

"Checking up" on Medical Training Programs and
Assessments for Foreign Trained Doctors through
Multimodal Analysis

by

Katharina Klassen

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in
partial fulfillment of the requirements for the degree of

Master of Arts

in

Applied Linguistics and Discourse Studies

Carleton University
Ottawa, Ontario

© 2019
Katharina Klassen

Abstract

Studies have investigated the culturally-bound characteristics of active listening across several disciplines, including psychology, business and conflict mediation (Lamiani, et al., 2008). Active listening is also valuable in multicultural and multilingual medical consultations as it improves doctor-patient relationships within a patient-centred care model of practice (Vogel et al., 2018). However, there is a dearth of research regarding the extent to which individual behaviours pertaining to active listening are present in clinical interactions. The present study evaluates multimodal active listening performances of non-native English-speaking medical doctors during Objective Structured Clinical Examinations against ideal models of active listening behaviours. Results indicate that non-native speakers' active listening behaviours differed from the baseline study in a number of verbal and non-verbal areas, the ramifications of which could impact perceptions of doctors' indifference regarding patients' health experience. Explanations for the findings and research and pedagogical applications are offered.

Keywords: active listening, multimodality, patient-centred care model, cross-cultural communication, Objective Structured Clinical Examinations (OSCEs)

Acknowledgements

This thesis project would not have been possible without the amazing support of many individuals both near and far.

Firstly, I would like to give a very heartfelt thank you to Dr. Michael Rodgers for supervising this project. Not only did you guide me through major setbacks in my project, but you encouraged me to challenge myself in terms of my writing and my professional development, beyond what I thought I was capable of doing. Thank you for believing in me. In addition, my deepest thanks go out to my review committee members, Dr. Timothy Wood and Dr. Natasha Artemeva, for taking the time to consider my work and offer insightful feedback.

I also need to give my sincerest thanks to Dr. Geoff Pinchbeck for connecting me with Ms. Lake when my original study project was no longer accessible. Of course, this leads into a great appreciation for Ms. Lake, my project partner at AIMGA and C2I in Alberta, as well as the participants at these facilities for being kind enough to share their preparatory videos with me. This kindness was also extended by the Research Department at the Medical Council of Canada, and I am extremely grateful for their interest in and support of my research as well.

My overall ability to thrive at Carleton University would not have been possible without the incredible support of the ALDS administrative staff and faculty members, from whom I was able to grow academically throughout a number of classes prior to beginning this project. Your passion for the field of Applied Linguistics is continually inspiring, and I feel blessed to have been a part of this community for the past two years. This spirit also extends to my dear ALDS cohort friends, who have acted as both commiserates and sources of good humour in the often isolating struggle of completing a thesis project.

Last, but certainly not least, I would never have finished this degree, much less this project, without my fantastic family back home in Manitoba as well as my partner, Colin, here in Ottawa. Each of you are keenly aware of the personal impact that this undertaking has had on me, and I will always be grateful that you pushed me to move forward through every low point. I love you all very, very much.

Table of Contents

Abstract	i
Acknowledgements	ii
Chapter 1: Introduction	1
1.1 A history of westernized healthcare.....	1
1.2 Bridging the ontological gap: Introducing the patient-centred care model of medicine.....	5
1.3 Organization of thesis	7
Chapter 2: Literature Review	9
2.1 Multimodal communication.....	9
2.2 The multimodality of active listening	14
2.3 Applying multimodal listening: Active listening across various domains of discourse	21
2.4 Multimodal listening in healthcare: Realizing the gap between theory and practice.....	27
2.5 Multimodal listening in healthcare applied: The Canadian clinical context.....	33
Chapter 3: Methodology	43
3.1 Introduction.....	43
3.2 Developing an active listening framework	43
3.3 Video analysis tool selection.....	45
3.4 Methodology phase 1: Ethics and recruitment.....	47
3.5 Methodology Phase 2.....	51
3.5.1 Methodology phase 2, part A: Establishing a pilot study	51
3.5.2 Methodology phase 2, part B: Designing the analytical framework through a pilot study.....	52
3.5.3 Methodology phase 2, part C: Challenges with and revisions of the analytical framework.....	72
3.5.4 Methodology phase 2, part D: Third-party feedback regarding the piloted analytical framework	73
3.6 Methodology phase 3: Implementing the revised framework and data collection workflow	74
Chapter 4: Results and Analysis	78
4.1 Introduction.....	78
4.2 Research question 1: Analysing NS medical candidates' and NNS medical candidates' performances of individual active and anti-active listening traits.....	78
4.2.1 Establishing assumptions and statistical approach for research question 1	78
4.2.2 Performances of non-verbal active and anti-active listening traits	82
4.2.3 Performances of verbal active and anti-active listening traits.....	85
4.3 Research question 2: A macro level analysis of NS medical doctors' and NNS medical doctors' performances of active and anti-active listening.....	91
4.3.1 Establishing assumptions and the statistical approach for research question 2	91

4.3.2	Macro level active and anti-active listening results and analysis.....	92
4.4	Research question 3: Realizing the relationship between active listening performance and OSCE success	101
Chapter 5: Discussion	103
5.1	Introduction.....	103
5.2	Responding to research question 1: Addressing the character of active listening on a micro level.....	103
5.2.1	Non-verbal modes of active and anti-active listening.....	103
5.2.2	Verbal modes of active and anti-active listening	109
5.3	Responding to research question 2: Addressing the character of active listening on a macro level.....	114
5.4	Responding to research question 3: Active listening in medical pedagogy for NNS, foreign-trained doctors.....	120
Chapter 6: Concluding Remarks	125
6.1	Study limitations and areas of future research	126
References	129
Appendices	148
Appendix A	Illustration of common MCCQE Part II errors.....	148
Appendix B	MCCQE Part II scoring scheme	150
Appendix C	Ethics clearance	152
Appendix D	List of illustrative control group video codes and corresponding video length and content.....	154
Appendix E	List of non-illustrative control group video codes and corresponding video length and content.....	155
Appendix F	List of test group video codes and corresponding video length and content.....	156
Appendix G	Results for individual active and anti-active verbal and non-verbal listening traits	158
Appendix H	Descriptive statistics for positive listening set	163
Appendix I	Descriptive statistics for negative listening set	163
Appendix J	Descriptive statistics for beats	163
Appendix K	Descriptive statistics for eye contact frequency	163
Appendix L	Descriptive statistics for eye contact duration	164
Appendix M	General reviewer commentary following mock OSCEs performed by test group	165

List of Tables

Table 1. Coding schema for ALNV, AALNV, ALV and AALV tiers	70
Table 2. Non-verbal active listening traits identified in the test group, illustrative control group and non-illustrative control group.....	83
Table 3. Non-verbal anti-active listening traits identified in the test group, illustrative control group and non-illustrative control group	84
Table 4. Verbal active listening traits identified in the test group, illustrative control group and non-illustrative control group.....	86
Table 5. Verbal anti-active listening traits identified in the test group, illustrative control group and non-illustrative control group.....	88
Table 6. A summary of positive (active) listening traits across both control groups and the test group	93
Table 7. A summary of negative (anti-active) listening traits across both control groups and the test group	93
Table 8. Dunn’s post hoc test results for positive set.....	94
Table 9. Dunn’s post hoc test results for negative set.....	95
Table 10. Dunn’s post hoc test results for beats hand gestures.....	97
Table 11. Dunn’s post hoc test results for eye contact frequency.....	98
Table 12. Dunn’s post hoc test results for eye contact duration	99
Table 13. Descriptive results of reviewers’ commentary regarding test group participants’ ability to communicate effectively with patients and provide accurate diagnoses.....	102

List of Figures

Figure 1. MCCQE part II categorical weighting.....	39
Figure 2. Transferring audio, video and framework files into ELAN	53
Figure 3. Creating tier archetypes	54
Figure 4. Programming a useable tier framework.....	56
Figure 5. Installation of the tier framework (segmentation mode)	58
Figure 6. Snapshot of overall tier framework in segmentation mode with annotations on tiers.....	58
Figure 7. Example of an annotation (i.e. “Hey Henry, come on in, have a seat”) on a tier with timestamp indication	59
Figure 8. Snapshot of all “ALNV” tiers in transcription mode, including timestamps	63
Figure 9. Illustration of an OSCE retrieved from a pilot study video (seated from left to right: medical candidate, acting patient, and examiner).....	71

Glossary of Key Terms and Acronyms

AIMGA: Alberta International Medical Graduates Association; project partner.

C2I: Communicate2Integrate; project partner.

ELAN: open-source software used to create time-aligned annotations from audiovisual recordings developed by the Max Planck Institute.

MCCQE Part 1: first part of a qualifying examination administered by the MCC; computer-based; includes a series of questions designed to test and apply candidates' medical knowledge.

MCCQE Part II: second part of a qualifying examination administered by the MCC; includes a series of stations or OSCEs that test candidates' clinical skills, professionalism and communication skills.

Medical Council of Canada (MCC): organization that oversees the assessment and evaluation of medical candidates across Canada and is tasked with granting licentiates so that successful medical candidates can practice in Canada.

Mode: a communication channel that is recognizable by a community of speakers (e.g. verbal, written, etc.).

Multimodality: theory that explores how people interact and communicate with each other through various, coinciding modes.

NNS: non-native English speakers; primarily test group participants in this study.

NS: native English speakers; primarily control group participants in this study.

Objective Structured Clinical Examinations (OSCEs): standardized stations used to assess medical candidates and designed to imitate real-life clinical scenarios (e.g. history taking, breaking bad news, etc.) with the use of actors as patients.

Patient-Centred Care Model (PCCM): approach to healthcare that focuses on exploring the illness experience from the patient's perspective in order to improve clinical interactions as well as increase accurate, initial diagnoses.

Chapter 1: Introduction

“Medicine is always a child of its time” (McWhinney, 2014, p. 18).

1.1 A history of westernized healthcare

The basis of many modern-day healthcare practices finds its roots in Greek antiquity. For early ancient Greeks, one’s health was the culmination of the well-being of one’s mind and spirit even more than the well-being of one’s body (Amundsen & Ferngren, 1983). Illness, then, was not only the absence of spiritual and mental well-being but also the maleficence of the gods. Therefore, healing came by way of superstitious “cures” (that were in fact fatal) or “medicines” suspect of the placebo effect. For example, amulets could be held against fractures and dislocations as a means of fixing the injured bone (Stafford, 2005, p. 123). The buildings which constituted hospitals in ancient Greece were also intimately linked to Grecian spirituality. Individuals would seek healing in one of five sites, situated in Epidaurus, Athens, Corinth, Kos and Pergamum, which were composed of temples for supplication and thanksgiving to the healing god, Asclepius, as well as shared sleeping quarters that were frequented by snakes, which were thought to be associated with Asclepius (Barefoot, 2005, p. 206; Stafford, 2005, p. 121). The doctors at these sites were tasked primarily with interpreting patients’ dreams, which were thought to be indications of cures from Asclepius, rather than physically adjusting ailments or injuries (Barefoot, 2005, p. 206). Overall, healthcare in Greek antiquity was comprised of non-standardized practices and was not a respected field.

By c. 460 BC, what would become the foundation of the modern-day healthcare institution began to take shape. Hippocrates, who later became known as the father of medicine,

produced a number of texts which rationalized that rather than appeasing a god's wrath, achieving good health could be accomplished through dietetics and the internal balance of bodily fluids (King, 2005, p. 152). As his understanding of the human body was rudimentary, he also understood that it was critical to install a standardized routine for medical interactions in order to develop a deeper relationship with the patient and ultimately to elicit more detailed information from the patient regarding their illness experience. Thus, in all encounters, Hippocrates stressed the importance of simply doing no harm to the patient. While it seems apparent by today's medical standards, this philosophy was revolutionary to the ancient Greek healthcare system in which patients distrusted so-called physicians and were often more likely to live if they did not seek out such practitioners. On this basis of trust, Hippocrates further enforced the practice of fully understanding each patient's unique situation in terms of food intake and lifestyle before prescribing a remedy, which crucially installed the tradition of diagnosis before prognosis (King, 2005, p. 152). The Greek morpheme "dia-", found also in contemporary English words such as "dialogue", emphasises the input of all parties in a conversation while the root "-gnosis", such as in "(g)nostology", refers to knowledge, thus forming a combined meaning which suggests that knowledge is gained through collaboration and the sharing of experiences rather than a singular "expert" opinion. Thus, Hippocrates' medical approach not only rejected the notion that illness was somehow divinely sent but also recognized the key role that the social and physical environment could play in the patient's experience of health and well-being.

With the rise of the Roman Empire, medical practice once again slipped into disfavour as many popular scholars, including Pliny the Elder, propagated the belief that it was unfair for Greek doctors to charge fees for their services (Amundsen & Ferngren, 1983). In addition, the exploration of the human body and attempts at new cures throughout the Roman Empire only

served to increase public fear of medical practitioners. For example, fatal bloodletting and starvation were common methods of attempting to rebalance bodily fluids (Amundsen & Ferngren, 1983). Thus, in order to escape the blame of such blatant medical failures, physicians reintroduced the divine into healthcare by suggesting that their prognoses were the “will of the gods”. As the Middle Ages were ushered in, this reintroduction of the role of divinity was further merged with medical care as healing largely took place in Christian institutions such as monasteries and overall the patient’s agency in the creation of their own treatment plan decreased dramatically (Montserrat, 2005).

In congruence with the European Enlightenment, the 17th century saw the return of the Hippocratic tradition of recording cases in order to better understand recovery and death as well as the rebirth of the modern medical interaction. In practice, this return was largely facilitated by the scholar Sydenham, who instigated a systematic bedside observation routine that sought to link symptoms with the progression of disease (McWhinney, 2014, p. 20). This routine was supported by paradigmatic shifts instilled by Bacon’s philosophy of man controlling nature (McWhinney, 2014, p. 17) as well as Descartes’ proposition of a separation of mind and matter, with preference being given to the mind (McWhinney, 2014, p. 19). This work was continued particularly by French physicians after the French Revolution (McWhinney, 2014, p. 21). With studies of cadavers using Laennec instruments, complemented by the identification of causal agents of disease by Pasteur and Koch, the medical tradition, as it is known today, of linking cause with reaction was cemented. Indeed, by the end of the 19th century the medical practice and the discipline of philosophy had become completely dissociated.

