Scherer, K. R. (1985). Evidence for the independent function of intonation contour type, voice quality, and F0 range in signaling speaker affect. *Journal of the Acoustical Society of America, 78*, 435–444.

Ladefoged, P. (1962). *Elements of acoustic phonetics*. Chicago: University of Chicago Press.

Ladefoged, P. (2006). *A course in phonetics* (5th ed.). Boston, MA: Thomson Wadsworth.

Laver, J. (1994). *Principles of phonetics*. Cambridge: Cambridge University Press.

Levinson, W., Roter, D. L., Mullooly, J. P., Dull, V. T., & Frankel, R. M. (1997). Physician-patient communication: the relationship with malpractice claims among primary care physicians and surgeons. *Journal of the American Medical Association, 277*(7), 553-559.

Lewis, S., & Stenfert-Kroese, B. (2010). An investigation of nursing staff attitudes and emotional reactions towards patients with intellectual disability in a general hospital setting. *Journal of Applied Research in Intellectual Disabilities*, *23*(4), 355-365. doi: 10.1111/j.1468-3148.2009.00542.x

Lieb, K., & Shah, D. (2010). Consumer culture theory, nonverbal communication, and contemporary politics: Considering context and embracing complexity. *Journal of Nonverbal Behavior, 34*(2), 127-136. doi: 10.1007/s10919-010-0085-y

Linville, S. E. (2004). The aging voice. *ASHA Leader*, *9*(19), 12-21.

Maguire, P., Booth, K., Elliott, C., & Jones, B. (1996). Helping health professionals involved in cancer care acquire key interviewing skills—the impact of workshops. *European Journal of Cancer, 32*(9), 1486-1489. doi: 10.1016/0959-8049(96)00059-7

Martin, J., Lloyd, M., & Singh, S. (2002). Professional attitudes: can they be taught and assessed in medical education? *Clinical Medicine, Journal of the Royal College of Physicians, 2*(3), 217-223.

McCormack, H. M., de L. Horne, D. J., & Sheather, S. (1988). Clinical applications of visual analogue scales: a critical review. *Psychological Medicine, 18*(04), 1007-1019. doi: doi:10.1017/S0033291700009934

Mennen, I., Schaeffler, F., & Docherty, G. (2007). *Pitching it differently: a comparison of the pitch ranges of German and English speakers*. Paper presented at the 16th International Congress of Phonetic Sciences, Saarbrücken, Germany.

Millar, M. G., & Tesser, A. (1986). Effects of affective and cognitive focus on the attitude–behavior relation. *Journal of Personality and Social Psychology, 51*(2), 270-276. doi: 10.1037/0022-3514.51.2.270

Milton, R. C. (1964). An extended table of critical values for the Mann-Whitney (Wilcoxon) two-sample statistic. *Journal of the American Statistical Association, 59*(307), 925-934. doi: 10.2307/2283111

Montini, T., Noble, A. A., & Stelfox, H. T. (2008). Content analysis of patient complaints. *International Journal for Quality in Health Care, 20*(6), 412-420. doi: 10.1093/intqhc/mzn041

Mozziconacci, S. (2001). *Emotion and attitude conveyed in speech by means of prosody*. Paper presented at the 2nd Workshop on Attitude, Personality and Emotions in User-Adapted Interaction, Sonthofen, Germany.

Munro, M. J., & Derwing, T. M. (1998). The effects of speaking rate on listener evaluations of native and foreign-accented speech. *Language Learning, 48*(2), 159-182.

Nolan, F. (2003). *Intonational equivalence: an experimental evaluation of pitch scales*. Paper presented at the Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona.

Nolan, F. (2006). Intonation. In B. Aarts & A. McMahon (Eds.), *The handbook of English linguistics* (pp. 433-460). Malden, MA: Blackwell Publishing.

Nooteboom, S. (1997). The prosody of speech: melody and rhythm. In W. J. Hardcastle & J. Laver (Eds.), *The handbook of phonetic sciences* (pp. 640-673). Oxford: Blackwell.

O'Toole, G. (2008). *Communication: Core interpersonal skills for health professionals*. Sydney: Elsevier Churchill Livingstone.

Ong, L. M. L., de Haes, J. C. J. M., Hoos, A. M., & Lammes, F. B. (1995). Doctor-patient communication: a review of the literature. *Social Science and Medicine, 40*(7), 903-918.

Parry, R. (2008). Are interventions to enhance communication performance in allied health professionals effective, and how should they be delivered? Direct and indirect evidence. *Patient Education and Counseling, 73*(2), 186-195.

Pfizinger, H. R., & Kaernbach, C. (2008). *Amplitude and amplitude variation of emotional speech*. Paper presented at the Interspeech 2008, Brisbane, Australia.

Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation.* (Doctorate Dissertation), MIT, Cambridge, MA.

Radford, M. L. (1998). Approach or avoidance? The role of nonverbal communication in the academic library user's decision to initiate a reference encounter *Library Trends, 46*(4), 699.

Rees, C. E., & Garrud, P. (2001). Identifying undergraduate medical students' attitudes towards communication skills learning: a pilot study. *Medical Teacher, 23*(4), 400-406. doi: 10.1080/01421590120057067

Roberts, C., Moss, B., Wass, V., Sarangi, S., & Jones, R. (2005). Misunderstandings: a qualitative study of primary care consultations in multilingual settings, and educational implications. *Medical Education, 39*(5), 465-475. doi: 10.1111/j.1365-2929.2005.02121.x

Roberts, C., Wass, V., Jones, R., Sarangi, S., & Gillett, A. (2003). A discourse analysis study of 'good' and 'poor' communication in an OSCE: a proposed new framework for teaching students. *Medical Education, 37*(3), 192-201.

Roberts, K. J. (2002). Physician-patient relationships, patient satisfaction, and antiretroviral medication adherence among HIV-infected adults attending a public health clinic. *AIDS Patient Care & STDs, 16*(1), 43-50.

Rosenberg, M. J., & Hovland, C. I. (1960). Cognitive, affective and behavioral components of attitudes. In M. J. Rosenberg & C. I. Hovland (Eds.), *Attitude organization and change: An analysis of consistency among attitude components* (pp. 1-14). New Haven: Yale University Press.

Royal College of Physicians. (2005). Doctors in society: medical professionalism in a changing world. Report of a Working Party of the Royal College of Physicians of London. London: RCP.

Ruiz-Moral, R., Perez Rodriguez, E., Perula de Torres, L. A., & de la Torre, J. (2006). Physician-patient communication: a study on the observed behaviours of specialty physicians and the ways their patients perceive them. *Patient Education & Counseling, 64*(1-3), 242-248.

Sato, K., Hirano, M., & Nakashima, T. (2002). Age-related changes of collagenous fibers in the human vocal fold mucosa. *Annals of Otology, Rhinology & Laryngology, 111*(1), 15-20.

Scherer, K. R. (1982). Methods of research on vocal communication: Paradigms and parameters. In K. R. Scherer & P. Ekman (Eds.), *Handbook of methods in nonverbal behavior research* (pp. 136-198). Cambridge: Cambridge University Press.

Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication, 40*, 227–256.

Scherer, K. R., Johnstone, T., & Klasmeyer, G. (2003). Vocal expression of emotion. In R. J. Davidson, K. R. Scherer & H. Goldsmith (Eds.), *Handbook of the affective sciences* (pp. 433-456). New York: Oxford University Press.

Scherer, K. R., Ladd, R. D., & Silverman, K. E. A. (1984). Vocal cues to speaker affect: testing two models. *Journal of the Acoustical Society of America, 76*(5), 1346-1356.

Schröder, M. (2001). *Emotional speech synthesis: A review.* Paper presented at the Seventh European Conference on Speech Communication and Technology.

Schröder, M. (2009). expressive speech synthesis: past, present, and possible futures. In J. Tao & T. Tan (Eds.), *Affective information processing* (pp. 111-126). London: Springer.

Seddoh, S. A. (2002). How discrete or independent are “affective prosody” and “linguistic prosody?”. *Aphasiology, 16*(7), 683-692.

Silverman, J., Kurtz, S. M., & Draper, J. (2005). *Skills for communicating with patients* (2nd ed.). Abingdon: Radcliffe Publishing.

Sliwa, J., Makoul, G., & Betts, H. (2002). Rehabilitation-specific communication skills training: improving the physician-patient relationship. *American Journal of Physical Medicine and Rehabilitation, 81*(2), 126-132.

Smith, R. (2010, 15 July). Foreign doctors responsible for overdoses: NHS watchdog, *The Daily Telegraph*.

Spackman, M. P., Brown, B. L., & Otto, S. (2009). Do emotions have distinct vocal profiles? A study of idiographic patterns of expression. *Cognition and Emotion, 23*(8), 1565-1588.

Stern, D. T. (2006). *Measuring medical professionalism*: Oxford University Press, USA.

Stewart, M., Brown, J. B., Weston, W. W., McWhinnery, I. R., McWilliam, C. L., & Freeman, T. R. (2003). *Patient-centered medicine: transforming the clinical method*. Abingdon: Radcliffe Medical Press.

Stewart, M. A. (1995). Effective physician-patient communication and health outcomes: a review. *Canadian Medical Association Journal, 152*(9), 1423-1433.

Taylor, D. M., Wolfe, R. S., & Cameron, P. A. (2004). Analysis of complaints lodged by patients attending Victorian hospitals, 1997–2001. *The Medical Journal of Australia, 181*(1), 31-35.

Thompson, W. F., & Balkwill, L. L. (2006). Decoding speech prosody in five languages. *Semiotica, 158*(1-4), 407-424.

Trumble, S. C., O'Brien, M. L., O'Brien, M., & Hartwig, B. (2006). Communication skills training for doctors increases patient satisfaction. *Clinical Governance, 11*(4), 299-307.

Van De Poel, K., & Brunfaut, T. (2010). Medical communication in L1 and L2 contexts: Comparative modification analysis. *Intercultural Pragmatics, 7*(1), 103-129. doi: 10.1515/iprg.2010.005

van den Berg, H., Manstead, A. S. R., van der Pligt, J., & Wigboldus, D. H. J. (2006). The impact of affective and cognitive focus on attitude formation. *Journal of Experimental Social Psychology, 42*(3), 373-379.

van Laerhoven, H., van der Zaag-Loonen, H. J., & Derkx, B. H. F. (2004). A comparison of Likert scale and visual analogue scales as response options in children's questionnaires. *Acta Paediatrica, 93*, 830-835.

Verdugo, D. R. (2005). The nature and patterning of native and non-native intonation in the expression of certainty and uncertainty: Pragmatic effects. *Journal of Pragmatics, 37*(12), 2086-2115.

von Fragstein, M., Silverman, J., Cushing, A., Quilligan, S., Salisbury, H., & Wiskin, C. (2008). UK consensus statement on the content of communication curricula in undergraduate medical education. *Medical Education, 42*(11), 1100-1107. doi: 10.1111/j.1365-2923.2008.03137.x

Watson, P. J., & Hughes, D. (2006). The relationship of vocal loudness manipulation to prosodic F0 and durational variables in healthy adults. *Journal of Speech, Language, and Hearing Research, 49*(3), 636-644.

Wennerstrom, A. (1994). Intonational meaning in English discourse: a study of non-native speakers. *Applied Linguistics, 15*(4), 399-420.

Wichmann, A. (2000). *The attitudinal effects of prosody, and how they relate to emotion*. Paper presented at the ISCA Tutorial and Research Workshop (ITRW) on Speech and Emotion, Newcastle, Northern Ireland, UK.

Wightman, C. W. (2002). *ToBI or not ToBI*. Paper presented at the Speech Prosody 2002, Aix-en-Provence, France.

Wilkinson, S., Bailey, K., Aldridge, J., & Roberts, A. (1999). A longitudinal evaluation of a communication skills programme. *Palliative Medicine, 13*(4), 341-348.

Winkelmann, R. (2004). Co-payments for prescription drugs and the demand for doctor visits – Evidence from a natural experiment. *Health Economics, 13*(11), 1081-1089. doi: 10.1002/hec.868

Wofford, M. M., Wofford, J. L., Bothra, J., Kendrick, S. B., Smith, A., & Lichstein, P. R. (2004). Patient complaints about physician behaviors: A qualitative study. *Academic Medicine, 79*(2), 134-138.

Woloschuk, W., Harasym, P. H., & Temple, W. (2004). Attitude change during medical school: a cohort study. *Medical Education, 38*(5), 522-534. doi: 10.1046/j.1365-2929.2004.01820.x

Wong, S. Y., & Lee, A. (2003). Communication skills and doctor patient relationship. *The Hong Kong Medical Diary, 11*(3), 7-9.

Yiu, E. M. L., & Ng, C. Y. (2004). Equal appearing interval and visual analogue scaling of perceptual roughness and breathiness. *Clinical Linguistics & Phonetics, 18*(3), 211-229. doi: doi:10.1080/0269920042000193599

Yuan, J., Liberman, M., & Cieri, C. (2006). *Towards an integrated understanding of speaking rate in conversation*. Paper presented at the Ninth International Conference on Spoken Language Processing, Pittsburgh, PA, USA.

Yuan, J., Shih, C., & Kochanski, G. (2002). *Comparison of declarative and interrogative intonation in Chinese*. Paper presented at the Speech Prosody 2002, Aix-en-Provence, France.

Zhang, Y., Nissen, S. L., & Francis, A. L. (2008). Acoustic characteristics of English lexical stress produced by native Mandarin speakers. *Journal of the Acoustical Society of America*, *123*(6), 4498-4513.

Glossary

Acoustic phonetics: the study of the physical properties of speech sounds with instrumental aid.

Articulation: the movement of speech organs to make speech sounds. This broad meaning of the term is adopted in this thesis. However, the use of the terms may strictly refer to the contribution of the "organs along the vocal tract from the larynx to the lips" (Laver 1994:132), as opposed to phonation referring to the generation of acoustic energy (e.g. by voicing) and initiation referring to the airstream mechanism (Laver 1994).

Bark: a psycho-acoustic pitch scale corresponding to the critical bands of hearing.

Condenser microphone: a type of microphone that can be small yet sensitive. Compared with the often cheaper and larger dynamic microphone, it is used when high recording quality is required and is considered to be suitable for recording voices. Modern lapel microphones usually belong to this type.

ERB (scale): a near-logarithmic psycho-acoustic scale for measuring pitch perception, similar to the Bark scale. It was developed based on calculating the equivalent rectangular bandwidth (ERB).

F0: fundamental frequency, the lowest frequency of a periodic waveform (e.g. that of voicing). The F0 of speech is the rate at which the vocal folds vibrate. It is the physical correlate of pitch.

Mel: a psycho-acoustic scale to measure perceived pitch which was developed based on experiments.

Intensity: a physical property of sound related to the magnitude of acoustic energy. It is the physical correlate of loudness.

PCM: pulse-code modulation, a method of digitally representing sampled analogue sound signals in recording, which is a standard form used in computers and digital audio equipment.

Phonetics: the scientific study of speech sounds.

Rate of speech: number of speech units (e.g. syllables) per time unit (e.g. per second), which can be calculated with either silent pauses included (speaking rate) or these excluded (articulation rate).

Sampling rate: the number of samples per second used to encode an analogue sound signal in a digital form. The sampling rate of 48 KHz is usually used by professional digital equipment. It may also be referred to as sampling frequency.

Semitone: the smallest musical interval in Western tonal music. A semitone is one twelfth of a musical octave. The semitone scale is a logarithmic transformation of the

Hz scale, and can be used as a psycho-acoustic scale for measuring pitch perception in speech.

Solid state recorder: a type of digital audio recorder that does not use moving mechanical parts such as tapes.

Spectrogram: a three-dimensional display of the frequency and magnitude of the acoustic energy in speech over time.

Waveform: a visual representation of the magnitude of sound pressure fluctuations over time.

Appendices

Appendix A. Ethical approval for the acoustic study

Faculty of Health Research Ethics Committee

University of East Anglia

Faculty of Health Research Office
Elizabeth Fry Building, Room 2.30
University of East Anglia
Norwich NR4 7TJ

Email: margaret.rhodes@uea.ac.uk
Direct Dial: +44 (0) 1603 59 7190
Research: +44 (0) 1603 59 1720
Fax: +44 (0) 1603 59 1132

Web: http://www.uea.ac.uk

Fan Wang
Room 0.27
The Queen's Building
University of East Anglia
Norwich NR4 7TJ

5th July 2010

Dear Fan

The Influence of Non-native Language Prosody on Health Communication: Reference 2009/10-062

The submission of your research proposal was discussed at the Ethics Committee meeting on Thursday 24th June 2010.

The Committee were happy to approve your application in principle but have the following concerns which they would like you to address and amend accordingly:

(1) Participant information sheet

- Page 1, paragraph 1 – delete '*do not carry on recording*'
- Paragraph 1 – '*record speech*' – rather than '*collect speech data*'
- Allow more than 10 minutes for the role play
- Destroying data – refer to the Data Protection Act (data is normally kept for 5 years)

(2) Consent form

- Ask the participants to initial the boxes (rather than ticking them)
- The consent form mentions lay listeners. Who will they be? The protocol only mentions actors

(3) Provide professional support for participants in case any of the role play scenarios cause distress

(4) Find a scenario from year 1 that everyone goes through

Please write to me once you have resolved/clarified the above issues. I require documentation confirming that you have complied with the Committee's suggestions. The Committee have requested that you detail the changes below the relevant point on the text in this letter and also include your amendments as a tracked change within your application/proposal. The revisions to your application can be considered by Chair's action rather than go to a committee meeting, which means that the above documentation can be resubmitted at any time. Please could you send your revisions to me as an attachment in an email as this will speed up the decision making process.

As your project does not have ethics approval until the above issues have been resolved, I want to remind you that you should not be undertaking your research project until you have ethical approval by the FoH Ethics Committee. Planning on the project or literature based elements can still take place but not the research involving the above ethical issues. This is to ensure that you and your research are insured by the University and that your research is undertaken within the University's 'Guidelines on Good Practice in Research' approved by the Senate in July 2004.

Yours sincerely

Maggie Rhodes
Research Administrator

Faculty of Health Research Ethics Committee

University of East Anglia

Fan Wang
Room 0.27
The Queen's Building
University of East Anglia
Norwich
NR4 7TJ

Faculty of Health Research Office
Elizabeth Fry Building, Room 2.30
University of East Anglia
Norwich NR4 7TJ

Email: margaret.rhodes@uea.ac.uk
Direct Dial: +44 (0) 1603 59 7190
Research: +44 (0) 1603 59 1720
Fax: +44 (0) 1603 59 1132

Web: http://www.uea.ac.uk

22nd July 2010

Dear Fan

The influence of Non-native Language Prosody on Health Communication: Reference 2009/10-062

The amendments to your above proposal have now been considered by the Chair of the FOH Ethics Committee and we can now confirm that your proposal has been approved.

Please could you ensure that any amendments to either the protocol or documents submitted are notified to us in advance and also that any adverse events which occur during your project are reported to the committee. Please could you also arrange to send us a report once your project is completed.

The committee would like to wish you good luck with your project.

Yours sincerely

Maggie Rhodes
Research Administrator

Faculty of Medicine and Health Sciences Research Ethics Committee

University of East Anglia

Fan Wang
Postgraduate Researcher (AHP)
Room 0.27 The Queen's Building
University of East Anglia
Norwich
Norfolk NR4 7TJ

Research & Enterprise Services
West Office (Science Building)
University of East Anglia
Norwich Research Park
Norwich, NR4 7TJ

Telephone: +44 (0) 1603 591566
Email: fmh.ethics@uea.ac.uk

Web: www.uea.ac.uk/foh/research/ethics-committee

11th November 2011

Dear Fan

Project title: The Influence of Non-native Language Prosody on Health Communication.
Reference 2009-10/62

Thank you for your e-mail dated 25th October 2011 notifying us of the amendments you would like to make to your above proposal. These have been considered by the Chair of the Faculty Research Ethics Committee and we can now confirm that your amendments have been approved.

Please can you ensure that any further amendments to either the protocol or documents submitted are notified to us in advance, and also that any adverse events which occur during your project are reported to the Committee.

Please can you also arrange to send us a report once your project is completed.

Yours sincerely

Yvonne Kirkham

Yvonne Kirkham
Project Officer

Appendix B. Gatekeeper's consent

RE: Your permission

Leinster Samuel Prof (MED)
Sent:23 July 2010 13:58
To: Wang Fan Mr (AHP)

Fan,

The email is sufficient. Full permission is granted for you to undertake the study.

You should keep a copy of this email for future reference.

Sam

-----Original Message-----
From: Wang Fan Mr (AHP) <Fan.Wang@uea.ac.uk>
Sent: 23 July 2010 11:59
To: Leinster Samuel Prof (MED) <S.Leinster@uea.ac.uk>
Cc: Raschka Christine Dr (AHP) <C.Raschka@uea.ac.uk>
Subject: Your permission

Dear Sam

Please see the attached letter asking for your permission.
Will that be okay, or do you need a signed hard copy?
I will also send you the amended documents used for obtaining ethical approval in a separate email.

Many thanks

Fan

Appendix C. Participant information sheet for student participants

University of East Anglia

PARTICIPANT INFORMATION SHEET

The Influence of Non-native Language Prosody on Health Communication

Dear MB/BS student

I am a PhD student based in the Education in Health Research Institute within the Faculty of Health. I am looking to record speech from medical students for my project. This document explains some details about the project and how you can help. If you are not interested, thank you for reading so far, you can simply disregard the document.

Before you decide whether or not to agree to take part, it is important that you understand clearly the elements of this study that are relevant to you. If you are still interested, please take time to read the following information carefully. If you have any questions, please do not hesitate to contact me fan.wang@uea.ac.uk.

What is the purpose of this study?
Despite communication skills training in medical schools, the number of patient complaints related to communication remains high. Given the linguistic diversity of medical professionals in the UK and evidence suggesting that languages differ in prosody (melody and rhythm of speech), we want to investigate:

- whether native and non-native medical students differ in English prosody, and
- whether the differences are important for listeners in medical encounters.

Why have I been invited?
We are looking to recruit MB/BS students to take part in this study and we need speech data from both native and non-native speakers of English to answer the research questions.

Do I have to take part?
No, it is entirely up to you to decide whether to volunteer. If you do volunteer, you may or may not be selected for further study. If you are selected, you will be asked to give consent to provide audio recordings of the tasks you perform. Even if you volunteer to take part, you can withdraw at any time without giving us a reason, and the data you provide will be immediately destroyed. If you decide not to volunteer, this will have absolutely no effect on your position within your department – to volunteer or not to volunteer is entirely up to you.

Who gets chosen?
If you volunteer to take part, we will send you a short questionnaire asking basic questions such as your language background. We will need to control for certain factors in this study, so the decision to invite you to participate will be based on your information. If you were to volunteer but not be chosen this would be no reflection of your ability to take part; it would be simply because our inclusion criteria for this particular study did not apply to you. You may well meet the criteria for future extended studies. The following chart explains how you can help us, from now until we finish collecting your data.

Researcher to explain and answer questions during a consultation skills training session – leave contact information if you want to volunteer

Short questionnaire sent to you – fill in basic information and return to researcher

Notified that you have been chosen for the study – arrange a time for recording session with researcher

Recording day at a mutually convenient time – sign consent form and perform two speech tasks that will be recorded (thank you for your help!)

What will I have to do?
You will be asked to carry out two tasks: reading out a short list of sentences and role-playing in a practice consultation scenario with a simulated patient, i.e. an actor. We need you to perform the role-play in the same way as you have been taught in your consultations skills sessions. Both tasks will be audio recorded on campus at a mutually convenient time. These tasks will take about 30 minutes all together.

Will my personal information be kept confidential?
Yes, all your personal information will be kept strictly confidential in accordance with the *Data Protection Act 1998*. All data you provide will be anonymised and you will be referred to by your participant number throughout the project. Your data will be stored in locked cabinet and on password protected UEA server space.

What will happen to my data?
The speech samples you provided will be analysed by the researcher using speech analysis software. These will be computer measurements of features of

your speech sounds. No ‘goodness’ or ‘correctness’ judgements will be made on the content of your speech, we will simply be analysing what you do. A selection of your audio recordings will also be presented anonymously to a group of listeners who are not trained in phonetic science. The actor (simulated patient in the consultation scenario) and the listeners will be asked to comment on the features of your voice and their perceptions of your speech.

Your data will only be used for future research purposes if you allow us to do so in writing. If you do not agree to us using your data for future research, they will be archived for five years after the end of the study and then destroyed. If you allow us to use your data for future research, your data will be kept securely with the same level of confidentiality for five years after the end of the study and will be destroyed after that.

What will happen to the results?
The results of this study will be included in the researcher’s PhD thesis and publications in academic journals. Your anonymity will be protected in all potential academic publications.

Who is funding / supervising the research?
This research is funded by the UEA Faculty of Health Studentship and the UEA International Scholarship, both awarded to the postgraduate researcher. A team consisting of Dr Christine Raschka (AHP), Dr Zoe Butterfint (AHP) and Prof Sam Leinster (MED) is supervising the research.

Who has reviewed the study?
The Faculty of Health Ethics Committee at UEA has reviewed and approved this study.

Are there any risks to me?
No, this study is completely non-invasive and will pose no risk whatsoever to you. However, in the very unlikely case that you experience distress from the role-play, you should see your personal advisor or the Dean of Students’ Office in the first instance.

What should I do if I have any questions or concerns?
If you have any questions or concerns about the study, please feel free to contact the researcher.

Fan Wang
Postgraduate researcher
Education in Health Research Institute
Room 0.27, The Queen’s Building, UEA, Norwich, Norfolk NR4 7TJ
fan.wang@uea.ac.uk
01603 593300

Appendix D. Participant information sheet for the simulated patient

University of East Anglia

PARTICIPANT INFORMATION SHEET

The Influence of Non-native Language Prosody on Health Communication

Dear Simpatico Actor

I am a PhD student based in the Education in Health Research Institute within the Faculty of Health and am looking to collect data on medical students' communication during simulated consultations for my project. In our previous discussion, you expressed an interest in being an employed role-player for the project as well as participating in the study by providing data. This document explains some details about the study and how you can help as a participant.

Before you decide whether or not to agree to take part, it is important that you understand clearly the elements of this study that are relevant to you. If you are still interested, please take time to read the following information carefully. If you have any questions, please do not hesitate to contact me fan.wang@uea.ac.uk.

What is the purpose of this study?

Since there is research evidence showing that prosodic features (features of the speech signal linked to pitch, volume and rate) are vital in conveying speaker attitudes and emotions, and that medical communication is often emotionally laden, we want to investigate whether prosodic features in English speech (by both native and non-native speakers) are important for the listener's experience of medical communication. We are particularly interested in the role-player's perception of the medical students' communication during simulated consultations.

Why have I been invited?

We need data on the simulated patient's experience of the medical students' communication. Since Simpatico actors like yourself have helped with consultation skills teaching at UEA and are experienced in giving feedback to students, we need you to evaluate your impression as a simulated patient of the student doctor's communication with you.

Do I have to take part?

No, it is entirely up to you to decide. If you agree to take part in this study, we will ask you to complete a short questionnaire of your perception of the students' communication. This will take no longer than 5 minutes and will take place immediately after the role play. Even if you agree to take part, you can withdraw from this at any time without giving us a reason. We will ask you to give explicit consent prior to your participation.

What will I have to do?
After the end of the role-play, we will ask you to go to a different room to complete the questionnaire of your impression of the student's communication during the consultation. This will be in the form of an easy-to-use evaluation scale consisting of 12 items in total and it should take no longer than 5 minutes to fill in.

Will you keep any personal information?
No, we will not collect any personal information about you. The responses in the evaluation scales will be anonymous.

What will happen to my data?
Your evaluation of the medical students' communication skills will be kept confidential. Moreover, if any student wishes to withdraw from the study, all their information, including data on your impression of their communication, will be destroyed without us informing you. Your evaluation data will be stored in locked filing cabinets and on password protected UEA server space, and will be viewed by the researcher and the supervisors only. We will use statistical software to analyse your impressions of different students' communication.

You will be asked if we can use your evaluation data for future research purposes. If you do not agree to us using your data for future research, they will be archived for five years after the end of the study and then destroyed. If you allow us to use your data for future research, your data will be kept securely with the same level of confidentiality for five years after the end of the study and will be destroyed after that.

What will happen to the results?
The results of this study will be included in the researcher's PhD thesis and publications in academic journals. Your anonymity will be protected in all potential academic publications.

Who is funding / supervising the research?
This research is funded by the UEA Faculty of Health Studentship and the UEA International Scholarship, both awarded to the postgraduate researcher. A team consisting of Dr Christine Raschka (AHP), Dr Zoe Butterfint (AHP) and Prof Sam Leinster (MED) is supervising the research.

Who has reviewed the study?
The Faculty of Health Ethics Committee at UEA has reviewed and approved this study.

Are there any risks to me?
No, this study is completely non-invasive and will pose no risk whatsoever to you. However, in the very unlikely case that you experience distress from the role-play, medical including counselling help will be available to you, usually by arrangement with your GP.

What should I do if I have any questions or concerns?
If you have any questions or concerns about the study, please feel free to contact the researcher.

Appendix E. Screening questionnaire for the acoustic study

Participant reference number __________ (to be completed by researcher)

Faculty of Health
Education in Health Research Institute

University of East Anglia

Participant Recruitment Questionnaire

The Influence of Non-native Language Prosody on Health Communication

Dear MB/BS student

You expressed an interest in taking part in the study when I came to talk to you in the consultation skills training session. You also had an opportunity to read the participant information sheet. Before we can select participants for the study, we need to collect some information. If you still want to volunteer for the study, please take a few minutes to fill in this questionnaire and return it to me by email. The questionnaire should be easy to fill in. You can use TAB key to move to the next box, or click to choose a box and type in your answer. Please leave any questions that do not apply to you blank.

1. Name:

2. Contact e-mail:

3. Age:
(Please state in years. We will only use this information for statistical purposes)

4. Can we contact you on your mobile phone? Yes ☐ No ☐
If yes, please provide your number

5. Are you a native speaker of English? Yes ☐ No ☐

6. If you are a native speaker of English, would you say you have a Standard Southern British English (SSBE) accent? This is a regionally neutral accent often found in newsreaders or TV presenters, e.g. Philip Schofield. It is the type of accent that would not allow people to guess where you came from (unlike somebody like Paul O'Grady).
Yes ☐ No ☐

7. If you would not describe your accent as SSBE, how would you describe it?

8. If your first language is not English, how long have you been learning English (in years)?

9. If your first language is not English, in which country did you learn English?

10. If your first language is not English, what is your first language?

11. Would you consider yourself to have:

a hearing difficulty Yes ☐ No ☐

a speech difficulty Yes ☐ No ☐

12. Please indicate when you would be most available to participate

October 2010 ☐ November 2010 ☐ December 2010 ☐

Thank you! Please return this as an email attachment to the researcher at fan.wang@uea.ac.uk

For more information, please contact:
Fan Wang, postgraduate researcher
Room 0.27, The Queen's Building
01603 593300 fan.wang@uea.ac.uk

Appendix F. Consent form for students participants

Participant Reference No. [ ]

Faculty of Health
Education in Health Research Institute

University of East Anglia

CONSENT FORM

Project title: *The Influence of Non-Native Language Prosody on Health Communication*
Name of Researcher: Fan Wang

Please initial the boxes as appropriate

1. I confirm that I have read and understood the participant information sheet (version 2.2) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. [ ]

2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason. I understand that should I wish to withdraw from the study, all data I have provided will be destroyed. [ ]

3. I agree to provide, as outlined in the information sheet, data including audio recorded speech samples for this study and, where relevant, that the results of analysis of my data including audio recorded speech samples may be included anonymously in academic publications. [ ]

4. I agree to my speech samples being presented anonymously to a group of lay listeners in the AHP SLT Communication Laboratory, as outlined in the information sheet. I understand that the other data that I provide during the study will only be viewed by the research team, that is, the researcher and the supervisors and will not be viewed by anybody else. [ ]

5. Please select **one** of the following as appropriate

   I **do not** wish the data that I provide to be used for future research purposes after completion of this study. [ ]
   I **agree** to using the data I provide during the study for future research purposes after completion of this study, for a period of up to 5 years. [ ]

6. I agree to take part in this study. [ ]

Name of Participant	Date	Signature
Name of Researcher	Date	Signature

Appendix G. Consent form for the simulated patient

Participant Reference No. ☐

Faculty of Health
Education in Health Research Institute

University of East Anglia

CONSENT FORM

Project title: *The Influence of Non-Native Language Prosody on Health Communication*
Name of Researcher: Fan Wang

Please initial the boxes as appropriate

1. I confirm that I have read and understood the participant information sheet (version 2.2) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason. I understand that should I wish to withdraw from the study, all data I have provided will be destroyed. ☐

3. I agree to provide, as outlined in the information sheet, data including audio recorded speech samples for this study and, where relevant, that the results of analysis of my data including audio recorded speech samples may be included anonymously in academic publications. ☐

4. I agree to my speech samples being presented anonymously to a group of lay listeners in the AHP SLT Communication Laboratory, as outlined in the information sheet. I understand that the other data that I provide during the study will only be viewed by the research team, that is, the researcher and the supervisors and will not be viewed by anybody else. ☐

5. Please select **one** of the following as appropriate

   I **do not** wish the data that I provide to be used for future research purposes after completion of this study. ☐
   I **agree** to using the data I provide during the study for future research purposes after completion of this study, for a period of up to 5 years. ☐

6. I agree to take part in this study. ☐

Name of Participant	Date	Signature
Name of Researcher	Date	Signature

Appendix H. Task materials for student participants

Task Information for Student Participant

Thank you very much for taking part in the study. As you know from the Information Sheet, you will perform two tasks for the study:

(1) a role-play with you as the medical student on placement and an actor as the patient, and
(2) reading aloud a short list of sentences.

You will start with Task 1 (role-play). Please carry out this task in the same way as you have been taught in the consultation skills training sessions. Please take some time to read the following information and to prepare yourself for the task.

Task 1 will take about 10 minutes to complete, although you may spend more time if you need to.

Task 2 (sentence reading) will then commence. This task should take no longer than 3 minutes to complete.

Task 1: Role Play

<u>Patient Consultation: Hypertension Treatment and Management*</u>

You are a medical student on your primary care placement, and you have been asked by your GP tutor to see Mr Jones, a 48-year-old man who was diagnosed with hypertension three months ago by his GP and given medication. He moved soon after and therefore cannot return to his original GP. He has come to this practice for the first time today.

Mr Jones has brought a note of the medication which says ‘Ramipril 1.25 mg one tablet once a day’. His blood pressure today is 162/110.

Your GP tutor has asked Mr Jones if he would mind talking to a medical student and he agreed to help.

The GP has asked you to

- Take the patient’s medical history
- Find out what the patient knows about high blood pressure and the prescribed medication
- Explain the prescribed medication and the risk factors associated with high blood pressure to the patient
- Discuss treatment and management of the condition

* Adapted from: Helen Ward and Julian Barratt (2009) *Passing Your Advanced Nursing OSCE*. Oxford: Radcliffe.

Task 2: Sentence Reading

Please read out the following sentences according to the instructions. For example, if you are asked to say

> *I can see that you're angry*

using an empathetic expression, you might find it helpful to imagine the researcher to be your patient. Then say the sentence to show that you recognise his feelings.

There are not necessarily good or bad ways of saying these sentences – we would like to hear them said in your way. If you hesitate or make a mistake, you can simply say the sentence again.

Here are the sentences:

1. (As a casual social formula)
   How are you?

2. (Accusingly)
   How much alcohol do you drink in a week?

3. (In a patronising way)
   I'm sorry to hear that.

4. (In a demanding way)
   Do you have any thoughts on this?

5. (As a command)
   Tell me more about that.

6. (Demonstrating respect/ attending to patient's comfort)
   How are you?

7. (In a non-judgmental way)
   How much alcohol do you drink in a week?

8. (Empathising with the patient)
   I'm sorry to hear that.

9. (Checking the patient's idea/ concern/ expectation)
   Do you have any thoughts on this?

10. (Asking open-ended question/ expecting divergent response)
    Tell me more about that.

This is the end of your task.
Thank you very much for your help.

Appendix I. Task materials for the simulated patient

Role-play – Information for actor*

Name: Dan Jones
Gender: male
Age: 48
Setting: GP surgery

Clinical details
Shortly before you moved house 3 months ago, you were diagnosed with high blood pressure by your original GP. He gave you some medication which he said you must take. You took it for a month until the pills ran out. You have not taken any in the last 2 months. When you took them you did not notice any side effects. You have not experienced anything special since you stopped taking the pills.

You have come to the surgery today as your neighbour said that you have to take blood pressure medication for life and you are now worried. You have brought a note of what your original GP prescribed; this says Ramipril 1.25 mg one tablet once a day. Your blood pressure today is 162/110.

You have not exercised since teenage years. You are not overweight.

Past medical history
High blood pressure 3 months ago. Nothing of significance apart from that. Your blood test results were normal when you last had them taken at your original GP, but you are not sure what they were for.

Medication
None

Smoking history
You stopped smoking 10 years ago – you used to smoke five to ten cigarettes a day for 20 years.

Diet
You think that you eat fairly healthily, although you like a lot of salt on your food.

Drinking
You drink two to three units of wine at the weekends only.

Family history
Your father had high blood pressure and died of a stroke (aged 72). Your brother also has high blood pressure (currently aged 52).

Social history
You are married with two adult children who have left home. You work in a supermarket near your new home. You quite enjoy it and do not find it stressful.

Temperament

You are friendly and enjoy travelling with your wife. You and your wife both like meeting new people and experiencing new cultures on the journey.

Ideas
You have not experienced anything unusual since your medication ran out 2 months ago, and you had assumed stopping taking the drug was not a big issue until your neighbour told you that blood pressure medication must be taken for life. You became worried since then.

Concerns
Although you did not experience any symptoms, you want to have your blood pressure under control. Since you have a family history of high blood pressure, you fear that what happened to your father will happen to you if the condition is not treated, but you are not very keen on taking drugs. You think you might miss your medication, especially when travelling on holiday, even if it is just one tablet once a day.

Expectations
You expect the doctor to tell you measures to take your blood pressure under control other than taking medication. However, you want to know how severe your condition is and if there is risk of stroke like the one that killed your father. You hope that your condition will be manageable without affecting the quality of your life.

Feelings
You feel a little nervous since you did not follow your original GP's instructions. Moving house was one of the reasons you stopped taking the medicine, but you do not want the new doctor to have the impression that you are a "difficult" patient.

The aim of this scenario is for the medical student to use their communication skills acquired in the training sessions to initiate the patient-centred consultation, to gather information and to give information for this general task. The student needs to establish the initial rapport at the beginning of the consultation and to develop the rapport during the process.

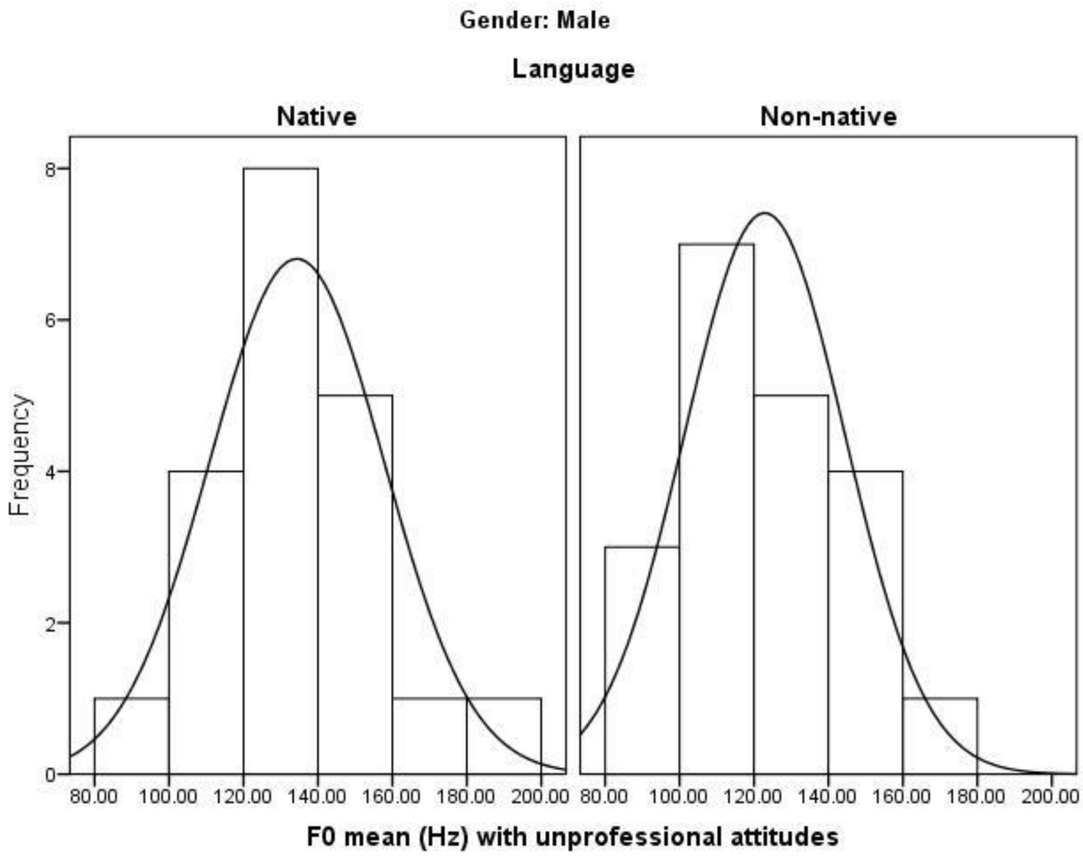
This is the end of the role-play information.

* Adapted from: Helen Ward and Julian Barratt (2009) *Passing Your Advanced Nursing OSCE*. Oxford: Radcliffe.