Since the 19th century, the medical field has rested upon two main models, with each gaining or losing emphasis in synchrony with shifting social paradigms and beliefs. The first is

known as the ontological model and portrays disease as "...an entity located in the body and conceptually separable from the sick person" (McWhinney, 2014, p. 21). It is associated with the conventional or academic clinical method and has received preference over more descriptive, natural clinical methods such as those proposed within the second medical paradigm: the physiological or ecological approach. In the ecological or physiological model, disease is thought to result from an "...imbalance within the organism, and between organism and environment; individual diseases have no real existence, the names being simply clusters of observations used by physicians as a guide to prognosis and recovery" (McWhinney, 2014, p. 22). One of the leading arguments against this approach is that the onslaught of information from the patient overburdens the physician emotionally, which may inhibit their ability to make difficult decisions or to work effectively. This may be accompanied by years of "bottling up" emotions as physicians may not have adequate time or appropriate outlets to relieve these emotions. The ontological or conventional clinical model, on the other hand, allows the physician to barricade themselves from the emotions of their patients, creating a calm, in-control disposition and reinforcing the superiority of the physician's clinical expertise and authority while at the same time subordinating the knowledge of the patient and relinquishing the power they hold regarding their own health. While this may be an effective means to arriving at a satisfactory treatment, it often leads to a loss of trust in the physician and a belief that the physician does not care for the patient's overall well-being. Balint, Hunt, Joyce, Marinker and Woodcock (1970) recognized that physicians should acknowledge their own emotions as well as those of their patients through listening and providing "overall" diagnoses as opposed to more discrete, dissociated diagnoses. This approach further evolved to include the interdisciplinary nature of the clinical process, highlighting the roles of social and psychologic data alongside biologic data (Stewart et al.,

2014, p. 7). Thus, garnering the lived experience of suffering through techniques such as active listening may not be in the best interest of the physician. Still, while the preference for the ontological model allows practitioners to more efficiently and systematically assess patients, its exclusion of the ecological model is highly problematic in that it also diminishes the agency of the patient. McWhinney describes this gap: “so successful was the model that its weaknesses only became apparent much later, as its abstractions became further and further removed from the experience of the patient. No abstraction is ever a complete picture of what it represents: it becomes less and less complete as levels of abstraction and power of generalization increase...[t]he danger is that we forget that our abstraction is not synonymous with the real world” (2014, p. 24).

1.2 Bridging the ontological gap: Introducing the patient-centred care model of medicine

In order to combat this omnipresent stress of perceived objectivity, which isolated the illness from the patient, medical scholars and practitioners began to take a more patient-centred approach (Balint et al., 1970; Hopkins & Balint Society, 1972). In the 1980s, this approach was introduced as the patient-centred care model (PCCM) (Stewart, Brown, McWilliam, Freeman, & Weston, 2014). The PCCM was originally viewed as a soft science, especially in contrast to the illness-centred approach within the ontological paradigm, due to its focus on humanitarian care of the patient (Stewart et al., 2014). Today, this model pivots on a balance of objectivity and subjectivity which is notable through its four components. The first component encompasses the exploration of health, disease and the illness experience through shared narrative building, including a realization of feelings, ideas, expectations and perceived meanings of the health and illness experiences (Stewart et al., 2014, p. 7). The second component suggests an understanding

of the patient based upon their individual traits and their socioeconomic, familial and environmental context (Stewart et al., 2014, p. 7). Thirdly, the PCCM emphasises finding common ground on the basis of problems, priorities, goals of treatment, and patient and doctor roles (Stewart et al., 2014, p. 7). Finally, the PCCM stresses that the patient-doctor relationship can always be enhanced through awareness of empathy, power, hope, introspection, transference and countertransference (Stewart et al., 2014, p. 7). The benefits of PCCM have been observed using several qualitative and quantitative studies (cf. Little et al., 2001a & 2001b; Epstein et al., 2005). For example, some benefits may include, but are not limited to, improved interactions between patients and physicians, particularly through positive perceptions of doctors' empathy and caring as rated by patients, and reduced costs for diagnostic tests and use of similar services due to an increase of more accurate, initial diagnoses by the doctor (Stewart et al., 2014, pp. 12-13). Boyle, Dwinnell and Platt (2005) further the PCCM methodology through the use of "patient-centered listening". This involves using a technique called "invite, listen and summarize" in order to create a comfortable space for the patient to share his or her health concerns (as cited in Stewart et al., 2014, p. 29). Overall, the PCCM incorporates a holistic view of the patient's illness experience through an emphasis on empathetic listening on the part of the doctor in order to arrive at more succinct treatment plans and to strengthen doctor-patient relations.

An alternative version of the PCCM is the social skills model (SSM). Argyle (1983) describes the SSM for healthcare practitioners in terms of its verbal and non-verbal communicative features as well as training and application methods. For example, one of the verbal features of the SSM is the expression of emotions and feedback signals, which must be augmented with non-verbal behaviours such as gaze and posture. These verbal and non-verbal

features are also embedded within a specific genre, within which dialogues should be sequenced in a familiar and effective manner based on the client's background and goals. Training medical candidates to use the SSM can be accomplished in a variety of ways, including supervised role-playing, modelling conducted by supervisors, video playback and on-the-job training. In summary, both the SSM and the PCCM encourage a move away from the ontological model, which aims to abstract the disease from the patient experience, in order to give agency to the patient through their experience of health and illness. As McWhinney states, "if we are to recapture our capacity to heal, we will have to transcend the literal-mindedness that seems to follow when we become prisoners of our abstractions. A new clinical method should find room for the exercise of imagination and for restoring the balance between thinking and feeling" (2014, p. 28).

1.3 Organization of thesis

The present study is organized into 6 chapters. Chapter 2 provides an in-depth discussion overviewing the theoretical background of active listening and multimodal analysis as well as the evolution of medical care. This chapter then explores the role of active listening across disciplines, focusing specifically on how active listening can be applied within a more patient-centred medical care model globally and within the Canadian medical context. Subsequently, Chapter 3 delineates the methodological process used in the present study, elaborating upon the analytical framework developed and used within the software system selected for this study. Chapter 4 presents the results of the two phases of the present study in response to the guiding research questions, in addition to a discussion of the results in relation to literature pertaining to active listening in medical pedagogy and practice. Following this, Chapter 5 presents a critical interpretation of the study findings and how they relate to the characterization of active listening

as well as the implications for medical trainees and training organizers. Finally, Chapter 6 consists of concluding remarks, featuring policy implications, study limitations and recommendations for future research.

Chapter 2: Literature Review

2.1 Multimodal communication

In a medical interview, the patient's experience of illness and health are expressed not only through the patient's words but also through their non-verbal and paralinguistic behaviours (such as eye contact and intonation patterns respectively) which must then be interpreted and responded to by the doctor using similar modes of expression in order to convey an attitude of empathy. Thus, an understanding of a wide variety of behaviours not limited to the verbal mode are essential to creating a holistic view of any communicative event and any communicative event is inherently a multimodal act. Traditionally, communicative events were analysed from a strictly verbal perspective through the theoretical framework of communicative competence. The concept of communicative competence was developed by Dell Hymes (1974) and takes into account "real-life" factors of language performance and proficiency, such as social context, beyond merely cognitive perfection. Verbal communication, on a basic level, was marked as a system of arbitrary signs, which could be categorized and used differently cross-culturally.

As the model of communicative competence grew, the roles of pragmalinguistics and sociopragmatics also developed within this framework (Bachman, 1990; Leech, 1983; Thomas, 1983). Pragmalinguistics is understood as "the study of the particular resources which a given language provides for conveying particular illocutions" (Leech, 1983, p. 11). Sociopragmatics, on the other hand, involves "the conditions that constrain and govern speakers' language use. Specifically, it focuses on how different social variables such as power or status, social distance, and rank of imposition or severity of offense...affect language use" (Beltrán-Palanques & Querol-Julián, 2018, p. 81). The concepts of pragmalinguistics and sociopragmatics are mutually dependent. An interlocutor's assessment of the sociopragmatics of a given communicative

context, including politeness and cultural norms, is encoded in their pragmalinguistic choices (Roever, 2011, p. 464). In general, the development of these competencies allows speakers to recognize the social factors guiding language use and choice and further promotes successful communication in a wide variety of communicative environments (Beltrán-Palanques & Querol-Julián, 2018).

Until recently, the communicative competence model still focused solely on the verbal mode and was summarised within the following sub-competencies: discourse, which examines the appropriate choice and sequence of words and sentences in a given verbal or written text; sociocultural, which recognizes the speaker's knowledge of what is or isn't appropriate within the context of the communicative event; linguistic, which portrays elements such as morphological, phonological and sentence constituents necessary to realize communication; actional, which suggests an ability to understand communicative intent; and strategic, which highlights an awareness of all available communication strategies and a speaker's proficiency in using them (Celce-Murcia, Dornyei, & Thurrell, 1995). However, scholars have argued that these pragmatic structures are evident in all modes of communication; thus, all should be taken into account with respect to pragmatic competence. In fact, Beltrán-Palanques and Querol-Julián (2018) go so far as to state the term "multimodal pragmatic competence" should replace the unilateral "pragmatic competence", which is negligent of the non-verbal and extra-linguistic components that are involved in a communicative event (p. 83). Although the non-verbal competencies have only been recognised and included within the interactional competence stream (of which actional competence is one of three components, alongside conversational and non-verbal/paralinguistic competencies [Beltrán-Palanques & Querol-Julián, 2018]) and the field of multimodality itself was essentially formulated in the 1990s, the conceptualization and study

of multimodality it not novel (Camiciottoli & Bonsignori, 2016, p. 2). In the late 19th century, Darwin published a text, which has been contested since its release, depicting the patterns of human and animal facial expressions in order to understand the evolution and universality of facial posturing in relation to emotion (Ekman, 2006). Additionally, in the later 20th century, Efron (1941) studied Jewish and Italian immigrant gestures using a naturalistic method of observation that is still in place today (as cited in Camiciottoli & Bonsignori, 2016, p. 2). Further studies illustrate the role of spatial and body positioning, kinesics, paralanguage and other non-verbal aspects in communicative events (Camiciottoli & Bonsignori, 2016, p. 2).

In contemporary literature, multimodality is attributed to three foundational schools of linguistic thought: social semiotics, which is associated with Halliday's systemic functional linguistics (1985) and the research of Kress and van Leeuwen (2001); discourse analysis, which is linked to the systemic functional theory established by O'Halloran (2008; 2011); and multimodal interactional analysis, as developed by Norris (2004) and Scollon and Scollon (2003), which combines interactional sociolinguistics, intercultural communication and multimodal semiotics. More generally, multimodality is defined as "an approach used to understand the contribution of various semiotic resources (e.g., verbal, visual, aural, spatial) in studies of communication" (Fortanet-Gómez & Camiciottoli, 2015, p. 1). Tang (2013) further characterises multimodality as a series of phases, termed "the multimodal instantiation hierarchy", which progresses in complexity from the general system of language, which includes every possible meaning that can be created from a combination of semiotic modes to the receivers' interpretations of the communicative event.

The conveyance of meaning is distributed across all modes used in a given communicative act and then interpreted by the recipient of the act using their knowledge of the

multimodal instantiation levels (as described above). In some communicative events, different modes may carry more or less bearing on the interpretation of meaning and thus are classified as “primary” or “secondary” accordingly (Fortanet-Gómez & Camiciottoli, 2015, p. 4). In other events, all modes may be indispensable in terms of understanding the intended meaning of the communicative action. Thus, an understanding of the interrelated roles of modes within the wider sociocultural context is critical. Beltrán-Palanques (2016) stresses that “if a language user understands the sociopragmatic norms of a particular language but lacks the pragmalinguistic resources, he/she will not succeed in the communicative event and a pragmatic failure may occur. Hence, the areas of pragmalinguistics and sociopragmatics are critical during communication and are of special interest for SL/FL [second language/foreign language]” (p. 102). In summary, participating actively and effectively in a conversation goes beyond understanding the individual words and phrases that are being verbally expressed; participants must be able to recognise and react to non-verbal and paralinguistic features of the communicative event in order to fully comprehend and respond appropriately to a given situation.

A number of studies have attempted to illustrate the non-verbal components of communicative events. Originally, Kendon classified these behaviours as gestures which could be grouped into phrases and understood as “units of visible bodily action identified by kinesic features which correspond to meaningful units of action” (2004, p. 108) or more generally “part of what the person is trying to say” (1996, p. 8). Knapp (1972) also depicted seven broad categories of non-verbal communication: kinesics or body language (e.g. facial expressions), paralanguage or context-free vocalizations (e.g. interruptions), physical contact (e.g. handshakes), proxemics (e.g. personal space), artifacts or adornments (e.g. clothing), physical

characteristics (e.g. perceived ethnicity), and environmental factors (e.g. spatial design) (as cited in Campoy-Cubillo, 2016, p. 27). In more recent scholarship, a critical component of non-verbal research is founded upon the theory of gesticulation developed by McNeill (2005). Gesticulation is defined as

motion that embodies a meaning relatable to the accompanying speech. Gesticulation is by far the most frequent type of gesture in daily use, and it covers many variants and usages. It is made chiefly with the arms and hands but is not restricted to these body parts- the head can take over as a kind of third hand if the anatomical hands are immobilized or otherwise engaged, and the legs and feet too can move in a gesture mode. (McNeill, 2005, p. 5)

The composition of a gesture then is twofold: it not only carries meaning, but it also is co-expressive to the speech with which it is synchronized (McNeill, 2005, p. 22). McNeill (2005) stresses that although the speech and gesture are co-expressive, they are not redundant, as each will present a single, underlying idea in different ways (p. 22). This presentation of a unified idea through distinct channels is critical. The union of gesture with speech is symmetrical (Kelly, Barr, Church, & Lynch, 1999) such that information that is later recalled by an interlocutor will only be remembered as speech (McNeill 2005, p. 26). Indeed, this tight bonding of gesture with speech is observed to overcome barriers in communication such as stuttering and loss of fluency. McNeill states that “tight binding clearly fosters an imagery-language dialectic by creating unbreakable psycholinguistic units within which it can take place” (2005, p. 27).