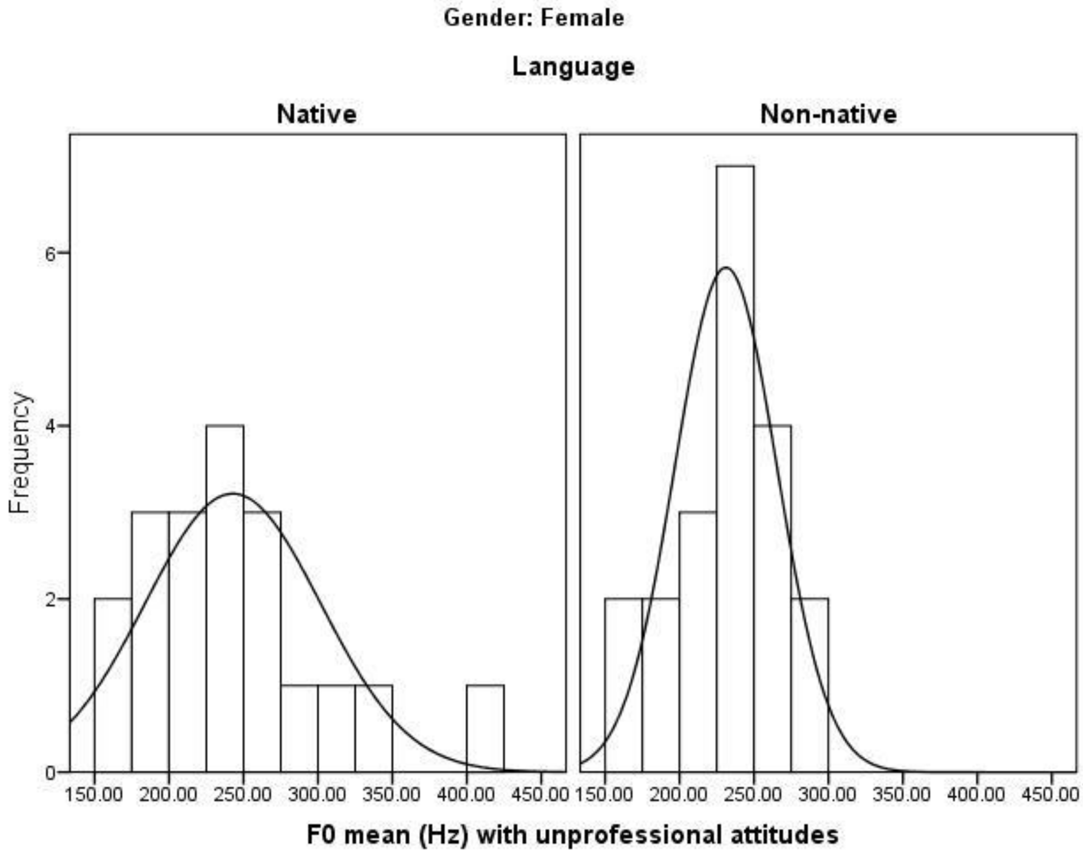
Appendix J. Demographic details of the acoustic study participants

Participant ID	First language	Gender	Year of medicine course	Age (years)
Native male (*n* = 4), ***M*** (*SD*) age = **23.50** (4.43)				
NS1-01	English	Male	5	30
NS1-03	English	Male	5	22
NS1-02	English	Male	3	20
NS1-04	English	Male	2	22
Non-native male (*n* = 4), ***M*** (*SD*) age = **29.50** (5.98)				
NN1-01	Cantonese	Male	5	24
NN1-04	Cantonese	Male	5	28
NN1-02	Soga	Male	3	28
NN1-03	Portuguese	Male	2	38
Native female (*n* = 4), ***M*** (*SD*) age = **23.75** (3.50)				
NS2-02	English	Female	5	25
NS2-04	English	Female	5	28
NS2-01	English	Female	3	22
NS2-03	English	Female	3	20
Non-native female (*n* = 4), ***M*** (*SD*) age = **26.50** (7.33)				
NN2-01	Malay	Female	5	37
NN2-03	French	Female	5	26
NN2-02	French	Female	3	22
NN2-04	Shona	Female	2	21

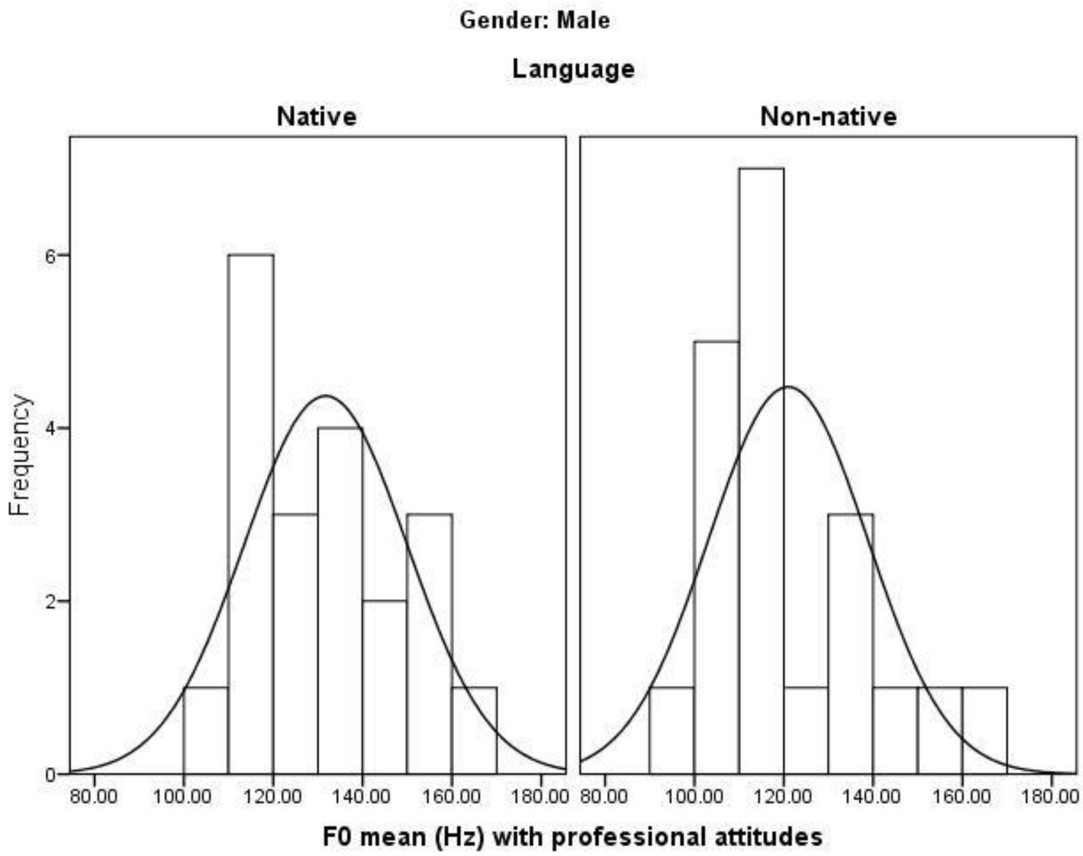
Appendix K. Frequency distribution of prosodic features in sentence reading data (histograms)

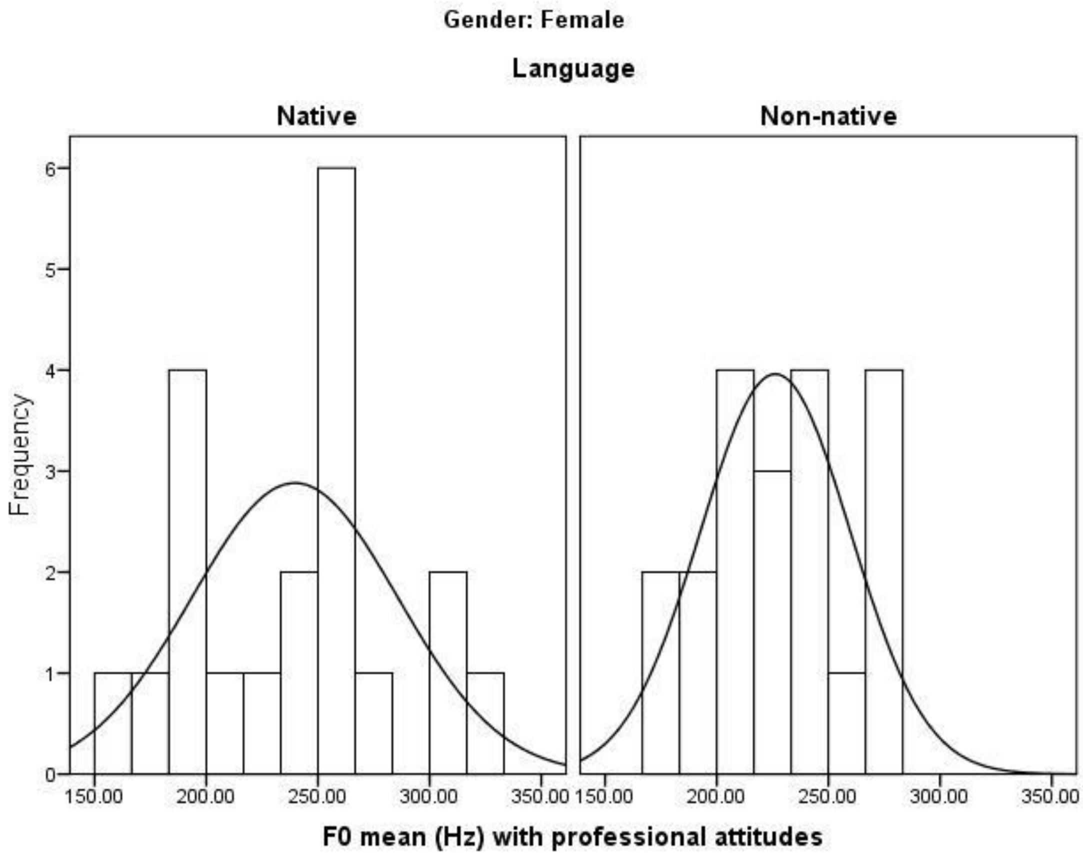


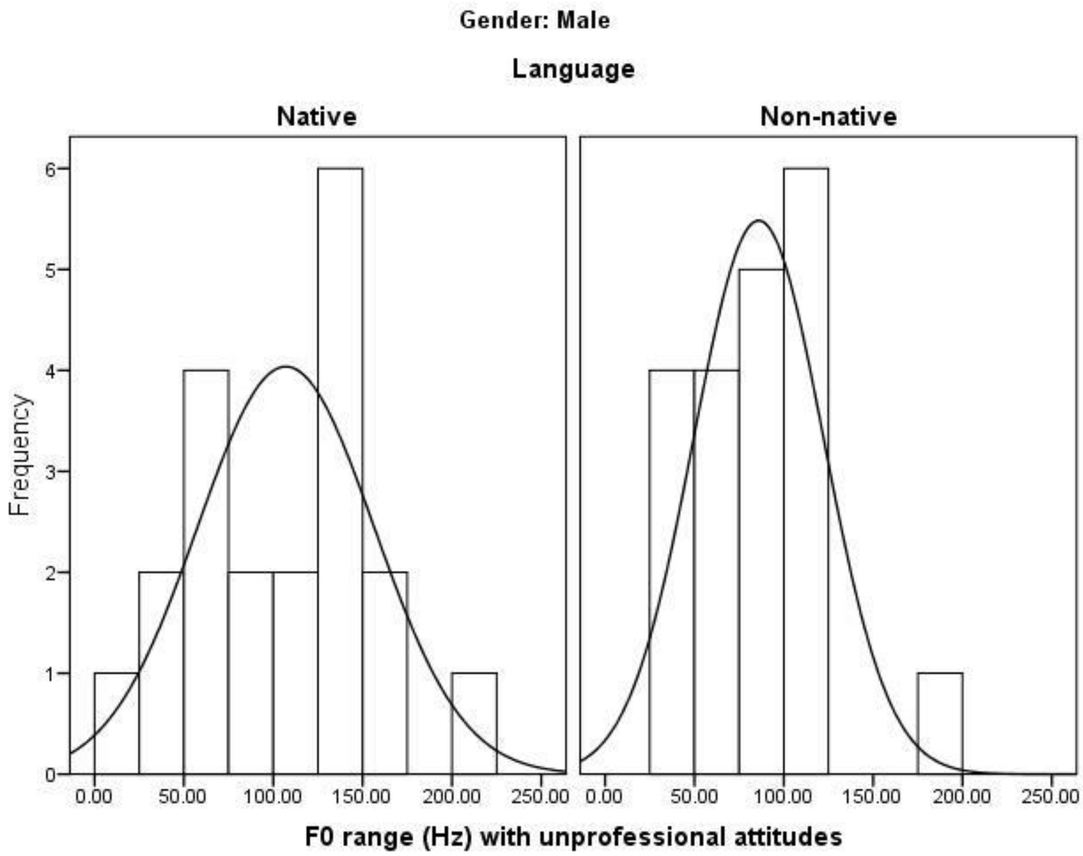
Gender: Male
Language
Native
Non-native
Frequency
0
2
4
6
8
80.00
100.00
120.00
140.00
160.00
180.00
200.00
F0 mean (Hz) with unprofessional attitudes

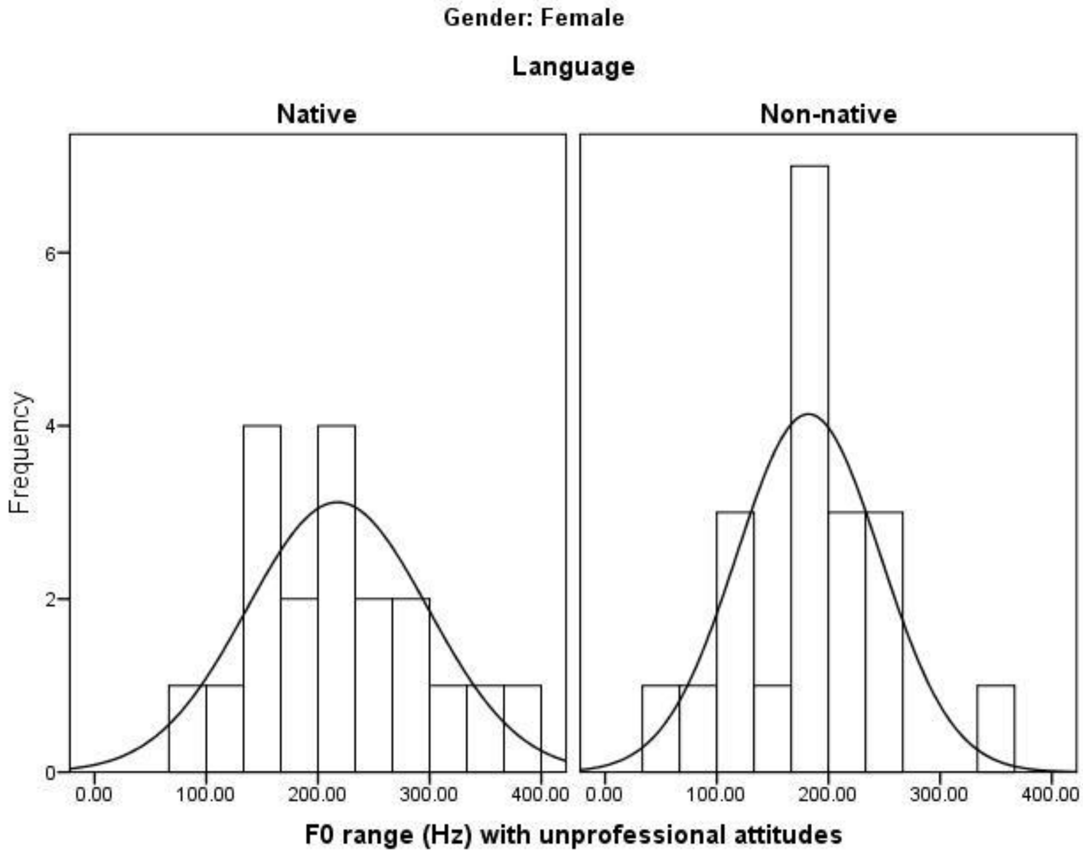


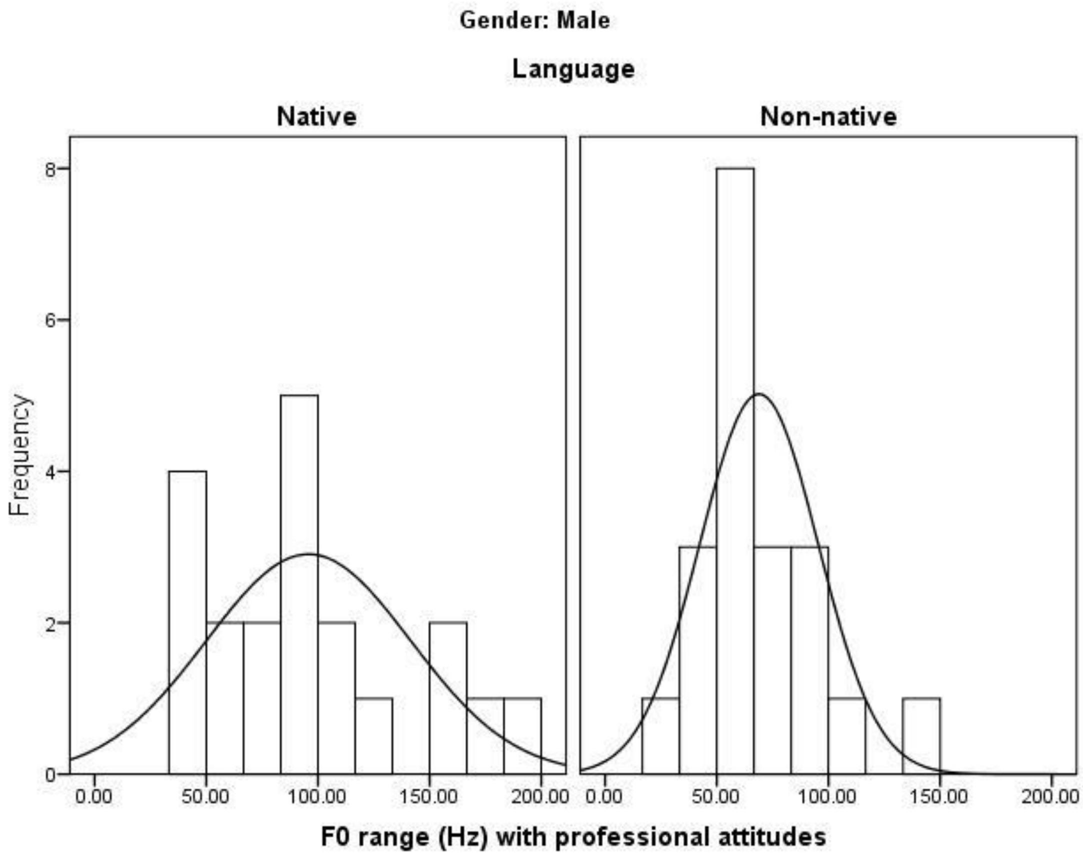
Gender: Female
Language
Native
Non-native
Frequency
0
2
4
6
150.00
200.00
250.00
300.00
350.00
400.00
450.00
F0 mean (Hz) with unprofessional attitudes


Gender: Male
Language
Native
Non-native
Frequency
0
2
4
6
80.00
100.00
120.00
140.00
160.00
180.00
F0 mean (Hz) with professional attitudes

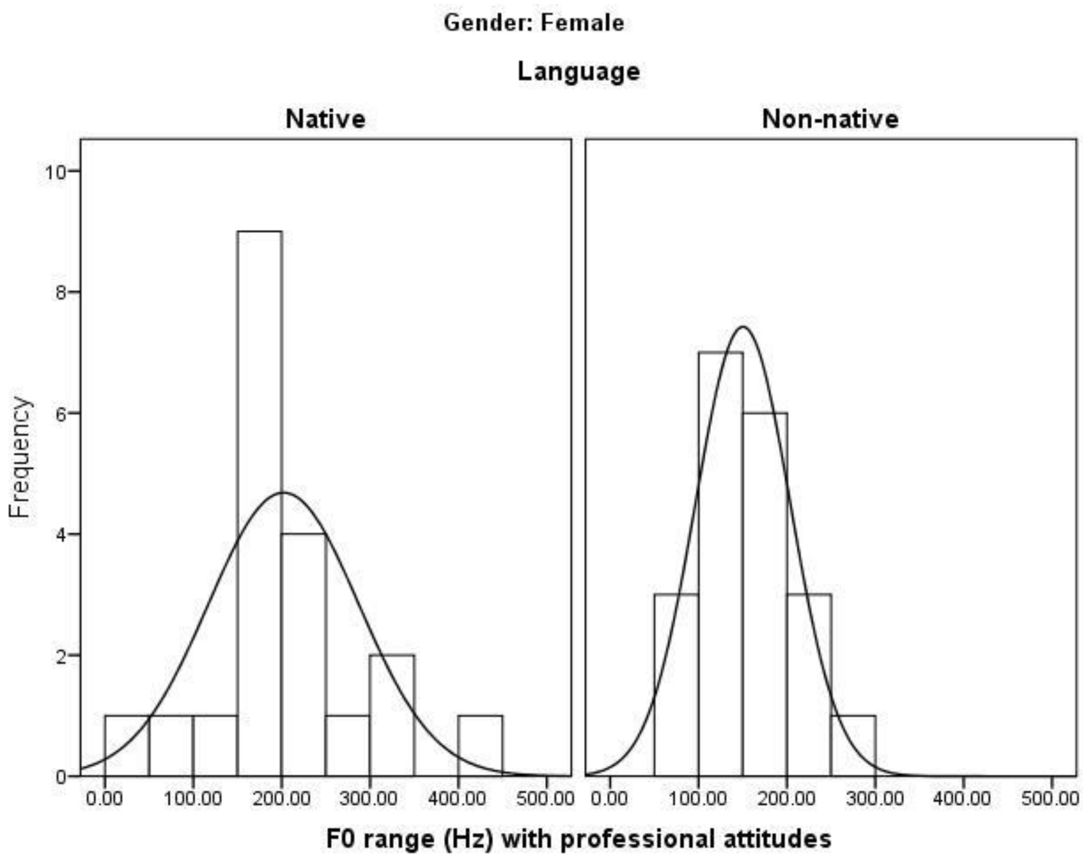

Gender: Female
Language
Native
Non-native
Frequency
0
1
2
3
4
5
6
150.00
200.00
250.00
300.00
350.00
F0 mean (Hz) with professional attitudes


Gender: Male
Language
Native
Non-native
Frequency
0
1
2
3
4
5
6
0.00
50.00
100.00
150.00
200.00
250.00
F0 range (Hz) with unprofessional attitudes

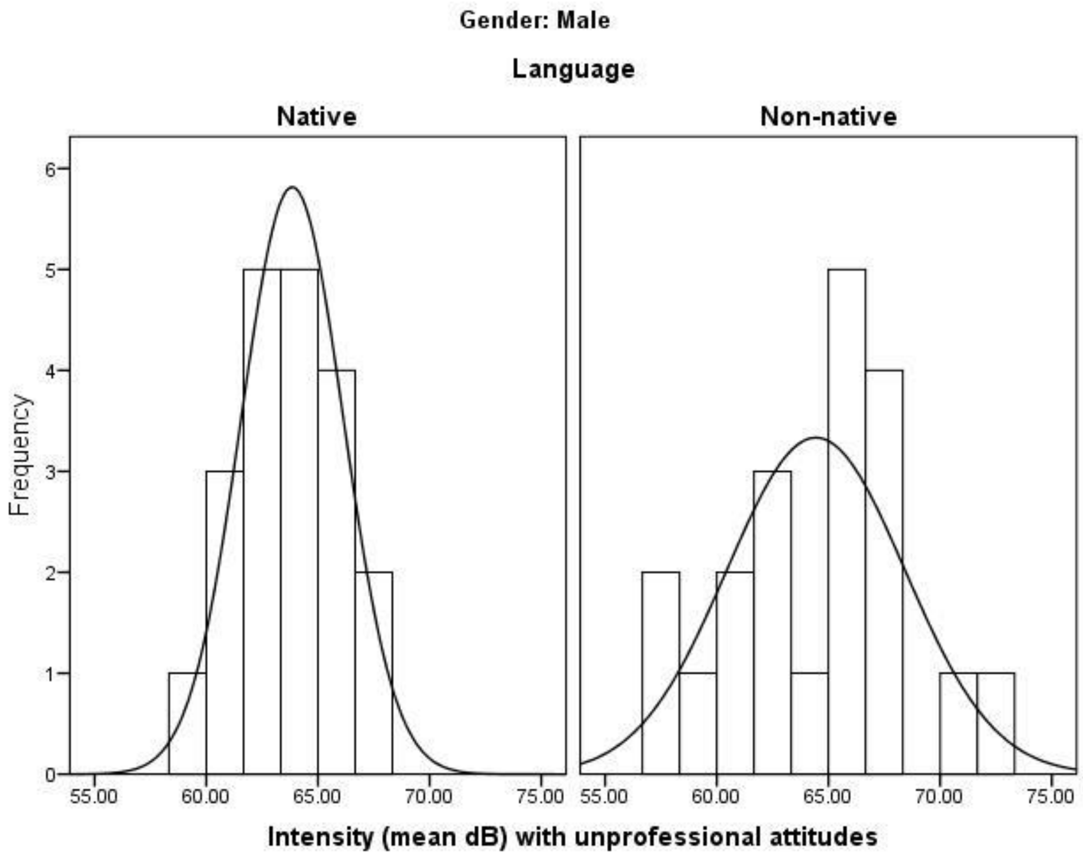

Gender: Female
Language
Native
Non-native
Frequency
0
2
4
6
0.00
100.00
200.00
300.00
400.00
F0 range (Hz) with unprofessional attitudes

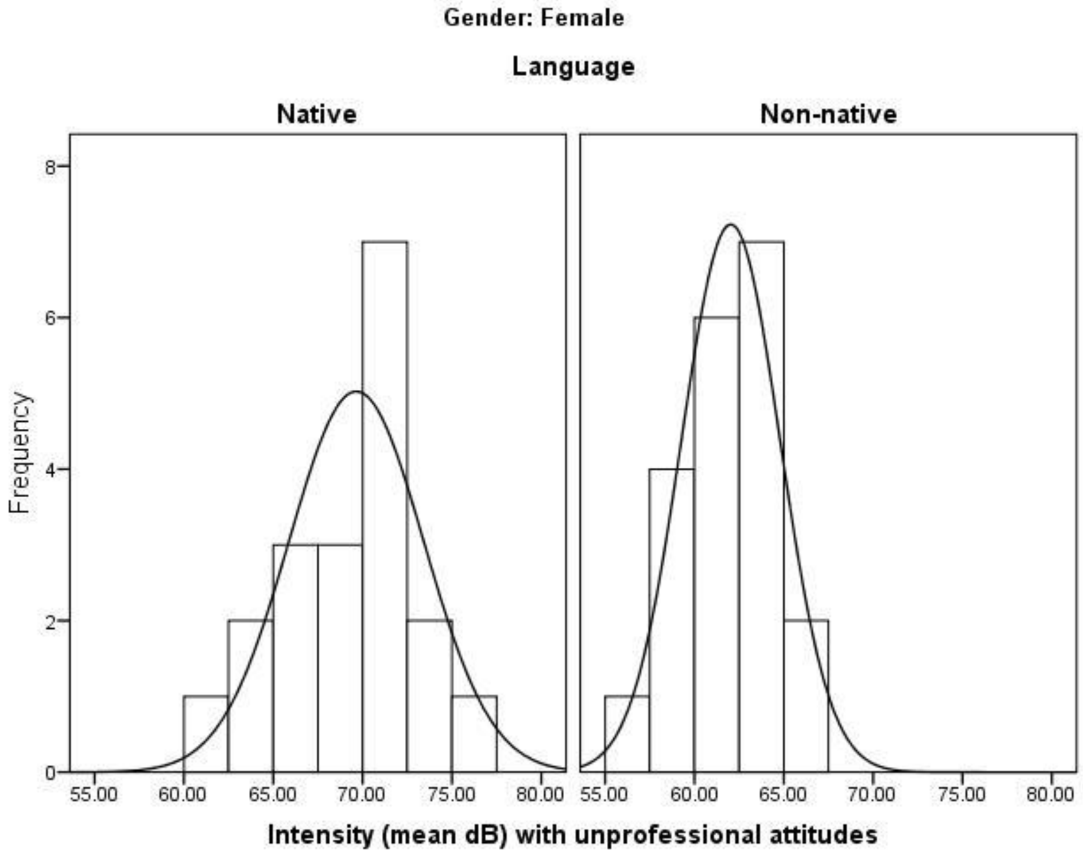


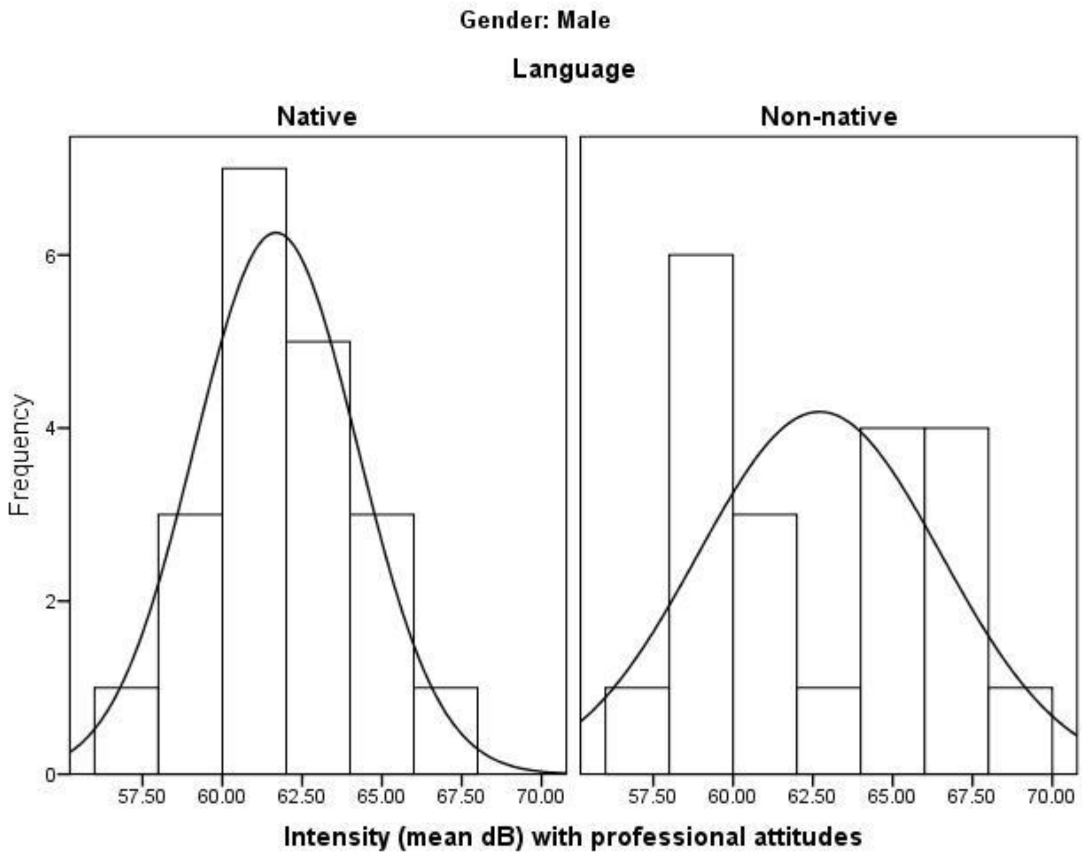
Gender: Male
Language
Native
Non-native
Frequency
0
2
4
6
8
0.00
50.00
100.00
150.00
200.00
F0 range (Hz) with professional attitudes

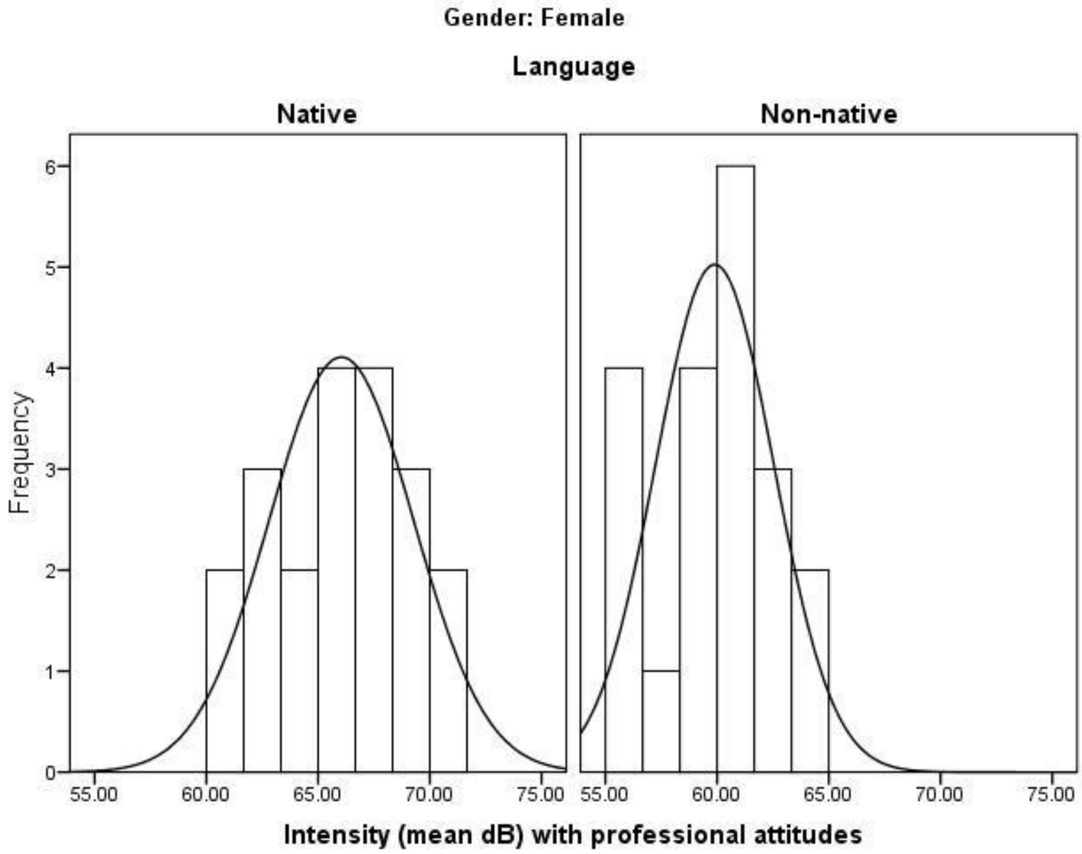


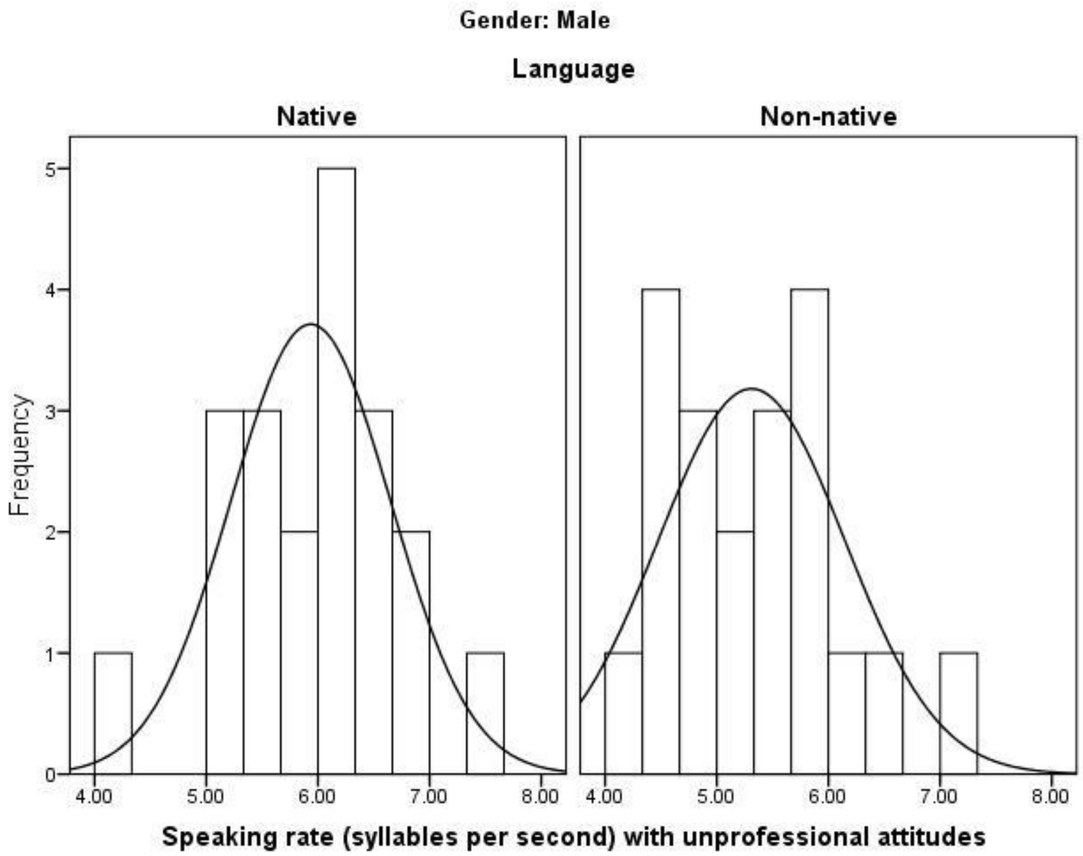
Gender: Female
Language
Native
Non-native
Frequency
0
2
4
6
8
10
0.00
100.00
200.00
300.00
400.00
500.00
F0 range (Hz) with professional attitudes

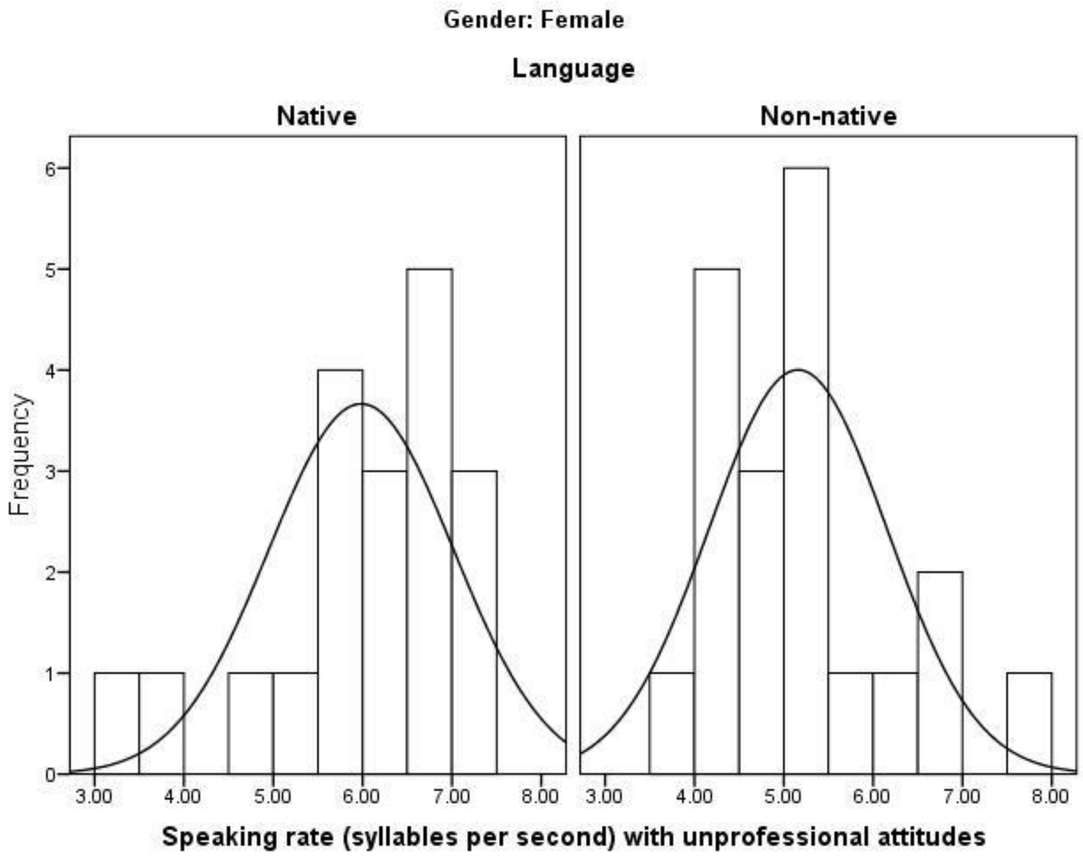

Gender: Male
Language
Native
Non-native
Frequency
0
1
2
3
4
5
6
55.00
60.00
65.00
70.00
75.00
55.00
60.00
65.00
70.00
75.00
Intensity (mean dB) with unprofessional attitudes

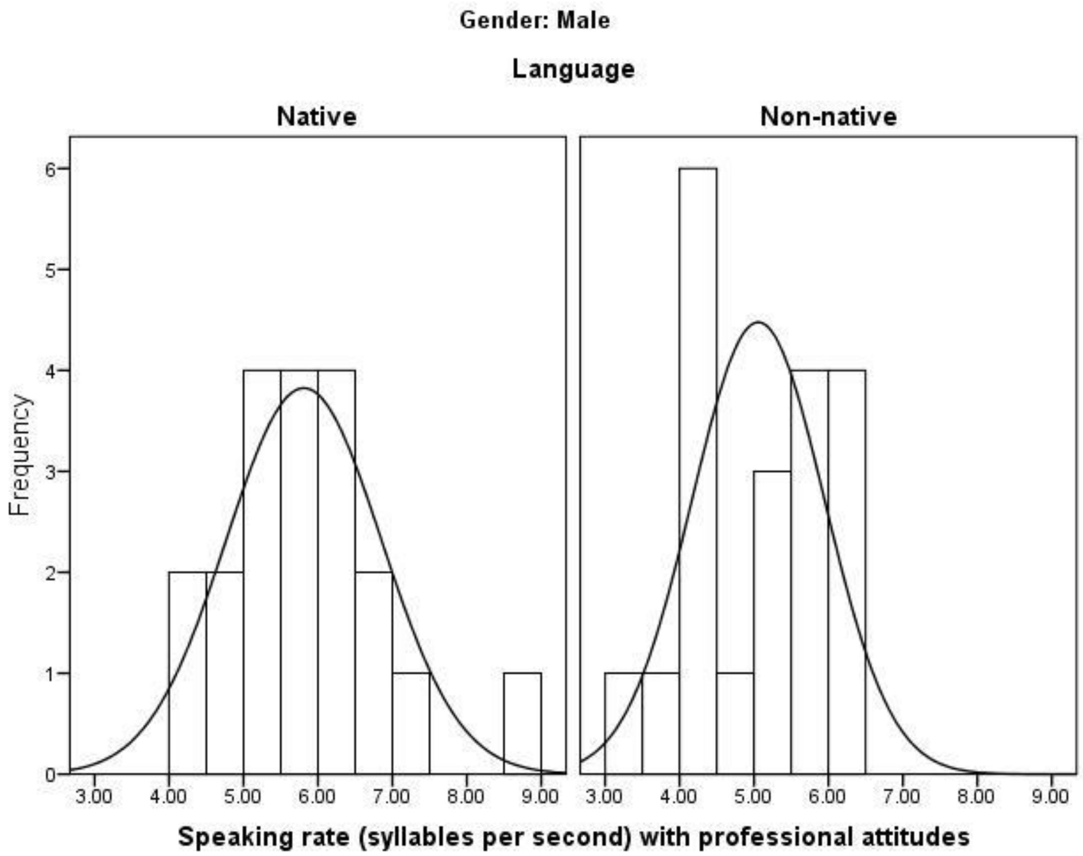

Gender: Female
Language
Native
Non-native
Frequency
0
2
4
6
8
55.00
60.00
65.00
70.00
75.00
80.00
55.00
60.00
65.00
70.00
75.00
80.00
Intensity (mean dB) with unprofessional attitudes