In addition to McNeill’s model, several other models have been proposed as means of categorizing and illustrating micro-level non-verbal behaviours. Ekman and Friesen (1972) code non-verbal behaviours according to their emotive value using FACS, a Facial Action Coding

System. In this model the “smallest visually discriminable facial movements” are defined as “action units” (as cited in Cohn, Ambadar, & Ekman, 2007, p. 203). These action units can be collectively utilised and classified as an event or “a set of action units that overlap in time (i.e. co-occur) and appear to define a perpetually meaningful unit of facial action” (as cited in Cohn et al., 2007, p. 211). Furthermore, Campoy-Cubillo explains non-verbal behaviours according to the intention or goal of the interlocutor, such that that these behaviours can act as manipulators, adapters or pacifiers (2016, p. 19). Finally, Patterson (1983; 1988) further specifies seven broad functions of non-verbal behaviours: providing information, regulating interaction, conveying intimacy, facilitating a goal or service, demonstrating social control, managing emotion, and self-portrayal or presentation (as cited in Campoy-Cubillo, 2016, p. 30).

While a handful of non-verbal behaviours such as deictic and metaphoric gestures tend to maintain their functional integrity regardless of the communicative setting, the majority of non-verbal behaviours are not permanently associated with a specific function or meaning (this will be discussed further below). The function of a non-verbal behaviour will vary depending on the context. Indeed, even in his discussion of the inseparability of speech and gesture, McNeill acknowledges that binding is not inevitable, especially when it is impacted by external factors (2005, p. 27).

2.2 The multimodality of active listening

Although most multimodal study is focused on the interlocutor in the position of “speaker”, the role of the interlocutor as the listener also constitutes a multimodal act. From a physiological perspective, sound, which is created when there is sufficiently quick movement within a medium such as air, carries four main properties: amplitude (resulting in loudness), frequency (resulting in pitch), spectrum (differentiating timbre) and phase (allowing for the

localization of sound) (Carpenter & Reddi, 2012). As sound enters the auditory canal through the external ear, the sound waves carry via vibrations through the eardrum and into the fluids of the inner ear. The primary structure of the inner ear, the cochlea, transduces these mechanical vibrations into nervous energy and is responsive to different sound frequencies (Carpenter & Reddi, 2012). Finally, these transformed sounds move into the brain via auditory fibres (Carpenter & Reddi, 2012).

When the sound, in energy form, reaches the brain, it activates a mechanism known as the speech sound map node. From this node, sound is sent to the auditory feedback controller where it is differentiated from other incoming sounds (sidetones) not relevant to the current dialogue. This “controller” is actually a chain of activity, as sound travels from the cochlea to the auditory cortex via brain stem nuclei and parts of the thalamus, which are then represented in an auditory state map and are further projected onto a target model of this map (provided there are no errors in the auditory chain along the way) (Guenther, 2016). The regular functioning of the auditory feedback controller is affected by environmental manipulations, such as shifts in pitch or loudness.

As this depiction of hearing suggests, listening in general must move beyond simply receiving sound input in order to actually engage with the interpretation of the sounds cognitively. In the field of applied linguistics, listening has received numerous definitions. For example, there are specific linguistic features that are characteristic of listening, which have been variously conceptualized as minimal responses (Fellego, 1995), backchanneling (Yngve, 1970; White, 1989), assent terms (Schegloff, 1982), feedback (Allen & Guy, 1974), accompaniment signals (Kendon 1967; 1996) and hearer signals (Bublitz 1988). However, it is agreed that listening is a response or reaction to a stimulus (i.e. a return of communication based upon the

output of the speaker) (Campoy-Cubillo, 2016, p. 22). Thus, listening serves a number of pragmatic functions (Nation & Newton, 2009; Weir & Vidaković, 2013). The primary purpose of listening is for general speech comprehension, or simply as a means of receiving verbal input from the speaker for the value of the speech alone (Campoy-Cubillo, 2016, p. 23). In second or foreign language acquisition classrooms, this is the most commonly assessed variant of listening. Learners are tested on their ability to accurately recognise sounds and words, perhaps with some inference from the context of the communicative event (Campoy-Cubillo, 2016, p. 23). The testing of higher levels of listening proficiency, however, require the individual to also think critically about the information they are receiving (Campoy-Cubillo, 2016, p. 28). In these circumstances, the listener is presented with the task of decoding the meaning of the information within a social dimension and must provide a thoughtful, immediate response that conveys a sense of deeper reflection, empathy and judgement (Campoy-Cubillo, 2016, p. 28).

Therefore, in critical thinking, relatively equal importance is assigned to the components of inner reflection and the surrounding context; however, some forms of listening require a greater emphasis on the role of context, namely the interpretations of the face value and emotions of the speaker. This is particularly the case when demonstrations of empathy are necessitated by the situation. Depraz (2001) suggests that there are four stages of this “lived empathy”: a passive association of one’s body to another’s; an imaginative self-transposal from one’s perspective to another; an understanding that one’s viewpoint may be alien to another’s; and an ethical responsibility to another person (as cited in Calloway-Thomas, 2010, p. 13). Calloway-Thomas (2010) underscores the role of empathy in such scenarios, stating that “the totality of human beings, including their life ways, customs, and habits of mind, all fall within the purview of empathy and can be used to further human understanding” (p. 13). Thus, this form of listening,

known as empathetic listening, can be defined as a "...complex listening type in the sense that the person is also expected to evaluate the emotions of another person who will demand an immediate response, as in medical professions... The listener in empathetic listening is also required to be supportive, and for this he or she will need to interpret not only what other people say, but also what they do, their facial expressions, changes in intonation in relation to what they say, the gestures they use when saying something, etc." (Campoy-Cubillo, 2016, p. 28). Indeed, Calloway-Thomas further supports this by arguing that "at any point in our relationship with others, we can use cultural data to embellish an argument, offer an example, clarify a point, or make ourselves accessible or inaccessible to others- depending on our proclivity toward empathy and to enmeshment with others" (2010, p. 18). The interpretation of these cues is instantaneous and must be responded to just as instantly by the listener lest the speaker infer that the listener is not acutely invested in the conversation (Campoy-Cubillo, 2016, p. 28). As this element of conversation is highly emotive and demonstrates a high level of trust and vulnerability on the part of the speaker, some elements of the speaker's communicative acts may be in contradiction in order to preserve their own face and create a safe space of self-preservation (e.g. when an individual says "I'm fine" even though they are wringing their hands), placing an even greater impetus on the listener to engage in the conversation effectively (Campoy-Cubillo, 2016, p 29).

As empathetic listening clearly demonstrates then, the act of listening itself does not occur in a vacuum; it requires an analysis of many modes of communication at once. In other words, in its natural, face-to-face occurrence, it requires a multimodal act. As Campoy-Cubillo explains "...multimodal listening skills may be considered a complex construct composed of a number of micro-skills that entail being able to understand and interpret the inner connection of several modes, as well as being able to provide a response to all or some of the information

modes” (2016, p. 22). At the macro level, the multimodal listening skillset includes acts of verbal and non-verbal communication (Campoy-Cubillo, 2016, p. 14). It follows then that the complexity of the listening task will depend upon the density of the information being processed, which may include the quantity of verbal and non-verbal input, as well as the message abundance, which includes the number of modes at play in the communicative event and their individual saliency (Campoy-Cubillo, 2016, p. 29).

Realizing the cohesion of verbal and non-verbal behaviours, rather than simply understanding their individual, microlevel roles, is key to performing effective multimodal listening. Campoy-Cubillo further stresses that social competence is a critical factor in achieving this balance of interpreting verbal and non-verbal behaviours, stating that “this message regulation and the ability to become skilled in controlling adequate non-verbal expression in an interaction are part of the process of mastering the social roles of encoding and decoding multimodal messages” (2016, p. 25). Philippot, Feldman and McGee (1992) also demonstrate that non-verbal communication plays a large part in conversational turn-taking, particularly when the communicative participant is in the role of the listener. For example, the shifting gaze of the listener could signal the beginning or end of the speaker’s turn. Loehr (2004) further identifies seven pragmatic functions that the verbal-non-verbal cohesion produces: completeness, incompleteness, information status, focus, emphasis, contrast and visual status of discourse (as cited in Campoy-Cubillo, 2016, p. 20). These are reinforced by the findings of Egolf (2012), who argues that the contrast between verbal and non-verbal communication appears across nine main dimensions: primitiveness, continuity, analog/digital, tense (in the sense of time), emotional/informational, reference to non-existent entities, brain hemisphere use, metacommunication, and complementary/contradictory. Thus, the symbiotic nature of verbal and

non-verbal cues within a communicative event cannot be understated. In the case of listening, the interpretation of both within a given social frame is essential to understanding all the levels of meaning implied by the speaker.

Beyond this cohesion, the verbal and non-verbal modes, as suggested earlier, must be interpreted within the message abundance and sociolinguistic context of the communicative event itself. Not all verbal and non-verbal cues hold the same meaning or function in every setting. On a macro level, these cues may be culture-bound (Scollon, Diener, Oishi, & Biswas-Diener, 2004; 2005), which is noted especially in the case of facial expressions (Friesen, 1972; Kita, 2009). Errors that result in culturally-based discrimination or conflict are a result of “things going wrong” in terms of an individual’s cultural competence. Cultural competence goes beyond basic sociopragmatic competence in that it is an individual’s total knowledge of the beliefs, information and skills needed in order to participate in society as well as the ability to select, from the totality of grammatically correct expressions available, forms that appropriately reflect the social norms defining behavior in specific happenstances. To some extent, miscommunication is also caused by errors of linguistic competence, which is an individual’s understanding of a language’s phonology, morphosyntax, lexicon, etc. These issues of linguistic error may be addressed using interactional sociolinguist Dell Hymes’ SPEAKING model (1974), which highlights the environment, code, tone, etc. of the communicative event. Overall then, each communicator brings not only linguistic knowledge to a conversation but also pragmatic performance capabilities, in addition to his or her learned experiences of culture. As a result, when one interlocutor is a native speaker of the language being used while the other is a second or foreign language learner, not only will the two communicators have different cultural backgrounds (differing degrees of cultural competence), but they will also encounter

impediments in the areas of phonology, lexicology, and semantics (linguistic competence).

Communicators who are unaware of the great number of sociocultural factors at play in a given interaction may then fall prey to accepting prejudices and enacting behaviours associated with assumed hierarchies of power. Consequently, the relationship between language and culture is not one-way; societal worldviews, beliefs and practices impact linguistic choices while at the same time linguistic choices perpetuate societal values and perspectives.

It is evident that sociocultural context plays a critical role in active listening. However, in addition to this context, individuals make idiosyncratic choices within the given norms of communication of a society or subculture within the society (Larina, 2015). As Royce (2007) states, “these meanings and choices, realized in differing modes, will necessarily have a message and interactional-focus, and will draw on the textual or compositional conventions appropriate to the mode” (as cited in Beltrán-Palanques, 2016, p. 101). For this reason, some scholars have argued that attempting to teach appropriate non-verbal cues is not useful, as practiced responses can seem artificial and cannot account for idiosyncrasies of speech nor the spontaneity of everyday conversation (Campoy-Cubillo, 2016, p. 24). The exception to this argument is in the case of professional settings in which pragmatic routines are relatively stable and standardized within a given society. Thus, as Campoy-Cubillo (2016) argues, “what is needed is a detailed analysis of non-verbal features and how they interact among each other and with verbal features in a particular situation...because non-verbal communication is mainly ‘situation-based’ or ‘situation-activated’” (p. 17). Although individuals cannot prepare for every communicative event, nor can they correctly interpret all meanings without fail, increased training and practice of multimodal listening scenarios would greatly improve accuracy of interpretation of communicative messages, particularly for second or foreign language learners (Campoy-Cubillo,

2016, p. 25). Calvo, Gutiérrez-García, Fernández-Martín and Nummenmaa (2014) also argue that the ability to repeatedly recognise non-verbal communication results in faster interpretations of conversational intent and messages, particularly in terms of recognising the meaning of facial expressions cross-culturally. In addition, this form of “interactive listening” or “participatory listening” creates a safe, secure environment for students to take risks and engage in problem solving (Vandergrift, 1997; Clementi, 2006). As a result, this model may shift the current methods of teaching listening, especially in the second and foreign language classrooms (Rost, 2011) and enhance the quality of listeners’ responses, including boosting their performance of active listening (Weger, Castle & Emmet, 2010). In summary, a multimodal construct of listening enables interlocutors, particularly within a multilingual or multicultural communicative event, to better understand and prepare for the multifaceted nature and process of listening that can be produced in with a variety of modes (Campoy-Cubillo, 2016, p. 30).

2.3 Applying multimodal listening: Active listening across various domains of discourse

As mentioned above, some communicative events can place more or less emphasis on different modes. For example, when a client is at a supermarket checkout conversing with the cashier about the total cost of their groceries, minimal effort is required in regard to the non-verbal mode, as the interlocutors are focused on exchanging information primarily in the verbal mode. In these scenarios, an appearance of a lower levels of non-verbal engagement with the other interlocutor is required because the conversation does not hinge on the interlocutors’ ability to interpret deeper, emotional meaning through the non-verbal mode. However, many other interactions may require a higher level of engagement across all communicative modes in order to convey such underlying beliefs and emotions held by the interlocutors. This active

engagement is essential in garnering information beyond what is conveyed verbally. It is arguable whether a listener can give the appearance of being engaged in a dialogue by providing the appropriate verbal and non-verbal cues, without actually being fully engaged in the conversation. However, evidence of the listener's lack of engagement may appear in their inability to ask for clarification, reiterate or paraphrase the speaker's claims, or effectively empathise with the speaker. In this situation, the speaker may or may not interpret the other individual as listening or as being engaged in the conversation, depending on the proficiency of that individual to mimic the feedback cues of active listening. However, Rogers and Farson (1988) argue that it is most likely that the speaker will interpret the listener's actions as "empty and inauthentic".