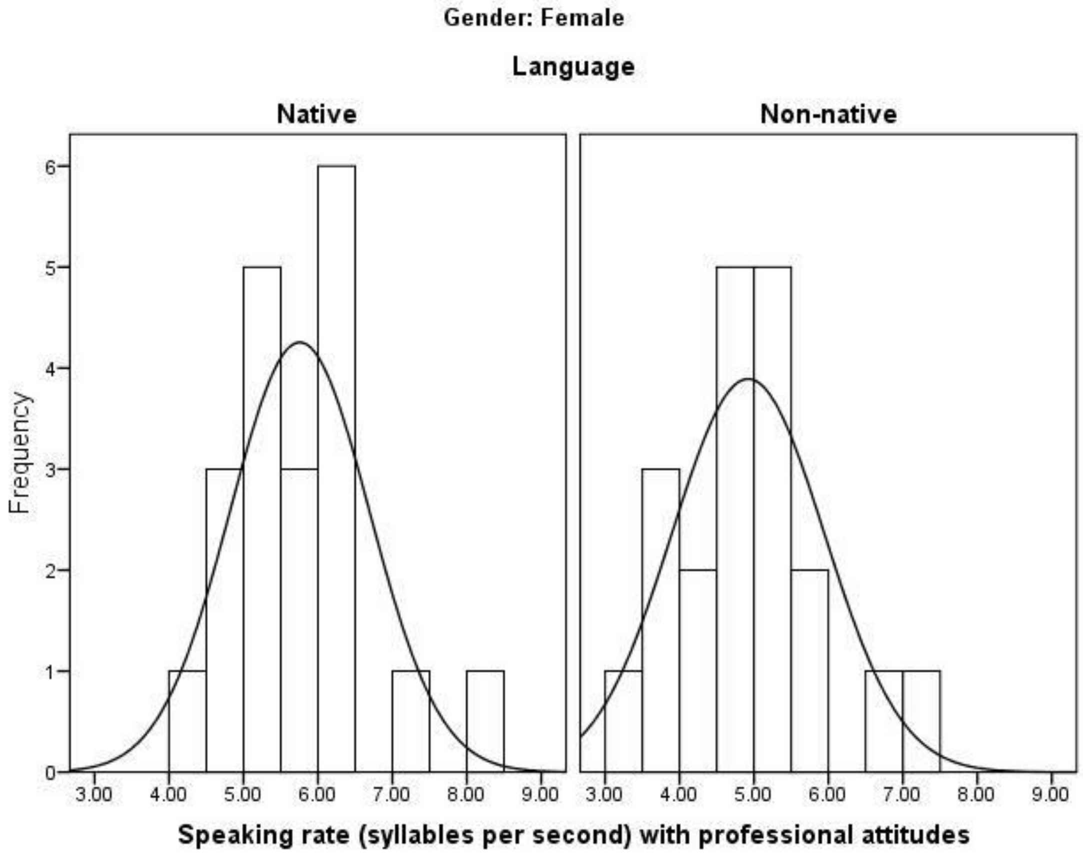

Gender: Male
Language
Native
Non-native
Frequency
0
2
4
6
57.50
60.00
62.50
65.00
67.50
70.00
Intensity (mean dB) with professional attitudes


Gender: Female
Language
Native
Non-native
Frequency
0
1
2
3
4
5
6
55.00
60.00
65.00
70.00
75.00
Intensity (mean dB) with professional attitudes


Gender: Male
Language
Native
Non-native
Frequency
0
1
2
3
4
5
4.00
5.00
6.00
7.00
8.00
Speaking rate (syllables per second) with unprofessional attitudes


Gender: Female
Language
Native
Non-native
Frequency
0
1
2
3
4
5
6
3.00
4.00
5.00
6.00
7.00
8.00
Speaking rate (syllables per second) with unprofessional attitudes


Gender: Male
Language
Native
Non-native
Frequency
Speaking rate (syllables per second) with professional attitudes


Gender: Female
Language
Native
Non-native
Frequency
Speaking rate (syllables per second) with professional attitudes

Appendix L. Ethical approval for the perceptual study

Faculty of Medicine and Health Sciences Research Ethics Committee

University of East Anglia

Fan Wang
School of Allied Health Professions
University of East Anglia
Norwich NR4 7TJ

Research & Enterprise Services
REN West (SCI)
University of East Anglia
Norwich
NR4 7TJ

Email: fmh.ethics@uea.ac.uk
Direct Dial: +44 (0) 1603 59 1566

Web: http://www.uea.ac.uk

8 October 2012

Dear Fan

The Effect of Speech Prosody on Attitudinal Perception in Medical Communication

Our Ref: 2012/2013-05

The submission of your research proposal was discussed at the Faculty Research Ethics Committee meeting on Thursday 27th September 2012.

The Committee considered that you had a very good application and they were happy to approve it in principle but have the following concerns which they would like you to address and amend accordingly:

1. Clarify how the social networks will serve as a "snow ball" technique for recruitment purposes.

2. Ensure that the information on the social media sites merely directs the interested participants to the relevant links on the Portal.

3. List the exclusion criteria on the Participant Information Sheet and on the Recruitment Flyer.

4. Correct the error on page 16 (Appendix 3), namely the sentence "On your CD, there is a PowerPoint slideshow to present to voice samples to you." should read "On your CD, there is a PowerPoint slideshow to present voice samples to you."

Please write to me once you have resolved/clarified the above issues. I require documentation confirming that you have complied with the Committee's suggestions. The Committee have requested that you detail the changes below the relevant point on the text in this letter and also include your amendments as a tracked change within your application/proposal. The revisions to your application can be considered by Chair's action rather than go to a committee meeting, which means that the above documentation can be resubmitted at any time. Please could you send your revisions to me as an attachment in an email as this will speed up the decision making process.

As your project does not have ethics approval until the above issues have been resolved, I want to remind you that you should not be undertaking your research project until you have ethical approval by the Faculty Research Ethics Committee. Planning on the project or literature based elements can still take place but not the research involving the above ethical issues. This is to ensure that you and your research are insured by the University and that your research is undertaken within the University's 'Guidelines on Good Practice in Research' approved by the Senate in July 2004.

The Chairman of the Committee is trying to amass a collection of applications which can be used for teaching purposes and would be very grateful if you would allow yours to be used as an example of a good application. It will of course be annonymised.

Yours sincerely

Yvonne Kirkham

Yvonne Kirkham
Project Officer

Faculty of Medicine and Health Sciences Research Ethics Committee

University of East Anglia

Fan Wang
School of Allied Health Professions
University of East Anglia
Norwich NR4 7TJ

Research & Enterprise Services
West Office (Science Building)
University of East Anglia
Norwich Research Park
Norwich, NR4 7TJ

Telephone: +44 (0) 1603 591574
Email: fmh.ethics@uea.ac.uk

Web: www.uea.ac.uk/researchandenterprise

09 November 2012

Dear Fan

Project title: The Effect of Speech Prosody on Attitudinal Perception in Medical Communication
Reference: 2012/2013-05

The amendments to your above proposal have been considered by the Chair of the Faculty Research Ethics Committee and we can confirm that your proposal has been approved.

Please could you ensure that any further amendments to either the protocol or documents submitted are notified to us in advance and also that any adverse events which occur during your project are reported to the Committee. Please could you also arrange to send us a report once your project is completed.

The Committee would like to wish you good luck with your project

Yours sincerely

Yvonne Kirkham

Yvonne Kirkham
Project Officer

Appendix M. Approval from UEA Survey Office

RE: approval for recruiting research volunteers at UEA

Sreeparna Ghosh (PLN)
Sent:20 August 2012 09:51
To: Fan Wang (AHP)

Dear Fan,
Yes the email can work as approval in principle. Could you kindly let FMH know that the form is no longer necessary for Postgrad researchers and others because I am no longer part of DOS?
Best wishes,
Sree
From: Fan Wang (AHP)
Sent: Monday, August 20, 2012 9:50 AM
To: Sreeparna Ghosh (PLN)
Subject: RE: approval for recruiting research volunteers at UEA
Dear Sree

Thank you very much for your help. FMH suggested using this form to obtain approval in principle from the survey office https://www.uea.ac.uk/polopoly_fs/1.151266!survey_form.pdf

Apparently this is no longer part of DOS. Is the form still applicable, or perhaps our email correspondence can be regarded as approval in prinicple?

Many thanks again for your support

Fan

Fan Wang
Postgraduate Researcher (AHP)
Room 0.27 The Queen's Building
University of East Anglia
Norwich
Norfolk NR4 7TJ
Email: fan.wang@uea.ac.uk
Mobile: 07578008130

From: Sreeparna Ghosh (PLN)
Sent: 20 August 2012 09:25
To: Fan Wang (AHP)
Subject: RE: approval for recruiting research volunteers at UEA

Dear Fan,
Thanks for your email. Yes please feel free to contact the Press Office with regard to advertising your project in the e-bulletin and the Portal.
I don’t manage the portal advertisements and they are best placed to help you.
Best wishes,
Sree

Dr. Sreeparna Ghosh
Survey Research Manager
Business Intelligence Unit
Planning Office, REG 2.10
University of East Anglia
Norwich Research Park
Norwich NR4 7TJ

Phone: 01603-593419 (x3419)
http://www.uea.ac.uk/planning/biu/survey-office
From: Fan Wang (AHP)
Sent: Sunday, August 19, 2012 9:33 PM
To: Sreeparna Ghosh (PLN)
Subject: approval for recruiting research volunteers at UEA
Importance: High
Dear Dr Ghosh

I am writing to obtain your approval for recruiting volunteers from the University community for a listening experiment for my PhD project.

The aim of the experiment is to ascertain the role of certain features in speech (pitch, volume and rate) in listeners' perception of doctors' professional attitude in clinical interactions. The participants will be asked to listen to a number of pre-recorded sound files and judge the attitudes they perceive on an easy-to-use scale. The estimated time for each experiment session is about 60 minutes, although an additional 10 minutes will be needed for half of the participants who need to listen and judge more sound files to establish the reliability of the instrument.

I should be most grateful if you would allow me to announce the call for volunteers through the University Portal and newsletters. Interested volunteers will be asked to complete a brief screening questionnaire after reading the information sheet. This will enable the selection of participants according to the inclusion and exclusion criteria. Suitable volunteers will be contacted to confirm their willingness and availability to participate.

This is the final stage of my project and I appreciate your support in this critical step of my research.

With best wishes

Fan

Fan Wang
Postgraduate Researcher (AHP)
Room 0.27 The Queen's Building
University of East Anglia
Norwich
Norfolk NR4 7TJ
Email: fan.wang@uea.ac.uk
Mobile: 07578008130

Appendix N. Recruitment flyer

University of East Anglia

Faculty of Medicine and Health Sciences
School of Allied Health Professions

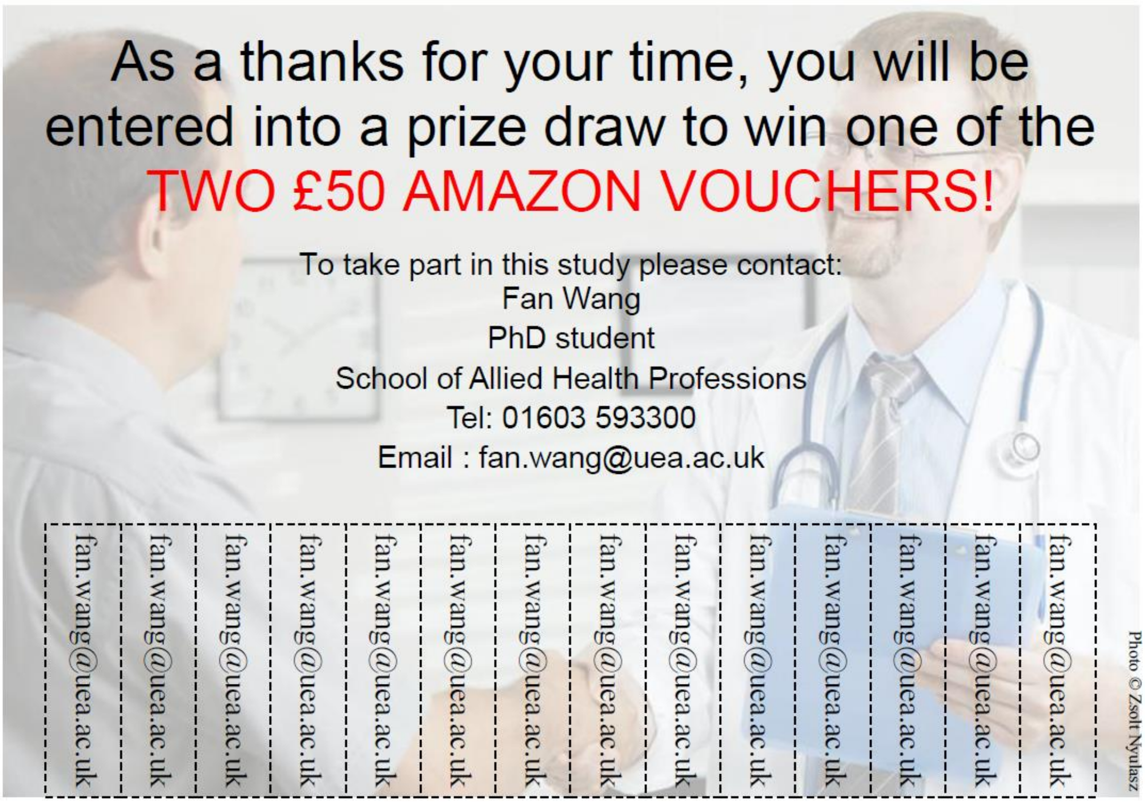
LISTENER PERCEPTION OF ATTITUDES IN MEDICAL COMMUNICATION

Have you sometimes thought that **how** people say something is more important than **what** they say? if so, we need your help!

We are conducting an experiment to investigate the link between the ways things are said and listeners' perception of the speaker's attitude, and we hope you can take part.

We will ask you to listen to a number of speech samples and then rate each one in terms of friendliness, respectfulness, etc. This will take about an hour of your time in our computer lab.

Results from this study may inform communication skills training for medical professionals in the future.



Appendix O. Participant information sheet

The Effect of Speech Prosody on Attitudinal Perception in Medical Communication

PARTICIPANT INFORMATION SHEET

I am doing PhD research that investigates the importance of prosodic features (pitch, volume and rate of speech) in medical communication. This document explains some details about the project and how you can help. Before you decide whether or not to agree to take part, it is important that you understand clearly the elements of this study that are relevant to you. If you are still interested, please take time to read the following information carefully. If you have any questions, please do not hesitate to contact me fan.wang@uea.ac.uk.

What is the purpose of this study?
We want to investigate the influence that prosodic features (such as pitch of the voice, volume and rate of speech) have on listeners' perceptions of attitude. In particular we are interested in this in the context of medical interactions, e.g. doctor –patient interactions. If these features of the 'way something is said' are shown to influence listeners' judgments we can make recommendations for the education of trainee doctors.

Why have I been asked?
We hope to recruit a cross section of the whole UEA community, including both student and staff members. You have been asked for help because you study or work at UEA.

Do I have to take part?
No, it is entirely up to you to decide whether to volunteer. If you do volunteer, you may or may not be selected for further study. If you are selected, we will ask you to listen to and to rate some voice samples in our Communication Lab. Even if you volunteer to take part, you can withdraw at any time without giving us a reason, and the data you provide will be immediately destroyed. If you decide not to volunteer, this will have absolutely no effect on your position within UEA – to volunteer or not to volunteer is entirely up to you.

Who gets selected?
You can volunteer if you are

- a native speaker of English
- without hearing or speech difficulty
- under 65 years old
- not a student or academic staff in the Faculty of Medicine and Health Sciences
- not trained in phonetics or speech sciences

Since we need to control for certain factors in this study, these criteria are necessary and we need you to confirm using the attached short questionnaire. Please complete it and return to the researcher (fan.wang@uea.ac.uk) if you want to participate. We will be in touch to let you know if you are selected.

If you were to volunteer but not be selected this would be no reflection of your ability to take part; it would be simply because our inclusion criteria for this particular study did not apply to you. You may well meet the criteria for future extended studies.

What will I have to do?
If you are selected to take part in the study , you will be invited to attend a data collection session at UEA . You will be given the choice of a range of dates and times. We will contact you to confirm your availability, and invite you to come to our Communication Lab (Queen's Building on campus)for the session. You only need to attend one session, and your session will last between 60-90 minutes. During the session you will hear a range of voice samples, and will be asked to rate a number of aspects of the voice you hear, e.g. 'how friendly do they sound?' We will give you the necessary instruction/training at the beginning of the session. That is all we need you to do for this study.

Will my personal information be kept confidential?
Yes, all your personal information will be kept strictly confidential in accordance with the Data Protection Act 1998. All data you provide will be anonymous and you will be referred to by a participant number throughout the project. Your data will be stored in locked cabinet and on password protected UEA server space.

What will happen to my data?
Statistical analysis will be carried out on the rating you provide. These findings will then be related to analyses of the voices themselves. We will not analyse the correctness or accuracy of your ratings. A statistical expert will be consulted in the data analysis, but he will **NOT** access your personal information.

According to the Data Protection Act 1998, your data will only be used for future research purposes if you allow us to do so in writing. If you do not agree to us using your data for future research, they will be archived for five years after the end of the study and then destroyed. If you allow us to use your data for future research, your data will be kept securely with the same level of confidentiality for five years after the end of the study and will be destroyed after that.

What will happen to the results?
The results of this study will be included in the researcher's PhD thesis and publications in academic journals. Your personal information will **NOT** be revealed in any publication.

Who is funding / supervising the research?
This research is funded by the UEA Faculty of Health Studentship and the UEA International Scholarship, both awarded to the postgraduate researcher. A team consisting of Dr Christine Raschka (AHP), Dr Zoe Butterfint (AHP) and Prof Jacqueline Collier (AHP) is supervising the research.

Who has reviewed the study?
The Faculty of Health Ethics Committee at UEA has reviewed and approved this study.

Are there any risks to me?
No, this study is completely non-invasive and will pose no risk whatsoever to you. Stress resulted from the study is highly unlikely.

What should I do if I have any questions or concerns?
If you have any questions or concerns about the study, please feel free to contact the researcher. If you have any complaints about your experience of the project, you can

contact Dr Nicola Spalding (Deputy Head of School) or Dr Christina Jerosch-Herold (PGR Lead).

Fan Wang
Postgraduate researcher
School of Allied Health Professions
Room 0.27, The Queen's Building, UEA, Norwich, Norfolk NR4 7TJ
fan.wang@uea.ac.uk
01603 593300

Appendix P. Screening questionnaire for the perceptual study

Participant No. (to be completed by researcher) ☐

Faculty of Medicine and Health Sciences
School of Allied Health Professions

University of East Anglia

Participant Recruitment Questionnaire

The Effect of Speech Prosody on Attitudinal Perception in Medical Communication

Before you continue, please make sure you have read the participant information sheet carefully and have understood all the relevant information regarding the study. In order to select participants for the study, we need to collect some basic information. If you still want to volunteer for the study, please take a few minutes to complete this questionnaire and return it to me by email. The questionnaire should be easy to fill in. You can use TAB key to move to the next box, or click to choose a box and type in your answer. Please answer all questions.

1. Name:

2. Contact e-mail:

3. Gender: Male ☐ Female ☐

4. Occupation:
   Are you a UEA student ☐
   or a UEA employee ☐

5. Age:
(Please state in years. We will only use this information for statistical purposes)

6. Can we contact you on your mobile phone? Yes ☐ No ☐
   If yes, please provide your number

7. Are you a native speaker of English? Yes ☐ No ☐

8. Would you consider yourself to have:
   a hearing difficulty Yes ☐ No ☐
   a speech difficulty Yes ☐ No ☐

9. Please select all sessions for which you would be available (we will then assign you to only one of them).

Wed 28 Nov	12.00-13.00	☐	Mon 03 Dec	10.00-11.00	☐
Wed 28 Nov	14.00-15.00	☐	Mon 03 Dec	12.00-13.00	☐
Wed 28 Nov	17.00-18.00	☐	Mon 03 Dec	14.00-15.00	☐
Fri 30 Nov	10.00-11.00	☐	Wed 05 Dec	12.00-13.00	☐
Fri 30 Nov	12.00-13.00	☐	Wed 05 Dec	14.00-15.00	☐
Fri 30 Nov	14.00-15.00	☐	Wed 05 Dec	17.00-18.00	☐

Thank you! Please return this as an email attachment to the researcher at fan.wang@uea.ac.uk

For more information, please contact:
Fan Wang, postgraduate researcher
Room 0.27, The Queen's Building
01603 593300

Appendix Q. Consent form

Participant Reference No. ☐

Faculty of Medicine and Health Sciences
School of Allied Health Professions

University of East Anglia

CONSENT FORM

Project title: ***Does speech prosody matter in medical communication? An exploration of speaker attitudes in professional context***

Name of Researcher: Fan Wang

Please initial the boxes as appropriate

1. I confirm that I have read and understood the participant information sheet for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason. I understand that should I wish to withdraw from the study, all data I have provided will be destroyed. ☐

3. I agree, as outlined in the information sheet, to provide data for this study by completing the rating scale during the research session and, where relevant, that the results of analysis of my data may be included anonymously in academic publications. ☐

4. I understand that the data that I provide during the study will only be viewed by the research team, that is, the researcher and the supervisors including statistical experts and will not be viewed by anybody else. I understand that my personal information will be kept strictly confidential. ☐

5. Please select **one** of the following as appropriate

   I **do not** wish the data that I provide to be used for future research purposes after completion of this study. ☐

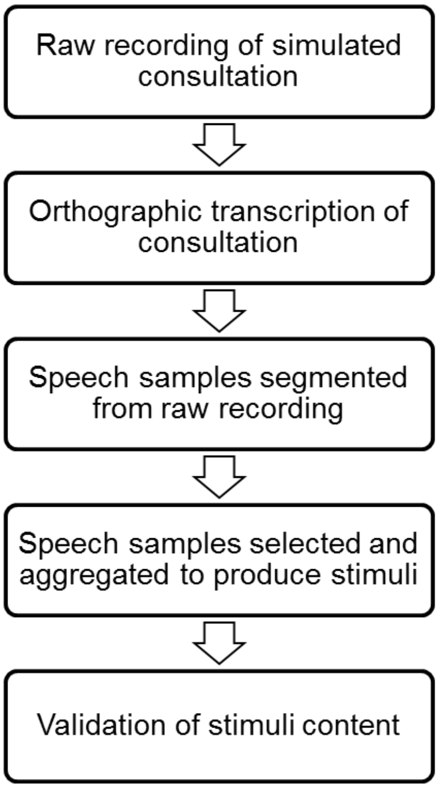
   I **agree** to using the data I provide during the study for future research purposes after completion of this study, for a period of up to 5 years. ☐

6. I agree to take part in this study. ☐

Name of Participant	Date	Signature
Name of Researcher	Date	Signature

Appendix R. Perceptual study stimuli selection procedure

The flow chart below illustrates the steps from the recording of the simulated consultation (see Section 4.4.2) to the validation of stimuli content for the perceptual experiment.