The alternative to this lack of engagement is to have interlocutors play a more active role in the conversation, particularly in regard to their listening-related actions. Although it was introduced within the applied linguistics framework above, this concept of "active" listening was initially developed within the Humanistic Theory in the late 1950s (Kourmoussi, Amanaki, Tzavara, & Koutras, 2017, p. 113). Active listening was originally defined as "giving free and undivided attention to the speaker", and is characterized through a number of verbal and non-verbal communicative cues (Rogers & Farson, 1988, as cited in Robertson, 2005, 1053). Although it is difficult to determine the physiological evidence in the brain to demonstrate that a listener is in fact actively engaged in the conversation, some studies have suggested that because the brain is a social organ that grows through interaction and experience, active listening can have positive implications for interlocutors from a neuropsychology standpoint (Cozolino, 2006). By increasing the degree of active engagement with a speaker, the listener demonstrates higher levels of psychological attachment and bonding with the speaker, which augments some

functioning and, in some cases, size of areas of various parts of the brain (Chugani, 2004) as well as increases the firing of premotor cortex mirror neurons, which are responsible for the ability to mirror the verbal and non-verbal behaviours of others (Iacoboni & Dapretto, 2006; Iacoboni, 2008). In a study conducted by Bodie, Vickery, Cannava and Jones (2015), 301 undergraduate students were randomly assigned to the role of problem discloser (i.e. speaker) or listener. Forty listeners were trained in active listening behaviours while a control group of 130 student listeners were not trained before engaging in short dialogues with a discloser. The videotaped interactions were coded for non-verbal immediacy cues (e.g. eye contact) and verbal behaviours (e.g. paraphrasing) and analysed using multiple regression. The results suggested that those verbal and non-verbal responses that responded specifically to the problem (as opposed to more generic feedback cues) tended to increase emotional awareness and improve affect.

In another study, Castro, Kluger and Itzhakov (2015) found that active listening correlated positively with speakers' psychological safety or willingness to disclose more private information to listeners than if the listener did not display active listening characteristics. In addition, the findings of a study of counsellor trainees completed by Levitt (2001) suggested that active listening not only increased the listeners' soft skills but also decreased their experiences of anxiety while the speaker disclosed highly emotional information. Similarly, Egger (2015) examined the differences in interprofessional communication (IPC) skills and self-efficacy (an individual's belief in their innate ability to achieve goals) between native and non-native English-speaking nursing students. Students attending various baccalaureate programs certified by the Accreditation Commission for Education of Nursing and/or the Commission on Collegiate Nursing Education in the United States were asked to complete three electronic surveys that collected data using an IPC Scale and a General Self-Efficacy Scale. The results showed that

both native and non-native English-speaking students rated themselves equally highly in terms of their IPC skills and self-efficacy. Furthermore, those with more professional healthcare experience, as well as senior students, had higher IPC ratings. These findings support other studies which have similarly shown that higher levels of bonding and attachment between interlocutors correlate with greater interpersonal communication skills (Klaus & Kennell, 1972; Karen, 1998; Levine & Heller, 2010). Condon and Ogston (1966), prior to the development of more modern brain scan technologies, have also shown that interlocutors demonstrate synchronous behaviours in dialogue, although the verbal and non-verbal communication may not exactly “mirror” that of the other interlocutor. Thus, an individual who is participating actively in the conversation not only through non-verbal cues such as backchannels but also through verbal cues such as requests for clarification demonstrates high levels of empathy and social bonding with the other interlocutor and the speaker will be fully aware that the listener is in fact interested in listening to them (Jones, Bodie, & Hughes, 2016). In summary, demonstrating active listening improves not only the efficiency with which information is exchanged but also has the potential to develop the emotional aspect of the relationship between the interlocutors.

Although its characterization and implications from both a neurological and psychological standpoint is clear, the precise realization of active listening is somewhat less straightforward and tends to shift based upon the social standards of the context and users, particularly when the setting is multicultural or multilingual (as will be addressed further below). Those operating in the domain of business stress the importance of what is termed as “active empathetic listening”, which is defined as “a form of listening practiced by salespeople in which traditional active listening is combined with empathy to achieve a higher form of listening” and is believed to be “at the heart of good communication” (Drollinger, Comer & Warrington, 2006,

pp. 161-162). Drollinger et al. (2006) suggest that there are three stages of active empathetic listening. These include sensing, which refers to “a listener’s attending to all of the explicit and implicit information expressed by the other person”, processing, which involves “synthesizing and remembering information in order to enable the construction of a narrative whole”, and responding, which includes “clarification and use of verbal and nonverbal means to indicate attention” (as cited in Kourmoussi et al., 2017, p. 114). Other researchers in the field of business have underscored the improvement of mutual understanding through active listening (Mineyama, Tstutsumi, Takao, Nishiuchi & Kawakami, 2007, p. 81).

The incorporation of active listening is critical in several intercultural conflict training (ICT) theories. Conflict face-negotiation theory developed by Ting-Toomey (1988) and Ting-Toomey and Kurogi (1998) encompasses seven key assumptions pertaining to the implications of both self-oriented or other-oriented and horizontal or vertical facework on dimensions of large or small power distances and individualism or collectivism. Most notable to this research is the seventh assumption, which states that “intercultural facework competence refers to the optimal integration of knowledge, mindfulness, and communication skills in managing vulnerable identity-based interaction scenes appropriately, effectively, and adaptively” (Ting-Toomey, 2007, p. 258). While knowledge informs notions of culture-based facework, this understanding is derived from mindfulness, which is defined as “attending to one’s internal assumptions, cognitions, and emotions and at the same time, becoming attuned to the other’s conflict assumptions, cognitions, and emotions” (Ting-Toomey, 2007, p. 259). Arriving at such a state of self-reflexivity and cultural relativism is achieved primarily through “a deep state of listening without judgement” or, more generally, active listening (Ting-Toomey, 2007, p. 259).

In another pivotal ICT theory developed by Hammer (2001) and Rogan and Hammer (2002), crisis negotiation is essential. Crisis negotiation is defined as “a unique form of conflict interaction in which law enforcement officers attempt to facilitate a (peaceful) resolution to an incident where an individual barricades him/herself, sometimes with a number of hostages, in an effort to elicit some desired want” (Rogan & Hammer, 2002, p. 229-230, as cited in Ting-Toomey, 2007, p. 260). This model suggests approaching these high-pressure situations by understanding the substantive demands and interests involved in the conflict, achieving a degree of trust, power, control and empathetic understanding between negotiator and subject, accounting for the face needs of the subject, and being aware and responding to varying levels of emotional distress expressed by the subject (Ting-Toomey, 2007, p. 260). The successful conduction of these four frames is achieved through effective communication. Ting-Toomey states that this model particularly “emphasizes the important communication skills of empathic listening, and to communicate to the perpetrator that his or her intrapersonal struggling situation is acknowledged and understood” (2007, p. 261). Thus, the expression of empathic or active listening is crucial to not only understand the face needs and demands of the subject but also to negotiate emotional imbalances and reach a fair conclusion for all parties involved.

Finally, a third ICT framework known as the A.E.I.O.U. collaborative negotiation training model developed by Raider, Coleman, and Gerson (2006) for a wide range of international clients including UNICEF and the United Nations also realizes the key role of active listening in intercultural conflict management. As the initialism suggests, this model emphasizes the negotiation practices of Attacking, Evading, Informing, Opening and Uniting (A.E.I.O.U.). The first two practices include several hostile, defensive means of negotiation which are not necessarily conducive to effective intercultural communication (Ting-Toomey,

2007, p. 262). The latter three practices however are more constructive. Specifically, behaviours included in the “Opening” style include “asking nonjudgmental questions about the other’s position and underlying needs and feelings, active listening, and testing understanding via paraphrasing statements without necessary agreeing” (Ting-Toomey, 2007, p. 262). While the researchers isolate these behaviours, all can essentially be summarized as modes of “active listening” as defined earlier.

2.4 Multimodal listening in healthcare: Realizing the gap between theory and practice

In contemporary medical practice, Kagan (2008a; 2008b) asserts that patients must believe that they are being heard and understood by healthcare professionals in order for clinical interactions to be effective. However, this listening, as has been noted throughout several other domains, cannot simply be passively done; it is essential that healthcare professionals demonstrate their care for the patient through active participation in the conversation (Shipley, 2010, p. 126). Listening in the medical field is often referred to as “therapeutic, empathetic and effective” (Clementi, 2006, p. 18). It involves the ability to infer meaning from verbal and non-verbal cues as well as the ability to be silent so that patients have the time and freedom to fully explain their situation (Stickley & Freshwater, 2006). In practice, active listening is further recommended by medical practitioners as an effective means of building rapport with patients, along with a number of other best practices coinciding with active listening such as establishing an attitude of helpfulness, being present in the moment as opposed to letting other thoughts distract from the time spent with the patient, focusing on affirming the patient and encouraging further information that may be pertinent to their diagnosis and overall well-being, and humanizing the situation by practicing empathy (Lensgraf, 1989). For example, Caris-Verhallen,

Kerkstra and Bensing (1999) conducted a study of non-verbal listening behaviours performed in home nursing and home-care for the elderly. The findings suggested that there were four non-verbal cues facilitating positive nurse-patient relations: patient-directed eye gaze, smiling, affirmative head nodding and affective touch. Studies of non-verbal behaviour in healthcare settings are ubiquitous due to the critical nature of building trust with the patient, which is largely achieved not through verbal clarity and confidence but rather in patients' interpretations of non-verbal cues that relate to different aspects of the healthcare practitioner's personal qualities, as discussed above.

Although there has been evidence of a move towards a greater degree of patient-centred care in medical theory, studies have shown that the application of this paradigm shift is not yet fully incorporated in actual medical consultations. For example, Neustein (1989) delineated interactional routines characteristic of medical history-taking dialogues. Using a number of interview case studies, this research recognized the following routines: truncation, which may include interrupting the patient and effectively blocking the patient from restating facts; interrogation, which occurs alongside truncation as it entails posing a decisive conclusion to the patient's datum listing; and dissimulative negotiation, in which the doctor makes conclusions about the patient's health status with the expectation that the patient will affirm the doctor's conclusions and provide more medical history if requested. This analysis suggests a continued emphasis on meeting patients' needs quickly but generically, and that there is still room to improve doctors' abilities to meet the social and emotional needs of their patients. Similarly, Henzel (1989) investigated the variegated linguistic means of establishing social distance between doctors and patients within the medical interview context. Data were collected from a longitudinal study of interactions at a medical centre in California. As in Neustein's study

(1989), the findings suggested that physicians use the following linguistic techniques to assert that superiority and authority in medical interviews: manipulating the use of professional register, specifically lexicon, which prevents patients from fully understanding their diagnoses; eliminating patient agency, particularly in regard to their body parts and medical actions; and preference for yes/no and tag questions as opposed to open-ended questions, in addition to other imbalances of turn-taking within the dialogue that restrict the patient from speaking and having access to choice of response.

In another study, Scott et al. (2001) focused instead on the difficulties of achieving common ground, a component of the patient-centred care model, when there is a lack of agreement and disjointed communication between the doctor and the patient during medical interviews. This study utilized a multimodal comparative case study of patients with acute respiratory tract infections interacting with their physicians. The findings implied that patients use pressure tactics such as making appeals to real-world circumstances and increasing the perceived severity of their illness in order to impede doctors' attempts to reach a consensus regarding the use of a particular type of antibiotics, which stemmed from a breakdown in the application of effective patient-centred care methods. Thus, the researchers stressed the importance of finding common ground through improved communication between the doctor and the patient, primarily through an increased willingness on the part of doctors to recognize the lived, emotional experience of the patient. Likewise, Yedidia (2007) challenged the notion that there must be a prohibition of emotional involvement by the doctor and reliance upon standardized paradigms for addressing barriers in communication. This study examined the practices of palliative care as a means of improving current doctor-patient relations such that they encourage patient-centred care. Data were collected from observations of 40 patient care

sessions and interviews with patients, family members, physicians and nurses. The findings suggested that medical training programs should prepare candidates in terms of their emotional management during clinical work and their ability to approach a variety of patient needs through improved personal awareness and teamwork training.

A number of incongruencies have also appeared in the literature pertaining to doctors' ideal and self-reported performances as opposed to the realization of active listening behaviours and correlated personal traits from the view of the patient. For example, Jaspars, King and Pendleton (1983) conducted a social-psychological analysis of the challenges experienced by doctors and patients during clinical consultations. The analysis was based upon a synthesis of secondary data including interviews with doctors and physicians as well as videotaped consultations that had been collected by the authors for previous research. It was found that doctors experience the highest levels of difficulty in eliciting adequate information from their patients and effectively transmitting information to patients, who may block the information cognitively or emotionally, both of which become obstacles for successfully prescribing treatments for patients. From the perspective of the patient, consultations became difficult when the doctor did not appear to take the patient seriously, did not accept the patient as a person rather than as their illness, did not demonstrate care, and did not appear to have time for the patient. The authors conclude that these challenges are a result of a breakdown in the reciprocal social influence between the interlocutors throughout the consultation and that an improved understanding of these breakdowns must be accompanied by further research regarding the communicative processes involved in the consultation process.

Overall, these barriers to deeper conversation within clinical interviews are further enforced when cross-cultural differences are introduced, particularly in bi- or multilingual

environments. The impact of these cross-cultural barriers in relation to perceptions of doctors' empathy and intelligibility is highlighted in the following studies. Fuse, Navichkova and Alloggio (2018) investigated the role of accents in patients' perceptions of physicians' intelligibility and positive personal qualities. In this study, native monolingual and bilingual English participants listened to four non-native accents (i.e. Spanish, Chinese, Russian and Indian) and completed an online survey that allowed them to rate their perceptions of the speakers' intelligibility and personal qualities such as patience and empathy. The results suggested that increased perceptions of intelligibility correlated positively with perceptions of positive personal qualities. In addition, those who were bilingual listeners rated non-native speakers as more intelligible than monolingual listeners. In contrast, Improta (2011) highlighted the correlation between anti-active listening behaviours such as interrupting and patients' negative perceptions of their physicians. The study delved into the realization of effective communication within cross-cultural dialogues between physicians and patients using four videotaped medical interviews between native and non-native English speakers as well as questionnaires examining participants' expectations and perceptions of the interaction. The findings indicated that although participants were satisfied with their interactions and believed that effective communication had taken place, a number of pragmatic failures, including pragmatic, socio-pragmatic and non-verbal aspects of the dialogues, had in fact occurred and rendered the conversations ineffective in meeting their goals. In addition, a prevalence of interruptions and conversational redirection on the part of the doctors correlated with patients' reports that the doctor lacked empathy.

On a broader level, cross-cultural barriers can also result in an inability to attain patient-centered care in clinical practice and particularly prohibit the doctor from providing

individualized treatment plans. For example, White and Robillard (1989) conducted a case study involving a white male doctor speaking with a part-Hawaiian female patient. An analysis of the dialogue suggested that although both interactants had a shared knowledge of typical medical consultation structure and pragmatic routines, there was a difference in the speech registers of the interactants. The patient used first-person narrative, a style known as “talk story” that is common to the Hawaiian speech community, while the doctor maintained a professional, socially distant register which aimed to create an objective listing of the patient’s complaints. The researchers concluded that, although it eliminates excessive time consumed by personal narrative, the physician’s method of truncating the patient’s speech prohibits the gathering of critical information that is in fact embedded within these narratives. May (2007) also illuminated the complications that can arise between non-native English-speaking patients and their doctors during check ups when interpreters are unavailable. The dataset consisted of multiple videotaped interviews at a university health clinic as well as the patient’s and doctor’s post-consultation ratings of their own comprehension and their perceived comprehension of each other. It was found that although both parties rated themselves highly in terms of comprehension, reviews of the video transcripts as well as third party ratings of interlocutors’ actual comprehension, suggested that in fact there were several cultural- and linguistic-based misunderstandings occurring. Lamiani et al. (2008) further illustrated misinterpretations of the implications of the patient-centred care model cross-culturally by asking medical professionals in Boston and Milan were asked to write a patient-centred dialogue based upon a specific clinical scenario. The dialogues were then exchanged and discussed via teleconference. The results of the discussion demonstrated that both parties emphasized the importance of responding to the patient's experience of and reaction towards disease and illness. However, the degree of patient autonomy

was interpreted differently by the two groups, which suggested a need for a more comprehensive definition of constructing patient-centred dialogues and greater awareness regarding the various interpretations of the patient-centred care model cross-culturally.

Overall, there is still a dearth of scholarship emphasizing the role of active listening during clinical interviews within the patient-centred care model cross-culturally. Some medical commentary suggests that active listening and related communicative cues are not trainable skillsets but rather unchangeable, personal attributes, the effects of which may also be correlated with external factors such as environment. For example, Tate (1983) argues that a doctor's style tends to be static and inflexible in terms of its ability to make micro adjustments based on each client's communication style. Furthermore, despite the recent move towards a patient-centred care model, the idiosyncratic characteristics of doctors' conversational style are still indicative of a doctor as the "expert" model, which creates a power and agency imbalance between the doctor and the patient. Examples of these authoritarian behaviours include the ways that doctors gather information, prevent patients from participating in the conversation (e.g. interruptions) and control time and conversational turns.

2.5 Multimodal listening in healthcare applied: The Canadian clinical context

In the Canadian medical context, this barrier between doctors and patients as a result of cross-cultural communication challenges, has been and continues to be particularly evident, specifically in terms of the sociolinguistic expectations of active listening within the patient-centred care model to be performed by the doctor. From pre-confederation to the early 20th century, medical practice in Canada was inconsistent, as medical training was based primarily on apprenticeship rather than formal education and years of dedicated classroom study, which often led to patient harm (Medical Council of Canada, 2019a). In 1912, after an 18-year campaign, Sir

Thomas Roddick and colleagues persuaded the Canadian government to pass legislation regarding a standardized, national qualifying exam for medical candidates with the hopes of increasing the standards for physician assessment and practice throughout Canada (Medical Council of Canada, 2019a). Thus, the Medical Council of Canada (MCC) was inaugurated. Still, in the post-WWII era and into the early 1970s, many Canadian-trained doctors migrated to the United States for more advanced training and it became necessary to bring foreign-trained doctors into Canada (Wright & Mullally, 2016). This transnational migration was only one example of a larger global shift towards the out-migration of doctors, most noticeably of British-born doctors leaving the National Health Service for various commonwealth countries (Wright & Mullally, 2016). Today, although medical practice and complete medical licensure are directed by and differ slightly based upon provincial regulatory bodies, the direction of medical assessment and prerequisites to medical licensure in Canada is largely designed and administered by a council of 54 members and administrative staff at the MCC as well as a Central Examination Committee (Medical Council of Canada, 2019b). The MCC's vision commits to "striving for the highest level of medical care for Canadians through excellence in evaluation of physicians" and maintains a commitment to developing and implementing effective assessments of medical candidates' competence whilst promoting liaisons and partnerships at every level of Canadian governance and encouraging further research regarding evaluation improvement, among other goals (Medical Council of Canada, 2019c).

From a theoretical standpoint, medical care across Canada is generally guided by the patient-centred care model and the Considerations of the Cultural-Communication, Legal, Ethical and Organizational Aspects of the Practice of Medicine objectives (C2LEO objectives). The latter is centred on four key principles: respect for patient autonomy, which is defined as the

patient's right to make decisions pertaining to his or her own health and life; beneficence, which encourages physicians to consistently act in support of the patient's well-being; nonmaleficence, which is characterised as the duty of the physician to not bring intended or unintended harm to the patient; and justice, which ensures that healthcare resources and protocols will be distributed equitably amongst all patients (Gao & Naugler, 2013, p. 5). The principles of C2LEO are enforced and revised in accordance with Canada's constitutional level through the Canadian Charter of Rights and Freedoms (Gao & Naugler, 2013 p. 5).

In congruence with the C2LEO objectives, the patient-centred care model of care in the Canadian context highlights several best practices for healthcare practitioners including truth-telling, the avoidance of negligence, and the obligation of disclosure, even in the case of adverse information. The latter is characterised by a procedure for breaking bad news, known as the SPIKES protocol [Baile et al., 2000]), which focuses specifically on collaborating with the patient through supportive dialogue in order to develop a treatment that is mindful of the patient's life circumstances. Negligence, on the other hand, is centred upon the physician's duty of care, which is defined as "the degree of care and skill which could reasonably be expected of a normal, prudent practitioner of the same experience and standing" (Gao & Naugler, 2013, p. 15). Demonstrating negligence requires that the patient "prove the existence of a duty of care, a breach of duty, causation and consequent patient harm" (Gao & Naugler, 2013, p. 15). Finally, disclosure includes "whatever a reasonable person in the patient's circumstances would want to know" including the "nature of the intervention, gravity of the patient's situation and of intervention, material risks and benefits, including special or unusual risks, alternatives and consequences of nonconsent and information regarding delegation of care" which allows courts

to "infer the content of an adequate disclosure independent of a patient's whims, while remaining responsive to particular circumstances" (p. 6).

In order to gain licensure from the MCC, medical candidates must successfully complete a two-part qualifying exam known as the Medical Council of Canada Qualifying Examination Part one and Part two, which will hereafter be referred to as the MCCQE Part I and MCCQE Part II, in keeping with the abbreviations assigned by the MCC. It should also be noted that the exam developed by the MCC has undergone and continues to undergo changes and therefore it may not follow the same design at the time that this thesis is being read. The goal of the MCCQE Part I is strictly to test the candidate's knowledge of basic medical information and clinical decision-making abilities (Medical Council of Canada, 2019d). At the time of this thesis, the examination was a computer-based assessment comprised of 210 multiple choice questions and 38 short answer questions based on clinical cases and must be completed within 3.5 hours (Medical Council of Canada, 2019d). Once this part of the qualifying exam is passed, candidates must then graduate from medical school in order to be qualified to enter into a designated term of supervised practice, typically in the form of clinical residency (Medical Council of Canada, 2019).

Following the completion of at least one year of post-graduate training, candidates are eligible to take the MCCQE Part II, which will again be described as it is formatted at the time of this thesis. In this exam, candidates are tasked with applying their medical knowledge, demonstrating their clinical skills, and developing thorough, investigative and therapeutic clinical plans while maintaining a professional attitude when addressing the mock clients and medical personnel associated with the exam. The MCCQE Part II is a 12 station Objective Structured Clinical Examinations (OSCE) administered across two days (Medical Council of

Canada, 2019e). OSCEs, developed in 1979 at the University of Dundee, were created to improve the medical community's assessment of candidates' skills beyond their clinical knowledge, such as their physical examination techniques and communication capabilities (Silverman, Kurtz, & Draper, 2005) and has since been utilised by numerous medical schools across Europe and North America (Silverman et al., 2005). Prior to the use of OSCEs, case studies (long case format involving a patient and questions which were unstandardized across medical candidates) and bell ringers (strictly timed, question-based stations infamously marked in length by a bell or other form of alarm) were used to assess medical candidates (Medical Council of Canada, 2019f). However, these examinations did not examine the candidate interacting with patients, nor were the formats consistent for all candidates (Medical Council of Canada, 2019f). In Canada, the MCC develops standardized OSCEs for completion during the MCCQE Part II and regularly conducts pilot OSCEs during this second examination (however, these pilot OSCEs are not counted towards the candidate's final score). That is, each candidate will face the same problems and tasks within the same set time allotment regardless of which facility they take the exam in across Canada (Medical Council of Canada, 2019e).

The MCC website describes a number of tips in terms of how to successfully complete the MCCQE Part II, such as:

Do communicate with the patient well. The examiners will assess you, not only on what you say and ask, but on how you communicate with the patient. For example, do not reel off a series of closed-ended questions to show thoroughness or because you hope that some of them will be on the checklist. This interrogation-style approach will not score well in the global assessment, which looks at your organization, attitude, and genuine connection with the patient. Instead, demonstrate your ability to use appropriate

questioning techniques, both open and closed questions, as well as other communication skills. (Medical Council of Canada, 2019g)

The MCC also details several mistakes to avoid in the completion of an OSCE, which include the following: avoid asking questions in a rapid-fire, disorganized manner; incorporate open-ended questions rather than using directed questions, which may not accurately elicit critical information from the patient; focus on listening to the patient, as patients are not obligated to share their entire story forthright; and avoid “showing off” or lecturing in an attempt to impress the reviewers (Medical Council of Canada, 2019g) (for a full depiction of these mistakes, see Appendix A). These mistakes implicitly highlight key features of active listening, as each stresses the importance of creating a safe, accessible space for the patient to share their story while still allowing the doctor to garner accurate information within the given timeframe. Thus, the “communication” category of assessment, in conjunction with the “professionalism” category, assess not whether the doctor is able to verbally demonstrate a high level of medical knowledge but rather that they are able to establish an interaction founded in patient-centred care, of which the cornerstone is active listening.

In 2018, the MCC examination structure underwent several key changes specifically regarding its evaluation “blueprint” or breakdown of weights assigned to each assessment category. For example, the MCCQE Part II categories of “communication” and “professional behaviours” were reassigned a combined, increased weight of 50% of the medical candidate’s

overall score (Medical Council of Canada, 2019e), as demonstrated in Figure 1 below.

		Dimensions of care				
		Health Promotion & Illness Prevention	Acute	Chronic	Psychosocial Aspects	Row %
Physician activities	Assessment/ Diagnosis					25±5
	Management					25±5
	Communication					25±5
	Professional Behaviours					25±5
Column %		20±5	35±5	25±5	20±5	100

Figure 1. MCCQE part II categorical weighting

These shifts went hand-in-hand with an increase of available training resources for medical candidates. However, these resources specifically targeted material necessary for successfully completing the MCCQE Part I, and the resources available for the MCCQE Part II were limited to the MCC’s partner institution, Physicians Apply, which is dedicated to improving cross-cultural proficiency and communication skills. Dr. Claire Touchie, chief medical officer of the MCC, stated that “in summary, the MCC assessment evolution is aligning knowledge, skills and testing with 21st century medicine and advances in healthcare delivery. It connects directly and realistically with the types of problems physicians deal with daily in Canada and the types of activities they perform... We are creating and delivering a completely updated package of learning supports. With enhanced examination preparation assets coming online, more authentic clinical scenarios and better, faster access to testing across the country, the experience for all will be improved. The MCC evolution also extends to those of you coming from international locations to the Canadian system for healthcare testing” (Medical Council of

Canada, 2017b). Thus, although the MCC clearly outlines the information necessary for candidates to understand the purpose and general structure of exams, it does not provide an inclusive set of problems and study materials. Instead, it recommends a series of online and print resources for candidates to review through affiliated organizations. As a result, the corpus of materials used for this analysis (which will be described further in the “Methodology” chapter) include those which are recommended by the MCC, in order to imitate the preparatory materials available to an actual medical candidate.

Additional, unendorsed, textual resources regarding general OSCE preparation are also available. For example, Gao and Naugler (2013) present a thematic list of 133 case studies that medical candidates can review in preparation for clinical skills exams. The case studies test the candidate’s knowledge in the following areas: general medicine, obstetrics and gynecology, pediatrics, preventative medicine and community health, psychiatry and neurology, and surgery. Each example reflects the design of a typical station that the candidate would encounter in a clinical skills exam such as the Medical Council of Canada Qualifying Exam II (MCCQE II). In another resource, Hurley (2011) details peripheral components of the exam such as professional conduct, general verbal and nonverbal communication skills, cross-cultural communication guidelines and conducting a medical history review. Additionally, Hurley (2011) provides a comprehensive set of case studies on the following topics: the cardiovascular system, the respiratory system, the gastrointestinal system, the genitourinary system, the nervous system, the musculoskeletal system, dermatology, hematology, endocrinology, psychiatry, women’s health, pediatrics, geriatrics, and ethics.