A worked example following these steps is given below. It is based on Participant NN1-01, a male, non-native English speaking, 5th-year medical student interviewing a simulated patient.

a) Transcript of recorded simulated consultation

Speaker NN1-01 (Male, non-native, Year 5 medical student)

Key to transcription symbols:

MS = medical student; SP = simulated patient

( )	Unintelligible fragment
(guess)	Best guest of unclear fragment; also used here for anonymity
(( ))	Non-verbal activity
(0.5)	Pause; duration shown in seconds
(.)	Pause shorter than 0.2 second in duration
[	Onset of overlapping turns
]	Offset of overlapping turns
=	Latching turns (turns with no gap in between)
:::	Stretched segment; the more colons the longer the stretching

·hh	Audible inhalation; the more h's the longer the breath
hh	Audible exhalation; the more h's the longer the breath
↑	Marked rising intonation
↓	Marked falling intonation

MS: Hello, Mr Jones .
SP: Hello.
MS: Hi. My name is (1-syllable name) . I'm a fifth year medical student.
SP: All right.
MS: I've been asked by your GP to come and, um, have a s:::peak to you. Is that okay with you?
SP: That's right, yeah, yeah.
MS: Right. Thanks. Uh. Right. So, um, can I ask you why … that you come to see your doctor?
SP: Well, what it is um, 'bout three months ago, uh, my previous doctor =
MS: =uh-huh.
SP: he sort of just gave me a, um routine blood pressure test. And that evidently came out a bit high. So he then gave me some medication to take, which I, I then took for a month, you know. But in that time, I actually moved, I, you know; I moved in the town from Brundall=
MS: =O↑kay
SP: so I didn't see that doctor no more. And I thought, well I, I've taken the medication. That, that ought to be enough. Evidently now, when I spoke to my neighbour, my new neigh↑bour, he said, uh, "you gotta take that the rest of your life, you know. You can't just [stop taking] that.
MS: [oh, well]
SP: So I thought, oh well, you know, I must have forgotten he said that or something. So I'd better come in and, and [I, I don't know. I'll have, have a check],
MS: [Okay (1.79) Okay]
SP: anyway, to see if that's right, you know=
MS: =Yeah. So, um, about three months ago, you've seen your GP and they found that you were having high blood pressure. [And, um::: um (0.9) so
SP: [That's right. (.) High blood what, what's that
MS: Uh, high blood pre↑ssure, (.) [your blood pressure's a bit high]
SP: [Oh yes, that's right.]
MS: And, um, they gave you some tablets.
SP: Yeah=
MS: =Yah, what, um, what's the name of [the]…
SP: [I] think they're called <u>Ramipril</u>=
MS: =Okay Ramipril [↑yeah]
SP: [Yeah]
MS: And, um, how long would that last you [for? Are you]
PS: [Well, I took it] for a month, you see.
PS: That's all. [That's just like the course of tablets. And I'll take them, and that'll do it, you know
MS: [An::d, okay (0.9) okay (1.0) okay
MS: And, it's run out already. [It's
SP: [That's right, yeah]

MS: Okay]
SP: It's run out 'cos that was (right) three months ago.
MS: O:::::kay. Soon, then you moved to, um, here.
PS: That's right. I was just chatting to the neighbour and then like, I::think his brother's got high blood pressure, and he's saying "you [um, you can't stop taking] that, you know=
MS: [Yeah, okay (0.7) okay]
MS: =Okay.So you are coming here to see the doctor [to um]
SP: [That's right]
MS: to have a new measurement and see, uh=
SP: =Yeah, he now took my blood pressure, I think that was a (.) [hundred and sixty-two over a hundred or something (like that)
MS: [okay, okay, so it, um
MS: Do you know what's the previous reading for the blood pressure, like the first time
SP: I think that was a hundred and sixty something over (0.6) [over a hundred and something (like that)
MS: [·hh Okay (0.5) okay
MS: so, it's still a bit on the high side, yeah?=
SP:=Yeah
MS: And=
SP: =Is that
MS: Yeah, and u:::::::m (0.3) ((tut sound)) ·hh Yeah, well, just to um: (0.8) make sure I, um, know what's your understanding of high blood pressure. Do you (kn:::), have you been told about, anything about (1.0) um=
SP: =Well (0.7) a thing is that my dad, he died of a stroke when he was 72. And, and, the doctor at the time said that (0.3) that's partly the high blood pressure what caused it [(you see), so] (0.8) you know I don't want=
MS: [uh-huh]
MS:=Yeah (that)
SP: to go the same way really, and, and the thing is that my brother (.) he lives in New Zealand, but he's evidently got high blood pressure [as well]
MS: [uh-huh]
SP: (0.5) but I don't know a lot about it 'cos obviously he's over there and, you know about like 15 years so=
MS: =Okay
SP: But that looked like it might run in the <u>family</u>. [Does that sort] of thing happen, does it?
MS: [↑hmmm]
MS: Yean, well, a::ctually, it's, it's quite common when, um, in the normal population=
SP: Is it?
MS: [Yeah]
SP: [(Uh)]
MS: But may I ask you for your age, please? How old are you?
SP: I'm 48=
MS: =48, and [you're still
SP: [Yeah, my brother's 52
MS: ↑Okay=

SP: =and my dad was 72 when he died.
MS: (0.3) Okay (.) And, um, are you still working (.) at the [mo]
SP: [Oh yeah, yeah]
MS: Okay. What, what's your [occupation
SP: [Work at supermarket, yeah, Tesco's, yeah]
MS: ah, okay]
MS: ·hhh And, u::m, (0.6) right. So <u>before</u> we talk about the::: high blood pressure and the:: (0.3) ((tut sound)) (0.3) the uh medi↑cations ·hhh um, (0.7) may I ask you something about your past medical history?
SP: [Okay]
MS [Like, like] Um have you been to the hospital before for any
SP: [No
MS:[thing
MS:↑No]
SP: No]
SP: No (previous mild really) I mean. I feel all <u>right</u>, you know. I [feltl] all right when they said "Your blood pressure is ↑high"
MS: [Hmmm]
SP: I felt all right when I took the (.) Ramipril=
MS: =Okay
SP: I feel all right <u>now</u>, you know.
MS: Okay, so you are not on any medications [or]
SP: [No]
MS: um an, an, any surgeries or::=
SP: =No.
MS: Okay. Right. So you mentioned family history. Uh, your father and your brother had, uh, high blood pressure
SP: Yeah
MS: How about any other things that might run in the family you're aware of=
SP: =oh no, not really, [no]
MS: [Okay]
SP: If that's right you see now that's about it really.
MS: Okay. And, um, I gonna ask you some questions about, um, your social history, like, um, we ask everyone, um ·hh whether you smoke or not.
SP: I I gave up 10 years ago, yeah=
MS: Okay ·hh and when did you start (0.3) smoking?
SP: Um, when, in my teens, you know
MS: Okay=
SP: = I smoked for about (0.3) 20 years [(long)], yeah.
MS: [Okay]
MS: And, um, on average, how many would you smoke (.) [a day?]
SP: [Um]
SP: Five or ten something [like that]
MS: [Okay]
MS: You gave up ten years a↑go=
SP: =Yeah. [That
MS: [Yeah
SP: was quite hard to give up (that like you know) yeah]
MS: Yeah]
MS: It's really good then. Well done. [Um (0.3)] How about the…alcohol?

SP: [( )]
SP: (0.4) I'm, I'll share a bottle of wine with my wife on a, on a weekend [you know]
MS: [↑hmmm]
SP: Yeah, that's about it, [really]
MS: [↑Okay]
SP: We don't, we don't [go out] that much.
MS: [↑Okay]
MS: Yeah, sounds good. An::d, allergic to anything?
SP: Don't think so, no.
MS: Okay, no, no drug allergies, or
SP: (1.0) No.
MS: Okay, (0.3) okay, ·hhh right, so, um::: (0.9) u:::m, let's talk ab', talk a bit about your u::m high blood pressure, and the uh, the drug that you're taking, ↑yeah? ·hhh Um, well, you are (0.5) quite right that high blood pressure might cause stroke an::d other things, but (.) it (.), it doesn't really cause it. May it's a risk factor=
SP: =All ↑right
MS: so, uh, that if, if you(r) (.) if your blood pressure is (.) a bit high, then it increases the risk (0.5) of stroke or (.) other heart diseases ·hh=
SP: =Yeah.
MS: And that's why, um, (0.7) um (0.6) it is::: (.) it would be a good idea to::: bring it back down with the medications=
SP: =Right.
MS: Um, (1.1) how::: did you feel when you were on Ramipril?
SP: I, I [didn't feel] that different really.
MS: [Did you feel]
SP: I felt [all right], yeah.
MS: [Yeah]
MS: Okay=
SP: =Yeah
MS: ·hh Yeah? Because, uh, a lot of drugs will have their own side effects. ·hh And Ramipril is, (.) well, that's (a) class of drugs (that) might give some patients a, a cough::=
SP: =All ↑right
MS: Yeah, [so but if it]
SP: [Just watch out for that you know]
MS: Yeah, just uh::, well it doesn't affect everybody, so, but if::, if you realise then you:: have a cough, then ·hhh you [m: you m: you might]
SP: [Guess that might be why]
MS: Uh the doctor might change the drug for you to something else (.) Yeah ·hh so but that's not a (0.4) big ↑problem. ·hh And uh (1.0) yep, so (1.0) I guess if you speak to the:: doctor today, he might start you on um, (0.5) Ramipril (0.3) as ↑well, and s::: um, (0.5) keep monitoring the blood pressure every now and then, maybe a few months' time, and if it's still high, they might (0.3) increase the dose or <u>add</u> another drug. ·hh And if it's actually going (0.3) back down normal, well then they might actually take it off=
SP: =Oh ↑yeah
MS: [So, um:, yeah, there's
SP: [That'd be good then yeah
MS: different

SP: I were a bit worried about having to take the drug all the time, like, you know, will I remember, and I'm not that old. You know, I don't want to start taking drugs all the time, you know, any sooner than I have to, you know]
MS: ↑Uh-huh, hmmm, ↓hmmm, yeah]
MS: ·hh And there're also other things that (.) uh, can help you with ((unintelligible vocalisation)) getting off the drug, let's say yeah, how about your ↑diet
MS: ·hh [Um]
SP: [All right]
MS: Do you eat a lot of salty ↓food?
SP: I do like salt on me food, yeah. My wife says I put too much on it, yeah.
MS: Yeah, because uh (a) (0.5) like a diet with a lot of salt in seems to um, increase the blood pressure=
SP: =All right.
MS: So um, (.) actually if you (0.8) if you wouldn't mind trying reduce the salt, [then you might see a]
SP:
[Yeah I can have a (turn)]
MS: difference in the blood pressure=
SP: =Really?
MS: [Yeah]
SP: [O↑kay]
MS: And um, any vegetables or meat (that) you eat
SP: Yeah, I mean, I do, we do eat vegetables, you know, I like, we like to have a proper tea everyday=
MS: Hmmm
SP: (I cook), cook up something
MS: Yeah. Basically, just (0.3) uh healthy meal and ↑exercise. Do you do a lot of exercises?
SP: Not re::ally. [No, I work in a supermarket] the till, you know
MS: [Hmmm. Yeah]
SP: so I am moving around unless I'm on the till
MS: Hmmm
SP: but I thought I could do more, yeah.
MS: Is the working hour (.) pretty long, though?
SP: Not too bad (in what I'm doin') that varies, you know, sometimes you go in early
MS: Hmmm=
SP: =um (0.4) but I never, I never go in earlier than about half past six. And then I never finish later than ten. (0.5) [You know you do sort of eight hours and, uh an hour for lunch, you know]
MS: [Yeah: because, yeah]
MS: Because some (.) some:: (.) people tend to find it um, helps (that) when they walk a lot more during the day. ·h Let's say uh, go out for a walk during lunch,
MS: [or]
SP: [Right, yeah]
MS: park your car further away and then walk towards the car park, uh, yeah.
SP: All right, yeah=
MS: =That (sort of) thing helps, so [exercise and::]
SP: [Suppose so yeah]

MS: diet. ·h And um (0.5) the medications (0.6) yeah so that sh:: that will definitely (0.3) um, decrease the risk of (.) stroke and:: other (.) heart diseases=
SP: Okay.
MS: Do you know about your cholesterol: level? Has it, [have you] been told
SP: [Um] (0.6) No,
MS: Okay=
SP: =not really.
MS: Yeah, so that [would be:: (.) good]
SP: [I think that, I think]
SP: Yeah, he mu, he mu, he must've taken it, I suppose it must've [come out] normal,
MS: [Yeah:]
SP: I suppose
MS: Yeah, because if you're not on any medication, then you shouldn't [be] worrying.
SP: [Yeah]
MS: So, that's, h good hh. ((tut sound)) Yeah, so u::::m is there anything else that you (.) want to know? [(Anything like that?)]
SP: [No, I can't think of nothing, no]
MS: Okay, (0.3) okay=
SP: =Okay
MS: I must've given you quite a lot of um (.) information. (0.8) Yeah, is there anything that you want me to clarify or:::
SP: No, I think I under, I understand [it, (nothing there)]
MS: [Okay]
MS: Okay, ·h yeah. So um::: (0.4) yeah (.) Thank you very much.
SP: Thank you.
MS: Yeah.
SP: Yeah.

(See Atkinson, M. and Heritage, J. (1984). (Eds.), *Structures of social action: Studies in conversation analysis*. Cambridge: Cambridge University Press, for standard conversation analysis transcription conventions from which the symbols are adapted.)

b) Speech samples segmented from recorded consultation

The recorded consultation transcribed above was segmented to remove the voice of the simulated patient and to preserve that of the medical student. Procedure was then followed to obtain speech samples categorised according to the consultation task (see Section 4.5.1).

No.	Task	Orthographic transcript
1	INI	And um they gave you some tablets
2	INI	Hello Mister Jones
3	INI	And, um, how long would that last you for
4	INI	I've been asked by your GP to come and um have uh speak to you; is that okay with you
5	INI	Hi, my name is (1-syllable name); I'm a fifth year medical student
6	INI	Okay, soon then you moved to, uh, here

7	INI	Okay, Ramipril, yeah
8	INI	Right, thanks, uh, right, so, um, can I ask you why that you come to see your doctor
9	INI	So, um, about three months ago you've seen your GP and they've found that you are having high blood pressure
10	INI	Do you know what's the previous reading for the blood pressure like for the first time
11	INI	So it's still a bit on the high side yeah
12	INI	So you've come in here to see the doctor to have a new measurement and see um
13	INI	Yeah, what uh, what's the name of the
14	GAT	And um any vegetables or meat that you eat yeah
15	GAT	And um I'm gonna ask you some questions about um your social history, like um we ask everyone, like whether you smoke or not
16	GAT	And allergic to anything
17	GAT	And um, are you still working at the mo…
18	GAT	And on average how many would you smoke a day
19	GAT	And um, right, so before we talk about the high blood pressure and the uh medications, um, may I ask something about your past medical history
20	GAT	But may I ask for the age please; how old are you
21	GAT	Do you eat a lot of salty food
22	GAT	Do you do a lot of exercises
23	GAT	Do you know about your cholestrol level; has it
24	GAT	Um, have you been to the hospital before for any
25	GAT	How about any other things that might run in the family you are aware of
26	GAT	How about the alcohol
27	GAT	How did you feel when you were on Ramipril
28	GAT	Is the working hour pretty long though
29	GAT	Yeah, and um, yeah well, just to uh, make sure I uh know what's your understanding of high blood pressure; do you, have you been told anything about…
30	GAT	Okay, no, no drug allergies, or
31	GAT	It's really good there, well done
32	GAT	Okay so you were not on any medications or, um any surgeries
33	GAT	Um, when did you start smoking
34	GAT	Yeah, sounds good
35	GAT	You gave up ten years ago
36	EXP	And if it's actually going back down normal, then they might actually take it off
37	EXP	And if it's still high, they might increase the dose or add another drug
38	EXP	And that's why um it is it will be a good idea to bring it back down with the medication
39	EXP	And there are also other things that can help you with getting off the drug let's say about you diet
40	EXP	Basically just um healthy meal and exercise

41	EXP	Because um like a diet with a lot of salts in seems to um increase the blood pressure
42	EXP	I guess if you speak to the doctor today, he might start you on um Ramipril as well, and um keep monitoring the blood pressure every now and then maybe a few months' time
43	EXP	So, but that's not a big problem
44	EXP	So um, actually if you, if you wouldn't mind trying to reduce the salt
45	EXP	So uh that if, if you, if your blood pressure is a bit high, then it increases the risk of stroke or other heart diseases
46	EXP	Because some, some people tend to find it helps uh when they walk a lot during the day, let's say uh go out for a walk during lunch, or
47	EXP	Right, so, um, um, let's talk a bit about your, um, the high blood pressure and the, um, the drug that you are taking, yeah
48	EXP	Yeah, so that sh-, that will definitely decrease the risk of stroke and other heart diseases
49	EXP	Um, well, you are quite right that high blood pressure might cause stroke and other things but it, it doesn't really cause it; maybe it's a risk factor
50	EXP	Yeah, just uh, well it doesn't everybody so, but if, if you realise, um you have a cough, then
51	EXP	You might, uh, doctor might change drug for you to something else, yeah
52	EXP	Yeah, well actually it's, it's quite common when, and, in the normal population
53	CLO	I must have given you quite a lot of information
54	CLO	Is there anything else that you want to know
55	CLO	So that's good
56	CLO	Thank you very much
57	CLO	Yeah is there anything that you want me to clarify or

c) Stimuli produced from selected and aggregated speech samples

The following stimuli were produced by selecting from the speech samples and aggregating the selection for each consultation task until reaching the required length. GAC and GAO were from the GAT samples, according to question types (open-ended vs. closed). The EXP samples were also further divided into EXP (providing explanation) and ADV (providing advice) (see Section 5.4.1 for details).

NN1-01INI

Hello Mister Jones. Hi, my name is X. I'm a fifth year medical student. I've been asked by your GP to come and um have uh speak to you; is that okay with you?

NN1-01GAC

Do you eat a lot of salty food? Do you do a lot of exercises? Is the working hour pretty long though? Um, when did you start smoking? Okay so you were not on any medications or, um any surgeries?

NN1-01GAO
How about any other things that might run in the family you are aware of? How about the alcohol? How did you feel when you were on Ramipril?

NN1-01EXP
So uh that if, if you, if your blood pressure is a bit high, then it increases the risk of stroke or other heart diseases.

NN1-01ADV
So um, actually if you, if you wouldn't mind trying to reduce the salt… And there are also other things that can help you with getting off the drug let's say about you diet.

NN1-01CLO
I must have given you quite a lot of information. Is there anything else that you want to know? Yeah is there anything that you want me to clarify or… So that's good. Thank you very much.

Appendix S. Transcript of voice stimuli

Part 1: 160 (10x16) stimuli

Sentences read with unprofessional attitudes

1. How are you?
2. How much alcohol do you drink in a week?
3. I'm sorry to hear that.
4. Do you have any thoughts on this?
5. Tell me more about that.

Sentences read with professional attitudes

6. How are you?
7. How much alcohol do you drink in a week?
8. I'm sorry to hear that.
9. Do you have any thoughts on this?
10. Tell me more about that.

Part 2: 96 stimuli (from simulated consultation)

NN1-01INI Hello Mister Jones. Hi, my name is X. I'm a fifth year medical student. I've been asked by your GP to come and um have uh speak to you; is that okay with you?

NN1-01GAC Do you eat a lot of salty food? Do you do a lot of exercises? Is the working hour pretty long though? Um, when did you start smoking? Okay so you were not on any medications or, um any surgeries?

NN1-01GAO How about any other things that might run in the family you are aware of? How about the alcohol? How did you feel when you were on Ramipril?

NN1-01EXP So uh that if, if you, if your blood pressure is a bit high, then it increases the risk of stroke or other heart diseases

NN1-01ADV So um, actually if you, if you wouldn't mind trying to reduce the salt… And there are also other things that can help you with getting off the drug let's say about you diet.