It should be reiterated that these resources focus primarily on preparation for the first MCC examination and not on the second, OSCE-style exam. As a result, there are few

illustrations of the patient-centred care model, particularly the role and performance of active listening, within the Canadian clinical context. For doctors who have received their training and residency terms in Canada, this is not very problematic, as the culturally based listening skills appropriate for eliciting information from patients both during the second MCC examination and in actual practice will have been well-practiced during their residency hours. However, for foreign-trained doctors, who had been and continue to be brought into Canada to fill the demand for more doctors across the country, this knowledge may not be as intuitive. This lack of cross-cultural knowledge is further complicated by the lack of resources available which specifically depict this unique skillset. As a result, foreign-trained medical doctors have overall lower pass rates on the MCC examinations in comparison to their Canadian-trained counterparts, particularly on the MCCQE II, which examines the interactional abilities of the candidates inasmuch as their clinical knowledge. In 2017, only 74% of first-time MCC exam takers passed the assessment, which is only a minor improvement from those passing the exam in 2013 (Medical Council of Canada, 2017a, p. 34). For those who trained outside of Canada, taking the exam for the second or greater time, the pass rate is even lower, with 44% of candidates passing in 2013 and 64% passing in 2017 (Medical Council of Canada, 2017a, p. 34). The disparity between the pass rates of these individuals and those who have been trained in Canada is great, as those trained in Canada had an overall pass rate of 95% in 2013 to 97% in 2017 (Medical Council of Canada, 2017a, p. 34). Thus, as both parties can be assumed to have obtained similar clinical knowledge (foreign-trained doctors are required to "catch up" to their Canadian colleagues, if needed, prior to undertaking the MCCQE Part I) in order to reach the second examination, it seems highly plausible that the disparity in pass/fail rates between foreign-trained and Canadian-trained medical candidates may lie not in their medical knowledge. To some

degree, the pass/fail rate may be attributable to foreign-trained doctors' lack of familiarity with the assessment style used in North America (i.e. OSCEs) as well as variances in English or French language proficiency. However, this rate appears to be largely connected to the doctors' soft skills of interacting with mock clients, primarily their ability to actively listen. Indeed, the majority of items listed on the MCCQE Part II scoring scheme assess active listening within the verbal mode and one grouping of assessment items titled "Interpersonal Behaviour" explicitly recognizes the non-verbal mode of active listening (see Appendix B for further details). Thus, this study will first observe interactions between foreign-trained, non-native English-speaking doctors and their native English-speaking, Canadian patients and compare the active listening strategies used to those modelled by native-English speaking, Canadian-trained doctors in videos provided through an MCC partner organization in order to gain an understanding of any cross-cultural differences in terms of active listening and to provide recommendations for improving cross-cultural active listening training to accommodate a diverse candidate population. In short, this thesis project will respond to the following research questions:

1. How are active and anti-active listening traits realized and used on a micro level by native and non-native English-speaking medical doctors in order to create a collaborative, empathetic relationship with patients during Canadian Objective Structured Clinical Examinations?
2. How are active and anti-active listening traits realized and used on a macro level by native and non-native speaking medical candidates in order to create a collaborative, empathetic relationship with patients during Objective Structured Clinical Examinations?
3. How does a cross-cultural understanding of active listening impact OSCE success and inform subsequent training programs in Canada?

Chapter 3: Methodology

3.1 Introduction

This chapter will further expand upon the methodological process that was used to answer the guiding research questions. Overall, this study utilized a cross-sectional research design within a quantitative method approach (Dörnyei, 2007). Thus, this chapter highlights this overarching research design and approach by delineating the following: the evolution of an active listening analytical framework from its inception in the literature to its adjustments in a pilot study and finally to its establishment within ELAN, an audio-visual software, and its application to the control groups and test group; and the ethics procedures and workflow used to guide the two main phases of this study.

3.2 Developing an active listening framework

The methodological application of this approach for the present study was largely unprecedented. Multidisciplinary approaches to active listening assessment prior to this project have focused primarily upon self-reports of listening and of personal qualities that coincide with good listening skills as well as evaluations of these listener qualities from the speaker's perspective. Three main methods of assessing active listening have been developed in this regard: the Active Empathetic Listening Scale (AELS) (Bodie, 2011), the Active Listening Attitude Scale (ALAS) (Mishima, Kubota, & Nagata, 2000) and the Active Listening Observation Scale (ALOS-global) (Fassaert, van Dulmen, Schellevis, & Bensing, 2007). The AELS is a self-report measure that includes 11 questions, scored on a seven-point Likert scale, pertaining to three main themes: sensing (e.g. "I am aware of what others imply but do not say"), processing (e.g. "I keep track of points others make"), and responding (e.g. "I ask questions that show my understanding of others' positions") (Bodie, 2011). The ALAS, scored using a four-

point Likert scale, consists of questionnaire with 31 items addressing the areas of listening attitude (e.g. “I hurry him/her into talking faster”), listening skill (e.g. “I pay attention to his/her unexpressed feelings”), and conversation opportunity (e.g. “People feel easy to talk to me”) (Mishima et al., 2000). Finally, the ALOS-global (Fassaert et al., 2007) is composed of seven items pertaining to speaker’s beliefs that they are being listened to, including: level of distraction, expression of hastiness, attentiveness, appropriate time and space given to the speaker, questions, ability to lead the conversation, and non-verbal expressions of understanding (Fassaert et al., 2007).

As stated, while these methods are useful in obtaining some understanding of active listening taking place in a variety of environments, they do not provide a microanalysis of individual components of active listening, specifically from a multimodal analysis perspective, as most do not address all manners of verbal and non-verbal active listening traits. Another key difference between this study and former active listening assessments is the choice of data. Furthermore, previous studies have focused on the listener’s own attitude towards themselves and their listening proficiency, or the speaker’s evaluation of the listener’s personal and communicative qualities without attempting to quantify the actual realization of these active listening components in comparison to the resultant ratings of doctors’ performances. Thus, it was determined for this study that pre-recorded videos would be assessed by the researcher, who is a third party in the doctor-patient interaction, for the sole purpose of quantifying anti-active and active listening traits displayed by participants in relation to the success or failure of participants during the mock examinations, which will be analysed based on the assessments provided by external reviewers following the mock examinations. As such, it is critical to also incorporate a theoretical framework that has examined time-aligned corpora in a multimodal

framework. Specifically, this study utilises the method of analysing listening developed by Tsuchiya (2013), who applied a combination of global pattern analysis and turn structure analysis to a time-aligned multimodal corpus.

Furthermore, these active listening analyses (e.g. Fassaert et al., 2007) do not fully consider the extent of the impact caused by lower levels of active listening on the communicative event. Low levels of active listening expression appear to still actively affect the evaluation of a listener's personal and listening qualities from the speaker's perspective, such as producing a perception of lack of listener empathy, rather than being passively noted by the speaker and judged as a merely an act of non-listening with no emotional or psychological bearing. These components will be referred to as behaviours of "anti-active listening" that produce the opposite effects of active listening. As a result, the present study attempts to produce a synthesized assessment with respect to all three models described above in amalgamation with specific qualities of verbal and non-verbal, active and anti-active listening mentioned throughout active listening literature.

3.3 Video analysis tool selection

The primary coding software utilised for this project was EUDICO (European Distributed Corpus project) Linguistic Annotator or, as it is more commonly known, ELAN. ELAN was developed by the Max Planck Institute and has been maintained by scholars at the institute and around the globe for nearly a decade as a free, open source tool for creating time-aligned annotations from audiovisual recordings (Cox, 2014). It was originally designed as a means of recording and transcribing signed languages; however, today it is a widely used tool for projects both within and outside of the field of linguistics. ELAN software was selected as the primary means of coding for this project for several key reasons. Namely, it provides coders with a highly

efficient means of coding lengthy, multimedia files with unlimited tier applications, multiple symbolic hierarchies that are used to create meaningful segments within an audiovisual text, without causing the program to fail. Many similar programs have been known to shut down or malfunction if videos exceed certain time lengths and may otherwise be compromised when having to deal with multiple annotation tiers. In addition, ELAN software is compatible with any computer system and will not inhibit data access if the coder shifts from one system to another (e.g. moving between a PC and a MacBook). Furthermore, it utilises Unicode-encoded XML and can therefore sustain the integrity of a file over time with few to no changes regardless of computer system updates (Cox, 2014). Finally, the researcher had obtained extensive training in ELAN functions during a master's level course and thus deemed it most prudent to apply this skillset as opposed to learning a new program for this project.

However, there are two notable limitations with ELAN software. First, it is not possible to remove undesirable aspects of audio recordings once the files have been opened within the ELAN program. Thus, if there were unusable sections of video such as, in the present study, extended periods of time in which the doctor was not even in the same room as the patient, they would need to be removed in a separate software program prior to coding. The second limitation with ELAN is that annotations are only available in plain text (i.e. it is not possible to underline, boldface or use colour within annotations) (Cox, 2014). As the discrepancies within the videos, such as false starts or hesitations, would be critical for this study, a detailed coding scheme was developed which did not depend on extensive font selections. As a result, neither of these limitations were found to be overly problematic for the current study.

3.4 Methodology phase 1: Ethics and recruitment

As the researcher did not have access to the facilities and resources necessary to observe OSCEs firsthand, it was determined that videos recorded at formal OSCE training facilities would be utilised for this project. This entailed utilising one set of videos for the pilot study component of this project and two sets of videos for the main study conducted for this project. The 8 videos used in the pilot study were obtained from the MCC website. These were videos that aimed to depict scenarios similar to those that would be found in the MCCQE Part II and depicted doctors who practiced in Canada and spoke either French or English as a first language. However, they were not necessarily exemplary demonstrations of listening and therefore could not be included within the control group video sets of the main study.

The first set of videos in the main study were selected based upon their demonstration of satisfactory and unsatisfactory characteristics of active listening. This set of 21 videos performed by 21 different participants, obtained from an organization affiliated with the MCC known as Physicians Apply, were scripted and featured doctors who practiced in Canada and spoke English fluently. As a result of these inclusion criteria, these videos could be utilised as an exemplary baseline comparison to recordings of actual medical candidates undergoing an OSCE (see Appendix D and Appendix E for video codes and corresponding video lengths and content). The use of the Physicians Apply videos in the main study as well as the MCC videos in the pilot study was governed by the Department of Research at the MCC. Upon reaching out to the MCC to describe the nature of this project, it was determined that written permission to use the videos would supersede any need for further clearance from the MCC research review board as the videos were available on publicly accessible websites.

The third set of videos in the main study included 5 participants performing in 5 different OSCEs similar to those present on the MCCQE Part II. These videos were recorded by Deidre Lake at various Communicate2Integrate (C2I) and Alberta International Medical Graduates Association (AIMGA) facilities across Alberta. Both programs aim to prepare non-native English-speaking individuals for the MCC qualifying examinations. These videos became accessible to the researcher following the project's clearance by the Carleton University Research Ethics Board¹ (see Appendix C for ethics clearance form), which will be described further below. The inclusion criteria for participants in this study included medical candidates who spoke English as a second language and who were eligible to take the MCCQE Part II. Conversely, the exclusion criteria for participation in this project included medical candidates who spoke English as a first language or who were otherwise ineligible to take the MCCQE Part II. The exact format of the OSCEs in the videos provided by Lake varied as they would in an actual MCCQE Part II OSCE. Thus, participants may have been asked to take a medical history, conduct a physical exam, counsel, or break bad news. All OSCEs, for mock and actual examinations, are performed on actors using fictional scenarios that are intended to imitate the scenarios presented during actual practice. As a result, no structured interview took place; however, examiners were able to informally question participants' performance at the end of the OSCE. Following the completion of their OSCE, participants would receive their pass/fail results as well as performance feedback from the reviewers.

In order to obtain these videos and results, the researcher held an initial, informal interview with Lake to discuss the nature of the project and to ensure that the direction of the project would meet the research goals of C2I and AIMGA as well as have the ability to provide

¹ Project ID # 109402

research participants with useful feedback. Following this discussion, the researcher provided Lake with a Letter of Invitation and informed consent document to give to potential participants. If consent was given, the researcher was given access to the video recordings of 5 OSCE scenarios that had been undertaken by the medical candidate. Ultimately, this resulted in a total of 25 videos featuring 5 participants. These were made available to the researcher on a private, password-protected Vimeo webpage. The webpage access was removed by Lake following the completion of the study to ensure that the videos were not reused for future projects, which would violate the permission given by the participants in the informed consent document. All videos were then coded via ELAN software which, throughout the course of the project, were maintained on a secure, password-protected external hard drive, and then, at the completion of the project, the videos will be permanently erased. Although permission to receive the participants' final performance scores and commentary was also granted in the consent form, it was the impetus of the participant to notify Lake with the results, as Lake does not receive these results formally. Unfortunately, none of the participants chose to inform Lake and, as an extension, the researcher, regarding these results. The implications this had on research findings and the need for future studies will be discussed further in the discussion chapter.

Overall, the study required minimal interaction with the participants as the study only consisted of observing the videos. In addition, no previous relationship existed between the researcher and the participants, and the researcher's relationship with both Lake and the MCC only existed in the formal interactions described above, which limited the potential for conflict of interest in this study. No interviews or additional direct interaction with the participants on the part of the researcher were conducted for this study. The only interaction with participants was to be initiated by participants if they wished to withdraw from the study by December 1, 2018,

and/or have a copy of the final research manuscript if it did not compromise the confidentiality of the other participants in the study. However, no participant contacted the researcher in order to have their information removed from this study.

A number of additional safeguards were set in order to prevent confidentiality breaches throughout the course of the study. Although the study itself was considered for minimal risk review by CUREB, the participants were also considered to be a potentially vulnerable population as they were not native English speakers and they were recipients of the C2I and AIMGA services and could have felt that the research was obligatory. The first vulnerability was mitigated by Lake prior to the participant's mock OSCE. Participants had been vetted prior to their registration such that only those with sufficient English proficiency would be able to participate in the mock OSCE. In order to mitigate the second vulnerability, which is more applicable to the present study, participants would be informed verbally by Lake as well as in the Letter of Invitation and the informed consent document that the study would not negatively impact their previous or future mock or actual OSCE. In addition, there were no monetary or other compensatory gains in place to persuade participants to give consent. The key benefit for participants was a final document regarding their performed communicative style in comparison to the standardized model of communication in the given setting. Furthermore, there were no elements of deception involved in this study; participants were given full disclosure regarding the nature of the study in the Letter of Invitation and informed consent document. A final consideration regarding the identifiability of the participants was given. It was determined that all secondary data (i.e. copies of video recordings on ELAN and performance scores) would be anonymized. Participants, including medical candidates and examiners, would be assigned alphanumeric codes, which also made reference to their L1s when applicable, and video

submission dates would be replaced with the timestamp 1969:07:20 00:00 in order to remove all names and identifying information (see Appendix F for a full list of video codes and corresponding video lengths and content). This anonymized information would be used as well in reports or presentations of the project. Additionally, the results of the study were often presented to describe groups as a whole rather than individual participants in order to further increase anonymity. Finally, direct images of participants from video sets were not included in this thesis, to protect participants' identities. A line drawing taken from one of the pilot study videos which obscures the faces of participants was included in this chapter strictly for the purpose of illustrating the physical set up of an OSCE. Also, full transcripts were not provided in this thesis as permission was not received from project partners to reproduce this material; however, short quotes were used to exemplify the verbal mode in the descriptions of the active and anti-active listening cues and were attributed only to the video set that they were derived from (i.e. the MCC) and not from specific participants, which was again done to protect participants' identities.