NN1-01CLO I must have given you quite a lot of information. Is there anything else that you want to know? Yeah is there anything that you want me to clarify or… So that's good. Thank you very much.

NN1-02INI Good afternoon, um Da…, is, is it Dan Jones? Okay, everything's gonna be confidential, and uh if you've got any concerns about anything we'll, we'll explain th't lat[er].

NN1-02GAC Have you got problem in your chest department? Is there anybody else in your family with blood pressure? [O]K, is there anything else that runs through the family or…? Okay, anything else?

NN1-02GAO Um, is briefly, is, did, do you, what do you know about blood pressure briefly? Uh do you mind telling me briefly what uh any ideas about your blood pressure…anything?

NN1-02EXP And then you've got other small issues, well they are still big issues that sometimes because the blood is running so fast, you can imagine if you are in the garden sprinkling your gar… your flowers

NN1-02ADV So the big thing here is actually the diet, and reduce the salt and exercise; those are the two alternative. You are still forty-eight; there's still a lot of things we can do to prevent a lot of things.

NN1-02CLO Is there anything that I've missed or you really want to tell me? And if you've got any more questions or any worries, you can always come back to the GP, confirm the alternative.

NN1-03INI Uh hello, good afternoon, uh my name's XXXXX, and I'm a second year medical student with the University of East Anglia. And the GP just asked me to have a, a word with you; is that all right.

NN1-03GAC Do you smoke? Okay, how, how much did you smoke? What made you quit? Okay, do you have any allergies? Do you take any other medication? Surgery for any reason? Have you ever been operated on?

NN1-03GAO Yeah absolutely; can we give a bit and talk about your family history of um any other medical conditions that you might have?

NN1-03EXP It is I'm afraid so, yes. Yeah you are clearly not overweight so you… you've got that risk factor out.

NN1-03ADV Well usually you have to take that for life. Uh, but details of that I think you should address them directly with the GP as well.

NN1-03CLO okay then, uh, I will, I will convey this information to your GP, and I'm sure he will manage your situation, he will address your concerns accordingly, and everything will be fine, I'm sure of it

NN1-04INI Uh good morning uh Mister Jones. Hi there; uh I'm one of the medical students. Just to remind you that everything's obviously confidential between um, us and the medical team here.

NN1-04GAC Um do you know anything about why you were on this particular tablet and not any others? Right not even walking round the Tescos and…?

NN1-04GAO Any parti-, why were you started on uh, high blood pressure? Um, so, what do you know of high blood pressure? Um how about drinking?

NN1-04EXP The analogy is, with high blood pressure, is like plumbing; it's like a set of pipes, um and if you put the pressure up too much in the pipes, then you know water bursts out' the sink, out' the taps

NN1-04ADV What we want to do is to control it at the early stage. You know, uh we can do a few tests just to make sure that you are on the right drug. That's, that's probably the best thing you can do.

NN1-04CLO And hopefully we can prevent these things from happening in the future. Thank you very much.

NS1-01INI Hello uh, is Mister Jones. I'm uh XXXX, I'm a fifth year medical stu[dent] Okay, um thank you for coming in today. Um I believe you come in to see you GP.

NS1-01GAC Have you had any medical problems in the past other than this high blood pressure? Okay have you had any operations at all? Okay, um and do you smoke at all or drink alcohol?

NS1-01GAO How has your weight been over the years? And how about exercise; do you… ? Um when you say that about um need to take if for the rest of your life, is it, have you got any thoughts about that?

NS1-01EXP It's, it's not um, it's not high as in it's gonna cause you a problem at the moment. But blood pressure is something that gradually affects the body um, over time, so it's over weeks and months and years.

NS1-01ADV So the other area really is exercise. Yeah um so there is um general sort of lifestyle changes which gonna all help with your blood pressure.

NS1-01CLO Right, okay well I'm gonna talk to your um GP um now and uh pass on this information then, um if that's okay with you.

NS1-02INI Hi there, my name is XXX, I'm a third year medical student at the University of East Anglia. Okay so just to reassure you that everything we discuss here is completely confidential; I may have to share with the GP afterwards if that's okay with you.

NS1-02GAC Have you got any other medical conditions yourself? And also do you smoke at all? Okay and how much were you smoking at that time? And when, when did you give up? And would you say you are under a lot of stress at work?

NS1-02GAO And um how about alcohol? And um how would you say your diet is? And what is your understanding of high, high blood pressure though and what it does? And how much exercise would you say you do at the mo[ment]?

NS1-02EXP Well salt can actually increase your blood pressure. Okay because stress again can cause your blood pressure to go up. So what these drugs do is they just uh allow those vessels to relax a bit.

NS1-02ADV Um yeah as you said uh, we will prescribe you the pills again but there are things such as uh you can improve your lifestyle through uh a healthy diet, doing more exercise.

NS1-02CLO Uh was there anything else at all? Okay, well thank you very much for coming in today; I'll pass your information on to the doctor. And I'll get him to come to see you then. Thank you very much. Bye.

NS1-03INI Hello, my name is XXXX. I am one of the medical students and I've been asked by the doctor to come to talk to you today. Will that be okay with you? Um as I say the doctor's asked me to come and talk to you and anything we do discuss is completely confidential of course.

NS1-03GAC And um, can I ask um, are you being treated for any other medical conditions at the moment? Um no diabetes or asthma or any...? And with the work that you're doing, do you find that stressful?

NS1-03GAO Okay um, are there any, apart from Ramipril which is something you're concerned about taking everyday or not, are there any other concerns that you have?

NS1-03EXP There are um, certain um, what we call risk factors that make us more prone to um, developing high blood pressure. Um sometimes that can cause headaches and things.

NS1-03ADV And I can um, put you in uh, contact with the dieticians up there…surgery, they can sort of give you some guidelines. And then in a couple of months time you can come back and we can see if that's made a difference or not.

NS1-03CLO Okay um yeah, thank you for coming in today and um. And as I say with the um diet and exercise and with the medication you should see a significant change in your blood pressure.

NS1-04INI Hi my name's XXX and I'm a second year medical student from University of East Anglia. Everything you say to us will remain confidential unless there's anything I'm concerned that I need to take to one of the GPs.

NS1-04GAC Have you ever had a heart attack? Um have you ever had a DVT or any clots? Um have you ever, do you have diabetes? Um is there a family history at all? So um do you smoke at the moment? And um do you drink?

NS1-04GAO Um what would you say your daily activities are like in a week, in a general week? We've got you on this Ramipril; um I don't know how much you understand about that?

NS1-04EXP It stops some of the um enzymes in your body from uh convert, creating what we call angiotensin, um which is a compound which can cause your blood pressure to um go up through um constriction of your arteries and vessels.

NS1-04ADV We do advise that you do take the Ramipril, because um it's like the most effective thing that we can offer you at the moment. But um it would be ideal if um you could maybe cut out with alcohol altogether.

NS1-04CLO So um, is there anything else you wanted to, do you feel that I answered your questions today? And um is there anything else that's um concerning you that you'd like to ask me bout? Okay, well, thank you.

NN2-01 INI And, uh, the GP actually just wants me to have a chat with you about um, about your high blood pressure and um, your medical history and things like that; is that going to be all right?

NN2-01GAC Is that full time? Have you got any siblings? Do you smoke? Okay and that's why you're here today? All right, um I just want to say how old are you please?

NN2-01GAO Hmm and what happened to him? And what about your, what about your mother? And um when, what happened when you were diagnosed with the high blood pressure? And what would you like to know today then?

NN2-01EXP Usually when somebody comes in and either because they've got some symptoms of high blood pressure, sometimes it's just headache sometimes, you know as you came in with a cough and…

NN2-01ADV Um so there's one thing you can do: exercise. Maybe once a week you can do that. Uh what else can you do; let's see; diet, um people talk about food and stuff so…

NN2-01CLO And I think that's it; I'm just trying to think if there's anything else about high blood pressure that would be useful for you to take away today. Well, right, thank you for talking to me.

NN2-02INI Hello Mister Jones. I'm uh XXX, one of the fifth year medical students. Um I've uh come here to ask you a few questions; that's all right?

NN2-02GAC Did we actually take a reading you while were on the Ramipril? So have you tried anything else apart from the Ramipril? Have you suffered from anything else with it?

NN2-02GAO I, I'd like to know um, um what you actually know about the medicine you have been prescribed? And how are you feeling since?

NN2-02EXP Obviously, there, there's quite a few different um choice of uh treatment for, for hypertension; this is just one of the things we could do.

NN2-02ADV Well, if you, if you, if you um uh did everything that you could in terms of your lifestyle, um that could help a lot.

NN2-02CLO Okay, okay, well I probably gonna set up another meeting so we can, we can uh… And uh, otherwise do you have any other questions? Thanks very much, thank you.

NN2-03INI Hello my name is XXXXX; I'm a second year medical student, at the University of East Anglia. And whatever you are going to say is going to remain confidential between you, me and the medical team.

NN2-03GAC And, um, can you, uh, tell me do you know anything about how the medication works, the one that you are taking

NN2-03GAO Um, so, um, how do you feel about having been diagnosed with high blood pressure

NN2-03EXP Um well, maybe um, well, one of the causes of hypertension is increased salt intake. Yes, it could lead to heart problems.

NN2-03ADV I'll just say, uh that if you have any more questions regarding how the stroke can occur and how the heart uh, conditions can occur, you can always ask the GP.

NN2-03CLO So basically I'll just summarise what you just said. Um, is there anything that I missed or you'd like to add? Uh, okay, well, thank you.

NN2-04INI Just been asked by your GP to come and talk to you today about why you've come to see him and, you know, all the things that involve that; um anything we do say today's completely confidential between us and your GP.

NN2-04GAC Um and smoking; have you smoked? Uh-huh, and did you use to smoke very much? Well, did you have your blood pressure done this morning? Um do you exercise much at all?

NN2-04GAO Do you mind if I ask about your diet at all? And also uh, what about your alcohol intake? So is there a particular thing that made you smoke or did you just sort[of]…

NN2-04EXP It's something that works to lower your uh, blood pressure; uh, you had your blood pressure done, you said it was a bit high; um so obviously there's a risk of having a stroke and heart attack.

NN2-04ADV Um see, again, people say this and doctor say this to patients that you've got to do some sort of activity and exercise. Because it, it is good for your health.

NN2-04CLO So was there anything else you wanted to know from um, today or anything I can help you with? Right, well thank you very much for talking me today. And hopefully you're sorted out soon.

NS2-01INI So um is it okay to address you as Mister Jones? Um and we are with the UEA and the GP said that you wouldn't mind talking to me at the moment… about some hypertension management you've been started, for blood pressure.

NS2-01GAC Do you exercise much? Do you have any family history of diabetes or anything like…? Do you have any history of high cholesterol in your family at all; have you had your cholesterol checked? Um do you smoke at the moment at all?

NS2-01GAO Um what about alcohol do you drink any…? Um, or any family history, you mentioned your dad had a stroke, any other history of…?

NS2-01EXP The pill is working, and you will probably see, I mean your blood pressure's quite high; we talked about taking your blood pressure more regularly, so that you can keep an eye on it, and work out when it's low enough.

NS2-01ADV The ideal thing would be for you to stay on the Ramipril at the moment which will help take your blood pressure down. Um you start on that for a while and you keep doing regular blood pressure checks.

NS2-01CLO Brilliant, uh, well we'll talk to GP; thank you for talking to me um, are there any questions you'd like to ask me? It's very nice to meet you.

NS2-02INI Hello Mister Jones. Hi I'm XXXX; I'm a medical student. Everything we talk about will um, of course be confidential and just between me and you and the GP if necessa[ry].

NS2-02GAC I don't know if, if you, if you do…ss.. if you smoke at all? How long were you smoking for? Do you, do you do a lot of exercise, you…? Um, and, have you got any allergies, any…?

NS2-02GAO Do you, what do you know about um high blood pressure, what…? And that uh, and you said you didn't have any symptoms before; how are you?

NS2-02EXP And we'd like to keep that low because um, there're certain parts of your body um, that, that don't just sort of, don't respond very well to having very high pressure through them; it can just be damaged, it can be you know as it

NS2-02ADV Yeah, if you take the drugs, if you just do the sort of lifestyle things I was talking about, you can reduce your blood pressure and therefore reduce your risk.

NS2-02CLO Do you feel, do you feel happy with that or…? I hope that's helped you. Okay and it's nice to meet you too, bye-bye. Okay well, thank you.

NS2-03INI Hello nice to meet you; my name's XXXX, and I'm a third year medical student at the University of East Anglia. I've been asked to come and have a chat with you today about your blood pressure; is that okay?

NS2-03GAC Um do you have any other medical conditions at all? Um is there anyone in your um family got any medical conditions? And um, are you on any other medications?

NS2-03GAO Um would you mind just telling me a bit about your his… your, um, your medical conditions in general? How do you feel about that? And what does that involve?

NS2-03EXP Like in a pressure cooker where everything's sort of compact and it's steaming up; that's sort of what's happening in your vessels.

NS2-03ADV Um well really, you can do lifestyle changes, so for example, you can do more exercise. The more enjoyable it is, the better it's for you.

NS2-03CLO Yeah, thank you and I will, I'll give you some leaflets to take them home. Hope you don't forget it; you could take this piece of paper if I make a little bit of note for you. Lovely to meet you. Well thank you very much.

NS2-04INI Hi, my name's XXX; I'm a fifth year medical student. So I've been told a little bit about your history. And I've been told that you're happy to have a chat with me today.

NS2-04GAC Do you exercise much? And, um, have you been unwell in the past at all? And you've had no funny turns where you've lost your ability to speak properly; your, your vision's gone a bit funny? No hospital admissions? Is that something that concerns you?

NS2-04GAO Do you have any ideas about what ca… causes your blood pressure to rise? What about your diet?

NS2-04EXP That's an ACE-inhibitor. There are other options; there's another three different classes of medicines, which, if you have a problem with the ACE-inhibitor, your doctor could choose from.

NS2-04ADV And you can take that home and read through it, and when you come back to the GP next week to you know start this management plan, you can have looked at it all through

NS2-04CLO Okay, so if I, I'll give you this hypertension, which is high blood pressure leaflet as well. Thank you for your time. Right, lovely to meet you.

Appendix T. Excerpt of participant workbook

The Effect of Speech Prosody on Attitudinal Perception in Medical Communication

PARTICIPANT WORKBOOK

Participant No. [Completed by researcher]

Thank you for taking part in this experiment. Please read carefully the following instructions, which will help you complete your task.

What do I need to do in the experiment?
We need you to listen to some voice samples (extracts from simulated medical consultations) and rate the attitude you perceive from the voice on a scale.

On your CD, there is a PowerPoint slideshow to present the voice samples to you. You need to click on the speaker icon to listen to the voice sample on each slide. There is a slide for each voice sample, so you need to do it slide by slide.

The rating scale is in the next part of this workbook. A page is prepared for each voice sample that you hear, and on the page there are 12 aspects to be rated about each sample. When you listen to a voice sample, please rate all 12 aspects on the scale. When you have finished the page and are ready for the next sample, please turn to the next page and click for the next slide.

How many times do I listen to each sample?
Please only listen to each sample once, unless you have not heard the voice sample clearly. Please use your first perception to complete the rating. As soon as you have heard and rated all the samples, you can give the completed workbook to the researcher and leave.

So you want to know my perception from the voice, but not from the content of what is said?
Correct – it should be the voice, not the content. For example, if you hear “It is a condition which can be very serious” said in a nice way, and are asked to rate:

The voice of the speaker sounds

Extremely unpleasant *Extremely pleasant*

Even though the sentence itself is likely to be about bad news, you might still feel that the voice is quite pleasant, and your rating should reflect this feeling.

How do I use the scale?
Just mark with a slash the point that represents the extent of your perception on the line. Either end of each line represents the greatest possible extent. The mid-point should be interpreted as “neutral” or “neither … nor …”. Please feel free to mark anywhere between the two ends as appropriate.

Please turn to the next page to start the experiment→

Voice sample 1

1 The voice of the speaker sounds

extremely unfriendly — extremely friendly

2 The voice of the speaker sounds

extremely disrespectful — extremely respectful

3 The voice of the speaker sounds

extremely uninterested — extremely interested

4 The voice of the speaker sounds

extremely judgmental — extremely non-judgmental

5 The voice of the speaker sounds

extremely unconfident — extremely confident

6 The voice of the speaker sounds

extremely discouraging — extremely encouraging

7 The voice of the speaker sounds

extremely insensitive — extremely sensitive

8 The voice of the speaker sounds

extremely unfeeling — extremely empathetic

9 The voice of the speaker sounds

extremely unpleasant — extremely pleasant

10 The voice of the speaker sounds

extremely unhelpful — extremely helpful

11 The voice of the speaker sounds

extremely inappropriate — extremely appropriate

12 The voice of the speaker sounds

extremely unsatisfactory — extremely satisfactory

Appendix U. Demographics of the perceptual study participants

Male students (*n* = 18), ***M*** (*SD*) age = **21.72** (3.43)

Participant ID	Staff/student	Gender	Age (years)
TM01	Student	M	21
TM02	student	M	20
TM03	student	M	18
TM04	student	M	22
TM05	student	M	21
TM06	student	M	21
TM07	student	M	21
TM08	student	M	21
TM09	student	M	21
TM10	student	M	20
TM11	student	M	20
TM12	student	M	20
TM13	student	M	21
TM14	student	M	18
TM15	student	M	21
TM-RM01	Student	M	32
TM-RM02	student	M	28
TM-RM03	student	M	25

Female students (*n* = 15), ***M*** (*SD*) age = **23.67** (8.08)

Participant ID	Staff/student	Gender	Age (years)
TF01	student	F	18
TF02	student	F	19
TF03	student	F	21
TF04	student	F	49
TF05	student	F	22
TF06	student	F	32
TF07	student	F	18
TF08	student	F	21
TF09	student	F	22
TF11	student	F	24
TF12	student	F	23
TF13	student	F	19
TF14	student	F	18
TF15	student	F	20
TF-RF01	student	F	29

Male staff (*n* = 8), ***M*** (*SD*) age = **33.88** (10.26) years

Participant ID	**Staff/student**	**Gender**	**Age (years)**
AM01	staff	M	31
AM02	staff	M	30
AM03	staff	M	45
AM04	staff	M	47
AM05	staff	M	25
AM06	staff	M	19
AM07	staff	M	44
AM08	staff	M	30

Female staff (*n* = 13), ***M*** (*SD*) age = **42.45** (10.59)

Participant ID	**Staff/student**	**Gender**	**Age (years)**
AF01	staff	F	31
AF02	Staff	F	35
AF03	staff	F	47
AF04	staff	F	49
AF05	staff	F	22
AF06	staff	F	53
AF07	Staff	F	Missing
AF08	staff	F	58
AF09	staff	F	46
AF10	staff	F	36
AF11	staff	F	Missing
AF12	staff	F	41
AF13	staff	F	49

Appendix V. SPSS output (instrument internal consistency of the perceptual study)

Inter-Item Correlation Matrix

	friendly	respectful	interested	non-judgmental	confident	encouraging	sensitive	empathetic	pleasant	helpful	appropriate	satisfactory
friendly	1.000	.788	.733	.691	.366	.747	.773	.767	.781	.721	.726	.758
respectful	.788	1.000	.725	.756	.368	.756	.782	.761	.777	.744	.800	.787
interested	.733	.725	1.000	.596	.516	.773	.753	.755	.734	.742	.723	.734
non-judgmental	.691	.756	.596	1.000	.321	.730	.741	.733	.731	.688	.724	.732
confident	.366	.368	.516	.321	1.000	.534	.355	.359	.419	.512	.483	.482
encouraging	.747	.756	.773	.730	.534	1.000	.827	.819	.834	.829	.812	.827
sensitive	.773	.782	.753	.741	.355	.827	1.000	.897	.847	.801	.815	.808
empathetic	.767	.761	.755	.733	.359	.819	.897	1.000	.842	.798	.778	.793
pleasant	.781	.777	.734	.731	.419	.834	.847	.842	1.000	.824	.836	.845
helpful	.721	.744	.742	.688	.512	.829	.801	.798	.824	1.000	.853	.859
appropriate	.726	.800	.723	.724	.483	.812	.815	.778	.836	.853	1.000	.900
satisfactory	.758	.787	.734	.732	.482	.827	.808	.793	.845	.859	.900	1.000

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
friendly	642.68	43223.678	.827	.727	.966
respectful	641.15	42905.967	.848	.769	.965
interested	644.58	42129.344	.819	.713	.966
non-judgmental	643.46	42609.010	.779	.668	.967
confident	640.14	45305.260	.481	.410	.974
encouraging	645.56	41525.429	.901	.821	.964
sensitive	646.23	41418.443	.891	.859	.964
empathetic	647.93	41299.599	.880	.846	.964
pleasant	643.80	41671.723	.898	.829	.964
helpful	645.19	41805.344	.887	.815	.964
appropriate	642.01	41894.254	.895	.858	.964
satisfactory	644.00	41869.542	.904	.861	.963

Appendix W. SPSS output (Hypothesis 4 of the perceptual study)

Between-Subjects Factors

		Value Label	N
Stimulus_type	3001	INI	54
	3002	GAC	54
	3003	GAO	54
	3004	EXP	54
	3005	ADV	54
	3006	CLO	54
Speaker_gender	1	Male	170
	2	Female	154
Speaker_English	1	native	164
	2	non-native	160
Speaker_training	1	lower	160
	2	higher	164

Levene's Test of Equality of Error Variances[a]