3.5 Methodology Phase 2

3.5.1 Methodology phase 2, part A: Establishing a pilot study

As the overarching analysis framework to be used for this thesis was comprised of an amalgamation of studies rather than a single, pre-tested method, it was determined that a pilot study would be conducted to verify the framework's ability to meet the objectives of the present study. The pilot study was achieved in two parts. The first stage consisted of realizing the framework in the ELAN program and testing its usefulness, including recognizing any redundancies or gaps in the framework's categories, the second stage of this pilot study consisted

of garnering third party feedback regarding the categorization of different annotations through informal discussions using archetypal annotations found in the initial stage of the pilot study.

In regard to the initial stage of the pilot study, the 8 videos from the MCC, ranging in length from one minute to fourteen minutes, were analysed. The video content covered a variety of topics, including in-depth interviews and more hands-on physical studies. These pilot study videos were deemed to be “neutral”, in their expression of active listening. That is to say that they were presented merely as a means of explicating the nature of the OSCE design and exemplary scenarios. They were not suggested to be models of “best practice” in terms of communicating or approaching OSCEs, and thus could not be used as any sort of baseline model for the main study. However, they followed a similar OSCE structure to those provided by the project partner at AIMGA and C2I and could thus be used to effectively establish an analytical framework.

3.5.2 Methodology phase 2, part B: Designing the analytical framework through a pilot study

Prior to any coding, the videos had to be paired with WAV files (i.e. separate audio tracts) in order to be readable by the ELAN software. This was done through separate, offline software in order to avoid any potential confidentiality breach, which could be possible through online, third-party software companies. Once the two file types were prepared, they could be inserted into ELAN to begin the tier framework set up (see Figure 2). Initially, only one video was transferred into and analysed with the ELAN software. This was done to allow the coder to set up and modify, as needed, a tier framework that would most effectively encode all the active and anti-active listening cues before copying and applying the tier framework to all other video files. To begin, the coding language was set to English. By doing so, if languages other than

English appeared in the videos, they could receive a distinct code, which would be easier to locate and tabulate, if present, during the findings phase of the study.

ELAN 5.1

File Edit Annotation Tier Type Search View Options Window Help

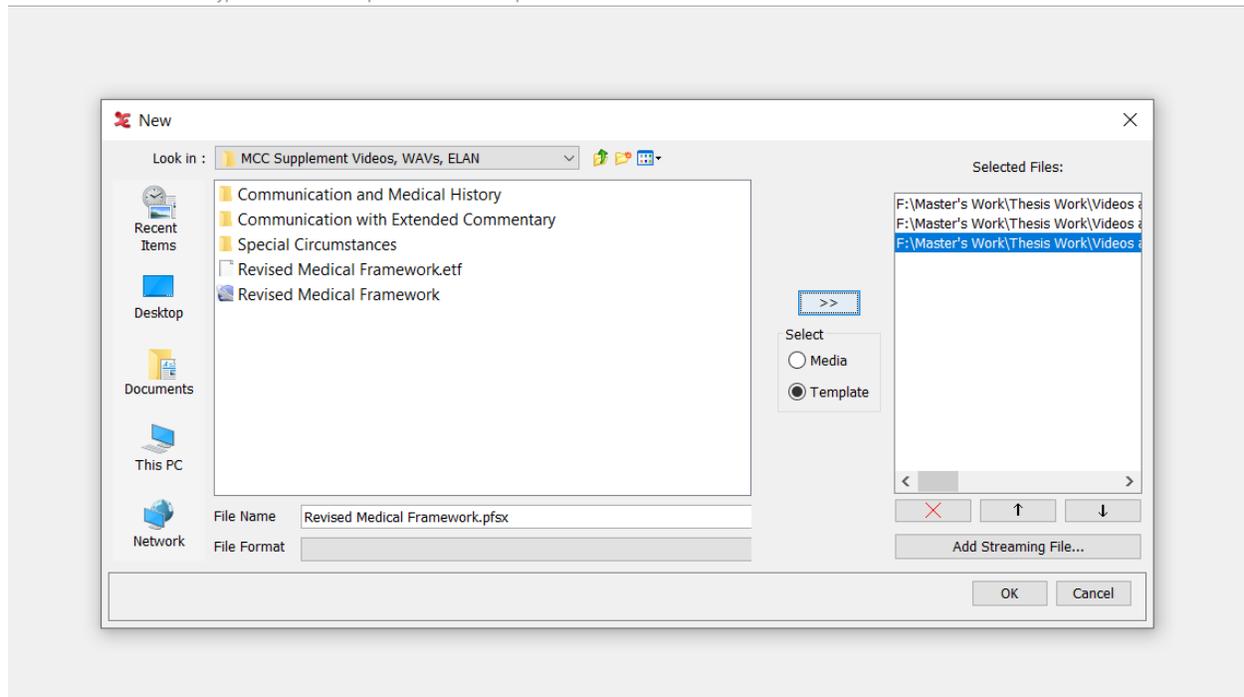


Figure 1. Transferring audio, video and framework files into ELAN

Next, the types of tiers to be used in this analysis were defined based on the literature which suggested that active listening and its opposite, anti-active listening, were multimodal acts, each of which could be broadly categorized into sets of “verbal” and “non-verbal” acts with several smaller subcategorizations (e.g. “asking mindful questions” and “providing affirmative vocalizations” could appear under the category of “active listening verbal”). Therefore, as a result of these studies, the following tier labels were introduced, which is further illustrated in Figure 3: "Active Listening Verbal", "Active Listening Non-Verbal", "Anti-Active Listening Verbal", "Anti-Active Listening Non-Verbal", "Transcriptions" and "Notes". Once these base archetypal tiers were defined, the main tier framework could be established. Initially, a total of seven tiers were introduced: “Doc ALV” (a secondary analysis of the doctor’s verbal active

listening traits based upon an initial transcription of the doctor’s speech), “Doc ALNV” (doctor’s non-verbal active listening traits), “Doc AALV” (a secondary analysis of the doctor’s verbal anti-active listening traits based upon an initial transcription of the doctor’s speech), “Doc AALNV” (doctor’s non-verbal anti-active listening traits), “Doc transcript” (transcription of the doctor’s speech), “Pat transcript” (transcription of the patient’s speech), and “Background notes”. For the scope of this thesis, the text spoken by the patient was sufficient for analysis as it is only being used to observe whether or not the medical candidate directly refers to the patient’s commentary later or not. Additionally, “Background notes” referred to any peripheral details such as noises in the background, shifts in the room layout during the mock OSCE, etc. that may be notable for a basic qualitative analysis of features of the discourse setting such as power.

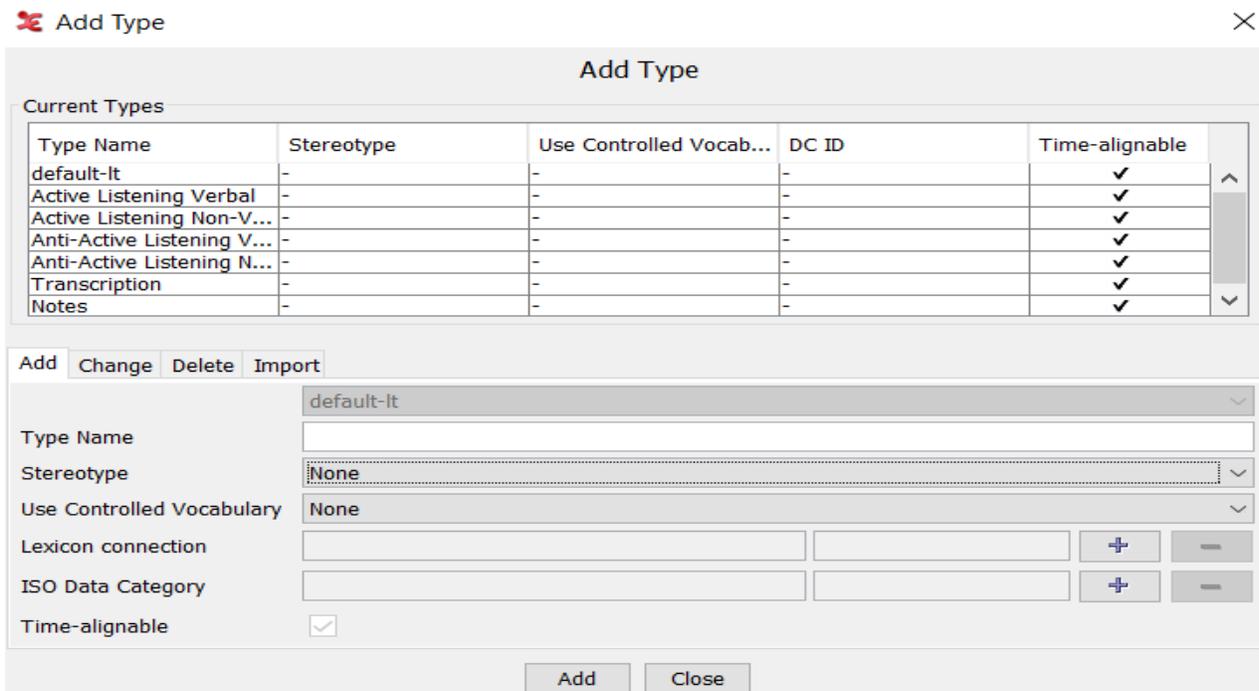


Figure 2. Creating tier archetypes

By assigning each of these seven categories to a separate tier, the coder was able to isolate and timestamp individual occurrences of each feature as opposed to having these segments automatically be assigned to tiers based on prescribed relationships within the

framework design. Each of these tiers were assigned a corresponding archetypal tier type, as established in the step prior, to make them more accessible as well as to establish their symbolic relationship to that archetypal tier label. This step also became crucial during the revisions phase of the pilot study as it allowed distinct behaviours within different categories to be later and to viewed as collective concepts, both within the individual video file and across multiple video files, rather than individual components later on in the transcription mode. It was found that some tier categories, namely Doc ALNV and Doc AALNV, were still too broad when coding began. Although they first pertain to the medical candidate, the individual components that would create the codes in each may not all occur at the exact same time. For example, while in the listening role, the candidate may nod only once while maintaining eye contact with the speaker throughout the speech act. In this example, the nodding will only occur for perhaps two milliseconds while the eye contact may be maintained for several seconds. As both are being coded on a single tier and font differentiation and highlighting functions are not available in ELAN, it was not possible to create sub-segments to indicate that these events, although both falling under the category of “Doc ALNV”, occurred within different timeframes. Thus, a new series of tiers pertaining to these subcategories of non-verbal communicative acts were created, including Doc ALNV-EYE, Doc ALNV-HEAD, Doc ALNV-HANDS, as were their anti-active listening non-verbal equivalents. Additionally, all categories were assigned the general tier type of “Active Listening Non-Verbal” or “Anti-Active Listening Non-Verbal” in order to maintain the symbolic integrity of the tier with its tier archetype (see Figure 4 for final tier breakdown). In transcription mode, this allowed all the subtypes to be viewed at once, with some context of their coalescence. However, this should not undervalue the importance of the interdependent relationship between individual active and anti-active listening components.

Add Tier ×

Add Tier

Current Tiers

Tier Name	Parent Tier	Tier Type	Participant	Annotator	Input Method	Content Language
Doc ALV	-	Active Listening Ver...	Doctor		-	eng
Doc ALNV-EYE	-	Active Listening Non...	Doctor		-	eng
Doc AALV	-	Anti-Active Listening...	Doctor		-	eng
Doc AALNV-EYE	-	Anti-Active Listening...	Doctor		-	eng
Doc Transcript	-	Transcription	Doctor		-	eng
Pat Transcript	-	Transcription	Patient Ken Scott		-	eng
Background Notes	-	Notes			-	eng
Doc ALNV-HANDS	-	Active Listening Non...	Doctor		-	eng
Doc ALNV-HEAD	-	Active Listening Non...	Doctor		-	eng
Doc AALNV-HANDS	-	Anti-Active Listening...	Doctor		-	eng
Doc AALNV-HEAD	-	Anti-Active Listening...	Doctor		-	eng

Add

Doc ALV

Tier Name

Participant

Annotator

Parent Tier

Tier Type

Input Method

Content Language

Figure 3. Programming a useable tier framework

Labels regarding the participant of each tier were also assigned at this phase of the tier framework establishment. All tiers pertaining to the medical candidate or the patient were assigned a participant label. These were codified as the participant role (i.e. “Doc” (medical candidate) or “Pat” (patient)) along with a numerical code beginning with “1”. In addition, if it was made evident to the coder that the speaker spoke a known first language other than English, this information was indicated in brackets following the alphanumeric code (e.g. “Doc1(Spanish)”). While these codes were used to differentiate between participants of different videos, if the same medical candidate performed in multiple videos, they maintained their given label. Participants of the “patient” role, even if they appeared in multiple videos, were assigned different codes as it was not relevant to track which videos they did or did not appear in. All participant codes and corresponding videos were maintained in a separate spreadsheet for ease of accessibility throughout the project. The “Background Notes” tier was the only tier to not receive

a participant. This was done because the “participant” in this category was not consistent across all videos. For some videos, the only background notes to make included street noises or buzzers signalling the beginning or end of the session. In others, the candidates were assisted by nurses or given in-session direction from the examiners. When nurses or examiners were “participating” in these videos, they were clearly labelled as such within the “Background Notes” tier transcription. All other non-human sounds were simply assigned a description (e.g. “long buzzer sound”).