Dependent Variable:VAS_sum_correct_by_avg

F	df1	df2	Sig.
1.543	47	276	.018

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + Stimulus_type + Speaker_gender + Speaker_English + Speaker_training + Speaker_gender * Speaker_English + Speaker_English * Speaker_training + Stimulus_type * Speaker_English + Speaker_gender * Speaker_training + Stimulus_type * Speaker_gender + Stimulus_type * Speaker_training

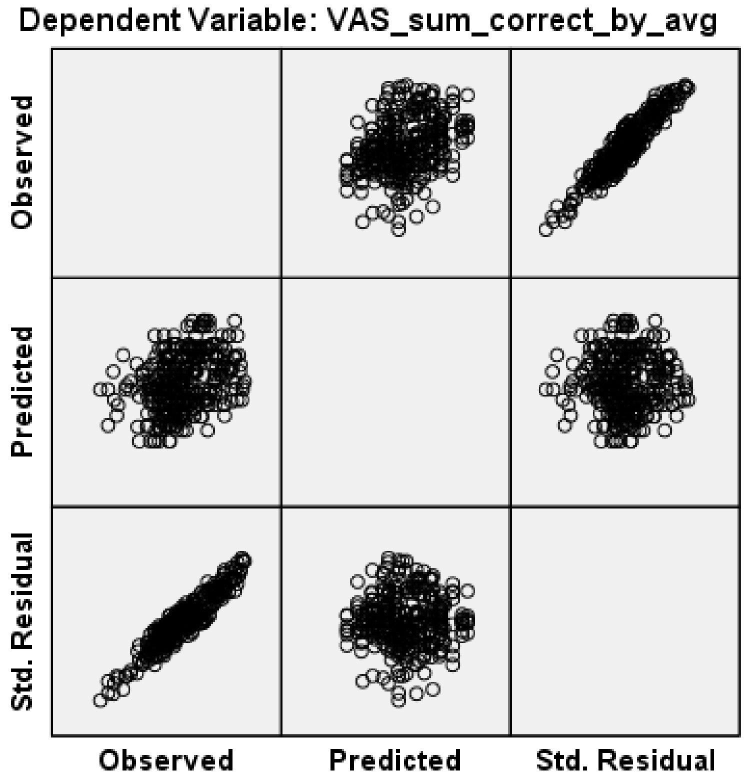
Tests of Between-Subjects Effects

Dependent Variable:VAS_sum_correct_by_avg

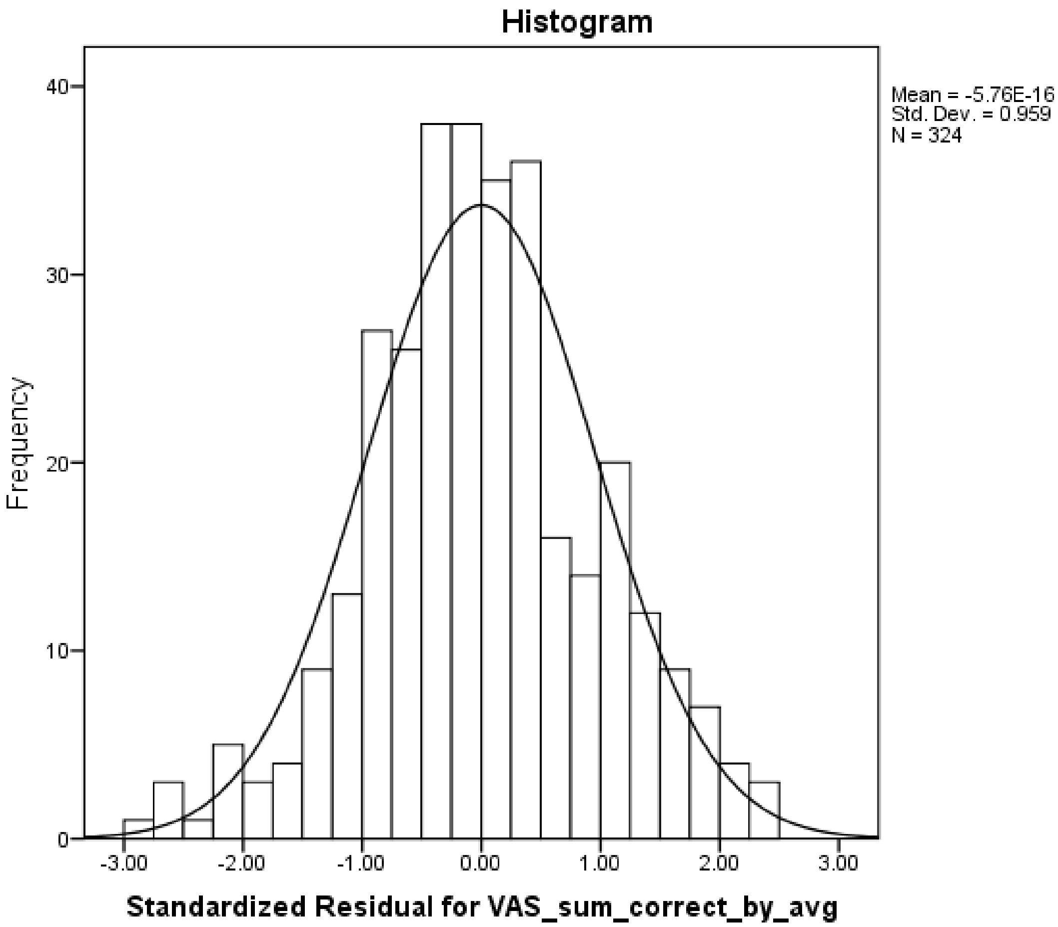
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	1.585E6	26	60958.875	1.473	.068	.114
Intercept	1.675E8	1	1.675E8	4046.725	.000	.932
Stimulus_type	631776.335	5	126355.267	3.053	.011	.049
Speaker_gender	36122.704	1	36122.704	.873	.351	.003
Speaker_English	280916.314	1	280916.314	6.788	.010	.022
Speaker_training	507569.737	1	507569.737	12.265	.001	.040
Speaker_gender * Speaker_English	587.455	1	587.455	.014	.905	.000
Speaker_English * Speaker_training	229.294	1	229.294	.006	.941	.000
Stimulus_type * Speaker_English	25277.883	5	5055.577	.122	.987	.002
Speaker_gender * Speaker_training	69.644	1	69.644	.002	.967	.000
Stimulus_type * Speaker_gender	83014.734	5	16602.947	.401	.848	.007
Stimulus_type * Speaker_training	41329.281	5	8265.856	.200	.962	.003
Error	1.229E7	297	41385.217			
Total	1.841E8	324				

Corrected Total	1.388E7	323				

a. R Squared = .114 (Adjusted R Squared = .037)



Model: Intercept + Stimulus_type + Speaker_gender + Speaker_English + Speaker_training + Speaker_gender * Speaker_English + Speaker_English * Speaker_training + Stimulus_type * Speaker_English + Speaker_gender * Speaker_training + Stimulus_type * Speaker_gender + Stimulus_type * Speaker_training


Histogram
Mean = -5.76E-16
Std. Dev. = 0.959
N = 324
Frequency
40
30
20
10
0
-3.00
-2.00
-1.00
0.00
1.00
2.00
3.00
Standardized Residual for VAS_sum_correct_by_avg

Appendix X. SPSS output (Hypothesis 5 of the perceptual study)

a) Native speakers

Tests of Between-Subjects Effects

Dependent Variable:VAS_sum_correct_by_avg

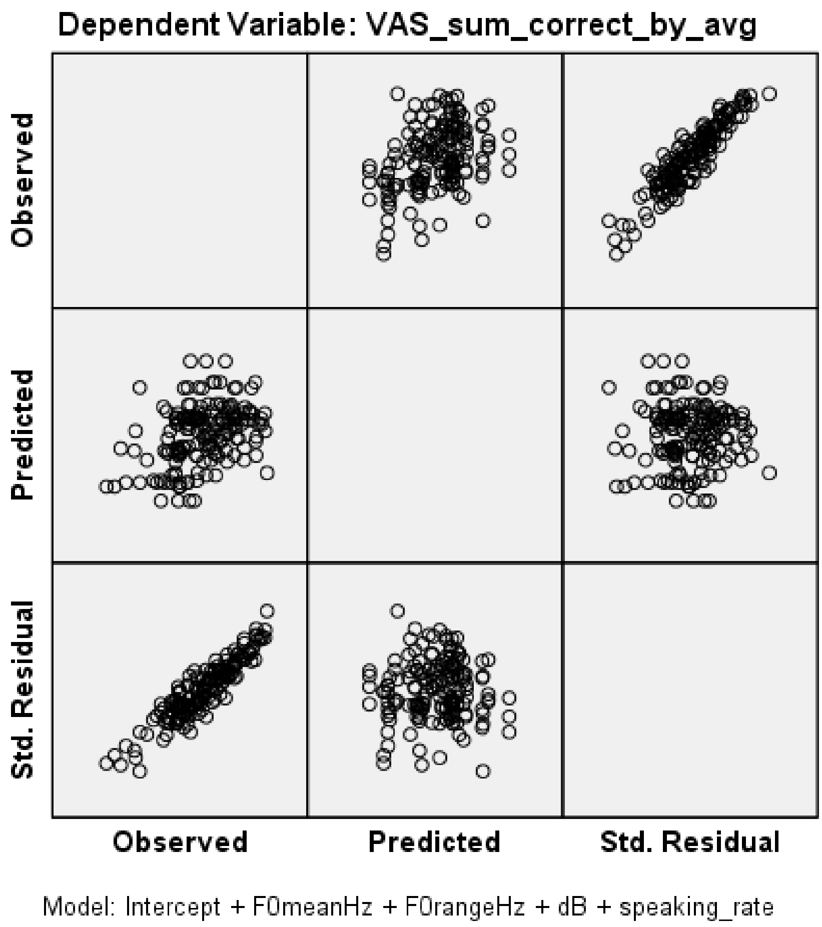
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	1.145E6	4	286342.893	6.950	.000	.149
Intercept	2213925.956	1	2213925.956	53.738	.000	.253
F0meanHz	34226.011	1	34226.011	.831	.363	.005
F0rangeHz	294918.943	1	294918.943	7.158	.008	.043
dB	963956.898	1	963956.898	23.398	.000	.128
speaking_rate	2407.250	1	2407.250	.058	.809	.000
Error	6550599.618	159	41198.740			
Total	1.008E8	164				
Corrected Total	7695971.189	163				

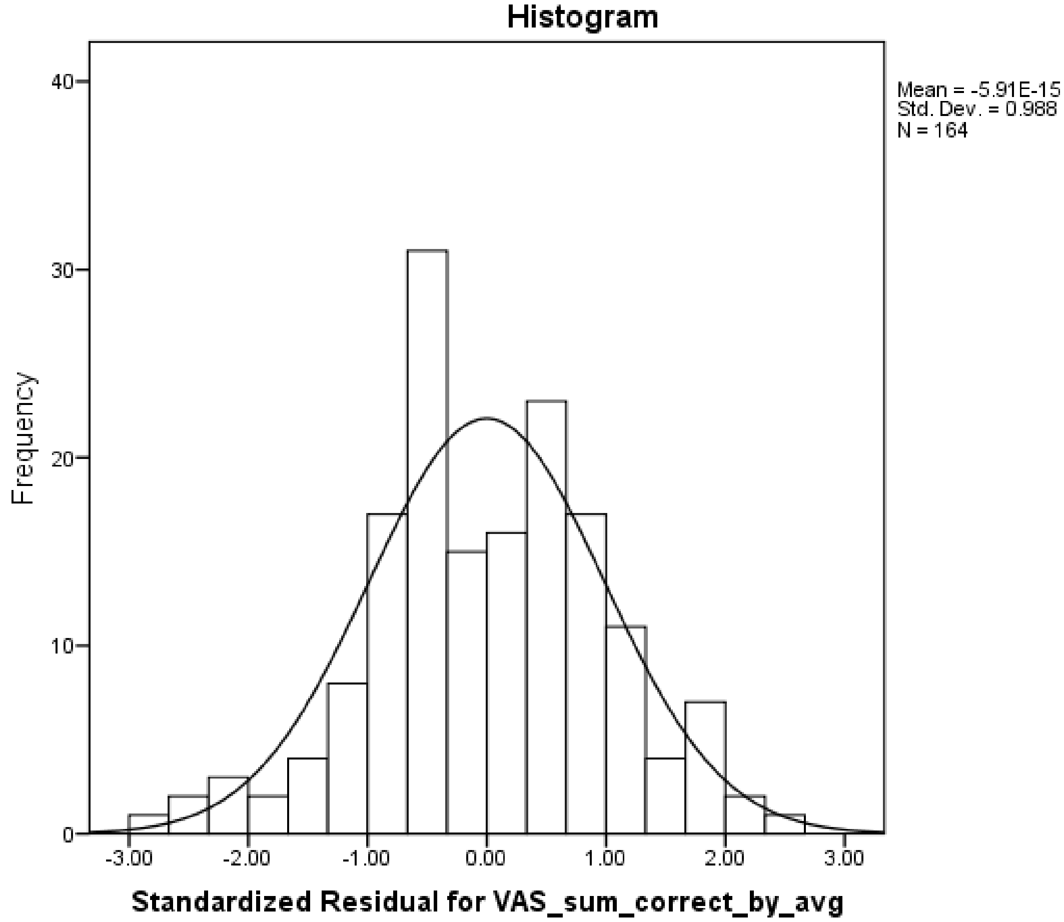
a. R Squared = .149 (Adjusted R Squared = .127)

Parameter Estimates

Dependent Variable:VAS_sum_correct_by_avg

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound	Partial Eta Squared
Intercept	2044.105	278.845	7.331	.000	1493.386	2594.824	.253
F0meanHz	-.712	.781	-.911	.363	-2.254	.831	.005
F0rangeHz	1.315	.491	2.676	.008	.344	2.285	.043
dB	-23.935	4.948	-4.837	.000	-33.708	-14.163	.128
speaking_rate	4.470	18.491	.242	.809	-32.050	40.989	.000


Dependent Variable: VAS_sum_correct_by_avg
Observed
Predicted
Std. Residual
Observed
Predicted
Std. Residual
Model: Intercept + F0meanHz + F0rangeHz + dB + speaking_rate


Histogram
Mean = -5.91E-15
Std. Dev. = 0.988
N = 164
Frequency
0
10
20
30
40
-3.00
-2.00
-1.00
0.00
1.00
2.00
3.00
Standardized Residual for VAS_sum_correct_by_avg

b) Non-native speakers

Tests of Between-Subjects Effects

Dependent Variable:VAS_sum_correct_by_avg

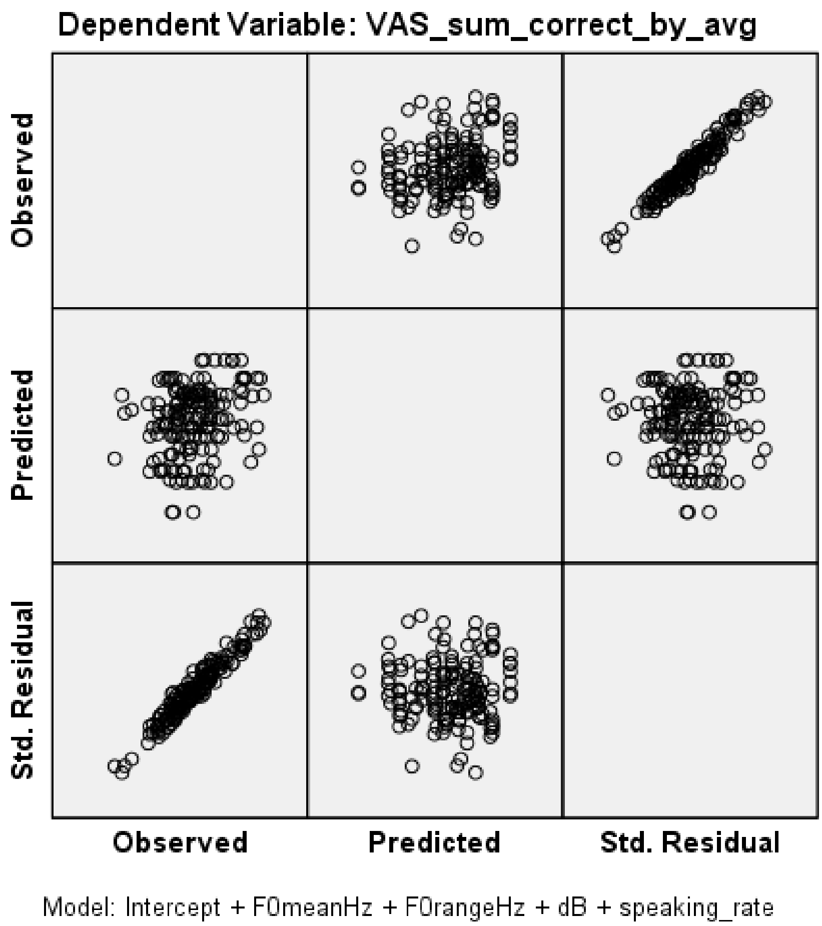
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	366914.405[a]	4	91728.601	2.565	.040	.062
Intercept	116645.228	1	116645.228	3.261	.073	.021
F0meanHz	33723.316	1	33723.316	.943	.333	.006
F0rangeHz	30348.076	1	30348.076	.849	.358	.005
dB	13888.719	1	13888.719	.388	.534	.002
speaking_rate	295900.024	1	295900.024	8.274	.005	.051
Error	5543495.570	155	35764.488			
Total	8.333E7	160				
Corrected Total	5910409.975	159				

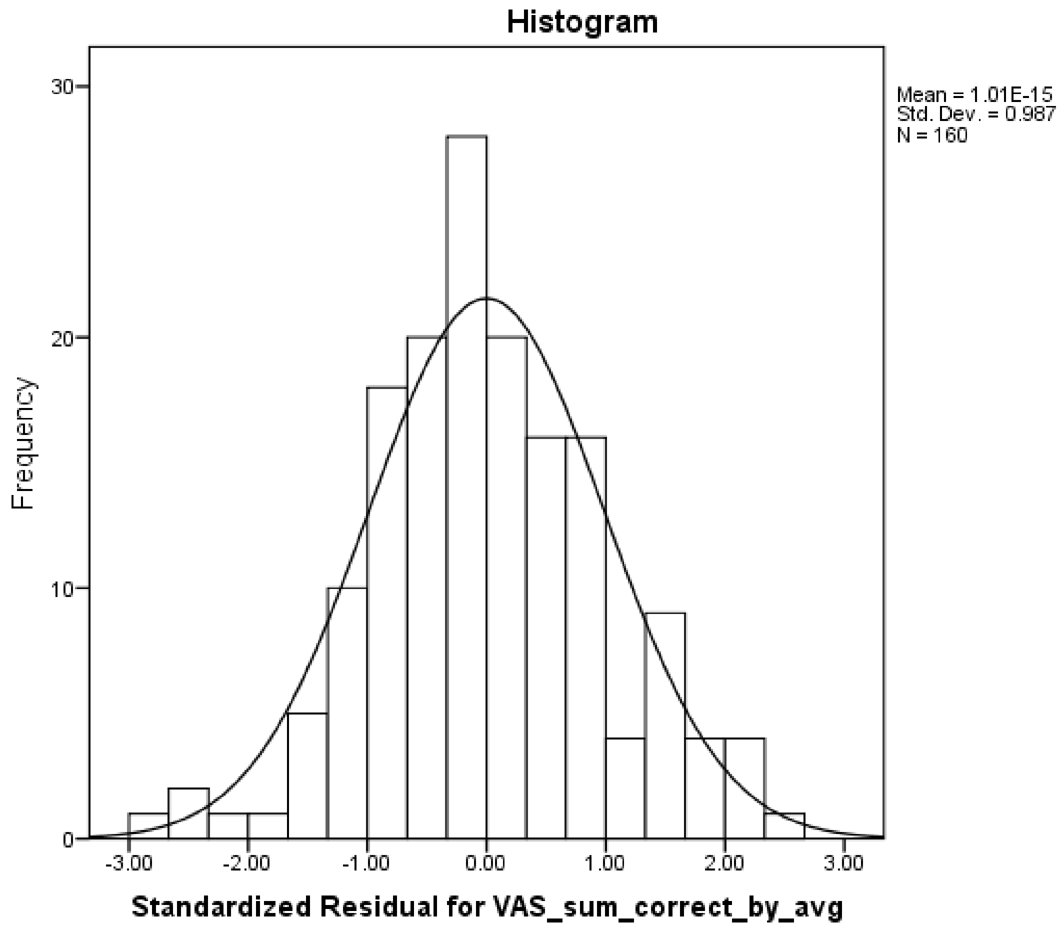
a. R Squared = .062 (Adjusted R Squared = .038)

Parameter Estimates

Dependent Variable:VAS_sum_correct_by_avg

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound	Partial Eta Squared
Intercept	587.733	325.441	1.806	.073	-55.139	1230.605	.021
F0meanHz	.516	.531	.971	.333	-.533	1.565	.006
F0rangeHz	-.329	.357	-.921	.358	-1.033	.376	.005
dB	-3.385	5.432	-.623	.534	-14.116	7.345	.002
speaking_rate	59.602	20.721	2.876	.005	18.670	100.534	.051


Dependent Variable: VAS_sum_correct_by_avg
Observed
Predicted
Std. Residual
Observed
Predicted
Std. Residual
Model: Intercept + F0meanHz + F0rangeHz + dB + speaking_rate



Histogram
Mean = 1.01E-15
Std. Dev. = 0.987
N = 160
Frequency
30
20
10
0
-3.00
-2.00
-1.00
0.00
1.00
2.00
3.00
Standardized Residual for VAS_sum_correct_by_avg