The next step in the workflow was to begin creating annotations or, more generally, to begin the video analysis process. This initial analysis phase was conducted in the Segmentation Mode in ELAN (see Figure 5 for isolated view of tier framework in Segmentation Mode and Figure 6 for broader view of creating annotations on tiers in Segmentation Mode). An annotation is an individual segment of audiovisual media that has been isolated by the coder on the basis of some individual feature, such as an actor speaking, in relation to the duration of time it took for that feature to occur or even to be absent. One exception to this principle occurred in the creation of annotations on the tier “Doc AALNV-EYE”. Intuitively, it would seem that annotations made on this tier would correspond with absences of annotations made on the “Doc ALNV-EYE” tier, as any period of absence on the latter tier would indicate a period that lacked eye contact from the doctor. However, in normal conversation interlocutors do not maintain constant eye contact throughout the conversation, and it would be unreasonable then to count these breaks in eye contact as acts of anti-active listening. However, some situations demand eye contact as a means of demonstrating active listening and consequently of demonstrating listener empathy and compassion for the speaker. In the case of these videos, these moments would be indicated by the speaker actively “searching” the candidate’s face while speaking, and potentially making a

subtle, non-verbal move to request eye contact from the candidate, such as leaning forward or making hand gestures towards themselves. Instances such as these were thereby annotated and described within the “Doc AALNV-EYE” tier.

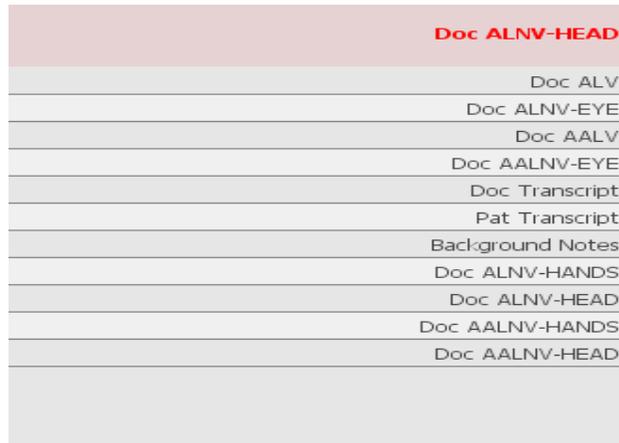


Figure 5. Installation of the tier framework (segmentation mode)

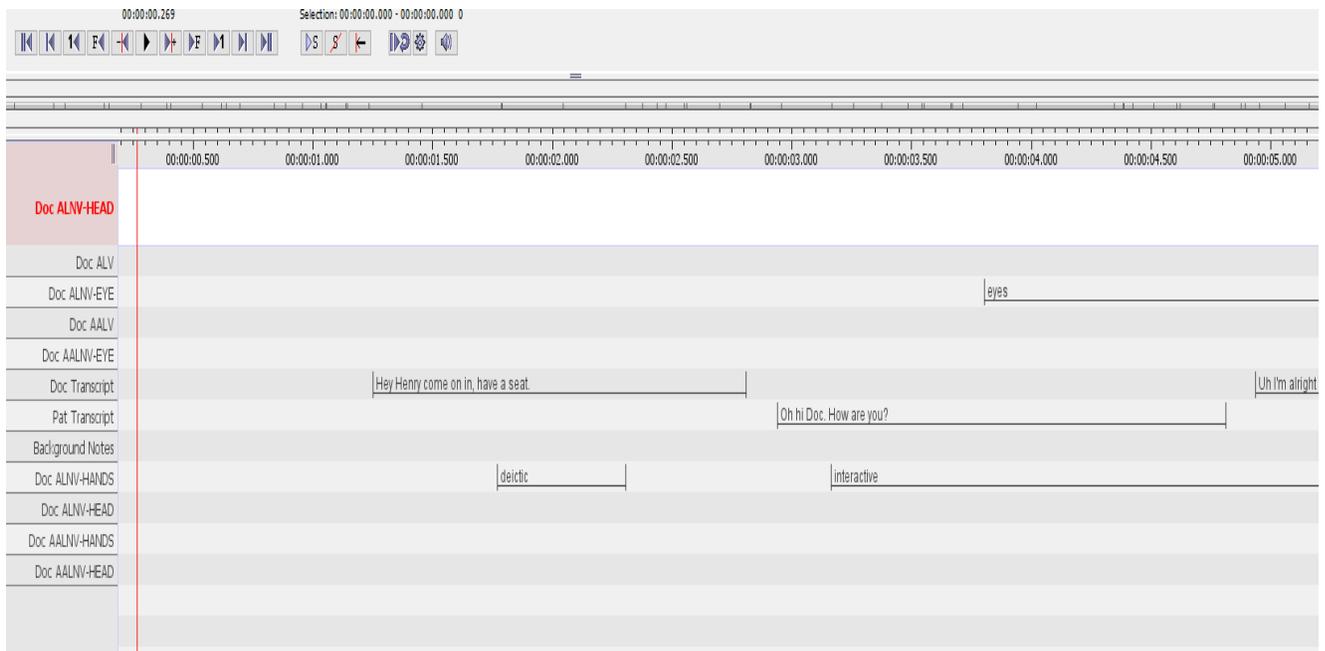


Figure 6. Snapshot of overall tier framework in segmentation mode with annotations on tiers

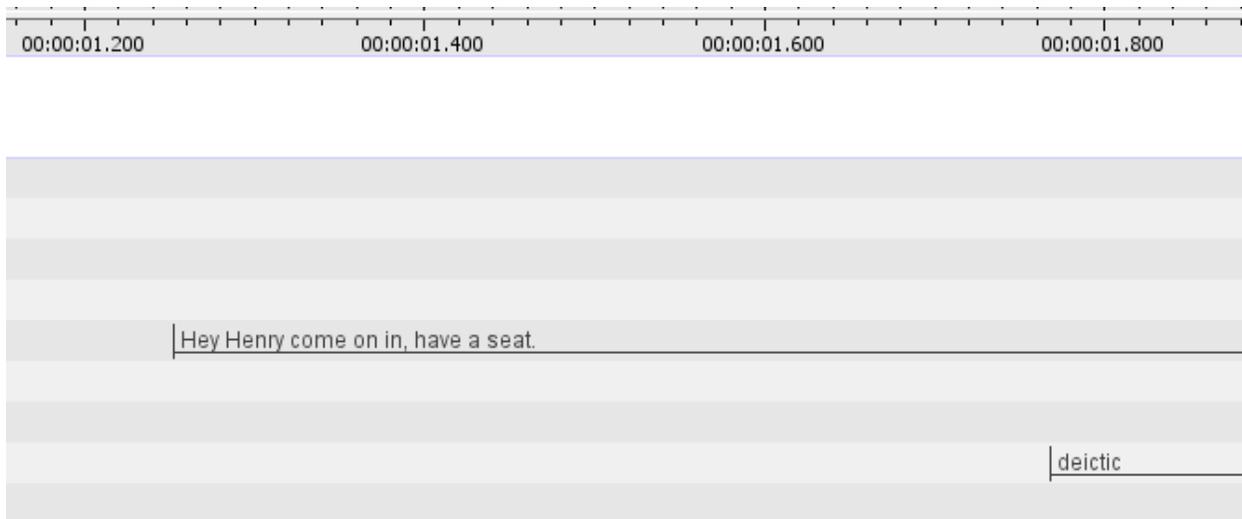


Figure 7. Example of an annotation (i.e. "Hey Henry, come on in, have a seat") on a tier with timestamp indication

After the annotations had been created, they could be assigned descriptors or, more specifically, transcriptions (see Figure 7 for example of annotation with descriptor in Segmentation Mode). In ELAN, this entails shifting from Segmentation Mode to Transcription Mode, which allows the coder to view a clear listing of empty text boxes that are associated with each annotation (i.e. the previously identified, time-stamped audio video file) as opposed to viewing one long stream of audio wavelength with intermittent annotation boxes. In other words, Transcription Mode enables faster and easier access to different annotations without having to scroll and search excessively through the audio wavelength. For most annotations, textual description was relatively straightforward (see Table 1 for full verbal and non-verbal coding schema). Transcriptions were representations of what was spoken by participants. Also, the various non-verbal annotations were labelled according to which action the participant was visibly performing. However, some annotations received additional data labels in order to make each as detailed and accurate as possible. For example, the non-verbal categories of “Doc ALNV-HANDS” and “Doc AALNV-HANDS” were further illustrated in a second round of

analysis of the annotations. While it was interesting to note that the candidate was making some sort of hand gesture, a deeper understanding of the interlocutor's engagement with the other speaker could be garnered from a more precise description of exactly what type of gesture was being performed. Thus, labels were assigned according to McNeill's (2005) delineation of gestures. Gestures are not a single action but rather a series of phases composing a gestalt of a movement known as a gesture. Generally, the gesture is composed of a preparation period, a stroke, and a retraction (2005, pp. 31-32); however, it may also include cessations of movement or moments when the individual has stopped the progression of the gesture phases in order to re-synchronise their speech with their non-verbal performance, indicating the tight interconnection of speech with gesture (McNeill, 2005, p. 33). These cessations are known as "holds" (McNeill, 2005, p. 32). Of the three main phases, only the stroke is understood to hold meaning and be obligatory in a given gesture; thus, a gesture is not thought to have occurred without the stroke (McNeill, 2005, p. 32). Nobe (2000) and Valbonesi et al. (2002) have demonstrated through analyses of gestures that strokes are synchronised with co-expressive speech 90 percent of the time. McNeill explains that "when strokes are asynchronous, they precede the speech to which they link semantically, usually because of brief hesitations, and the time gap is small. Strokes rarely if ever follow their co-expressive speech..." (2005, p. 32). The part of speech that is thought to correspond with the gesture in regard to meaning is known as the "lexical affiliate" (Schegloff, 1984). As noted above, in some cases the gesture may not always occur at precisely the same moment as its lexical affiliate. In some instances, the gesture, particularly the stroke phase, will anticipate the affiliate (McNeill, 2005, p. 37).

In attempting to place gestures into categories, a number of challenges can occur. In particular, positive results of intercoder reliability are often difficult to achieve as a result of

these restrictive labels (McNeill, 2005, p. 38). Thus, McNeill (2005) developed a four-dimensional model of interpreting gestures that allow for the flexible nature of gestures which will also be used for many of the non-verbal gestures, particularly those pertaining to hand gestures, found in this study. McNeill defines the four dimensions as follows (pp. 39-40):

- Iconic: “Such gestures present images of concrete entities and/or actions. They are gestures in which the form of the gesture and/or its manner of execution embodies picturable aspects of semantic content (aspects of which are also present in speech” (p. 39).
- Metaphoric: “Gestures can also present images of the abstract. Some gestures involve a metaphoric use of form- the speaker appears to be holding an object, as if presenting it, yet the meaning is not presenting an object but rather that she is holding an ‘idea’ or ‘memory’ or some other abstract ‘object’ in her hand. There is an iconic component (the form of the gesture resembles holding an object) and a metaphoric component (holding or presenting something is a metaphor for presenting a meaning). Other gestures involve a metaphoric use of space. A speaker, for example, divides the gesture space before him according to an appearance-reality dimension of morality being attributed to the story character” (p. 39).
- Deictic: “Although the prototypical deictic gestures are the (pointing hand symbol) hand with an extended index finger, almost any extensible body part or held object can be used for pointing. If the hands are employed otherwise, we can improvise pointing with our heads, noses, elbows, feet, etc., as well as with abstractions like the ‘eidola’ or gaze of ancient ophthalmology...Deixis entails locating entities and actions (end p. 39) in space vis -a-vis a reference point...” (p. 40).

- Beats: “Beats are among the least elaborate of gestures formally. They are mere flicks of the hand(s) up and down or back and forth that seem to ‘beat’ time along with the rhythm of speech. However, they have meanings that can be complex, signaling the temporal locus in speech of something the speaker feels is important with respect to the larger discourse” (p. 40).

A final gesture type called “interactive”, proposed by Bavelas, Chovil, Lawrie and Wade (1992) was also used in this analysis as it goes beyond merely providing information or progressing the conversation. Rather, these gestures help maintain the conversation as a social system (Bavelas et al., 1992, p. 470) and involves physically “reaching out” to the other interlocutor. These gestures are particularly critical to this study as they direct indications of the interlocutor’s intention to express some degree of empathy, one of the primary goals of active listening. In addition, notable absences of gestures when gestures are anticipated are interpreted as a result of the “complete predictability of the next step in discourse” (McNeill, 2005, p. 54). In terms of this study, these absences could be noted in the dearth of nodding performed by the doctors. These were interpreted as “predictable” in the sense that the doctor filled this gap through verbal acts of affirmation, such as saying “okay”.

It should be noted that head nods were initially given more detailed description in relation to the number of nods performed in succession, as it was believed that the disclosure of more highly sensitive information would require a greater level of active listening and perhaps affect the number of head nods performed. Thus, single nods were assigned the label “nod x1” whereas two or more consecutive nods were initially assigned the label “nod x2”, regardless of the quantity of nods performed. However, it was quickly discovered that the exact beginning and end of a nod was highly subjective and to count individual nods was trivial information not needed

for this study. Thus, “nods” in this study refers to segments of nodding, which may vary in terms of the exact number of nods performed.

No	Type 1 : Active Listening Non-Verbal
Doc ALNV-HANDS	
1	deictic
2	interactive
Doc ALNV-EYE	
3	eyes
4	eyes
Doc ALNV-HANDS	
5	deictic
6	deictic
Doc ALNV-EYE	
7	eyes
8	eyes
Doc ALNV-HANDS	
9	deictic
Doc ALNV-EYE	
10	eyes
Doc ALNV-HANDS	
11	beats

Figure 8. Snapshot of all "ALNV" tiers in transcription mode, including timestamps

As mentioned earlier, during this phase of data collection, it was possible to look at multiple, related tiers at the same time (see Figure 8), whereas in Segmentation Mode only a single tier could be accessed at a time. This was crucial for the assignment of the active and anti-active listening verbal category. These assignments were based entirely on the transcriptions of the candidate and patient, and thus the tier type groupings of “Transcriptions”, “Active Listening Verbal” and “Anti-Active Listening Verbal” had to be readily accessible within Transcription Mode within a single screen. The active and anti-active verbal categories were assigned following the “Transcription” text insertions. As with the non-verbal categories, the verbal categories were pre-determined based on several textual and discourse analyses of clinical

interviews. Thus, the following active listening verbal subcategorizations were utilised: requests clarification, accurately summarizes patient's commentary, gives reflective questions, gives reflective comments, affirms patient's statements, and uses appropriate language.

The first category, "requests clarification", was further developed to also include "requests confirmation". Thus, active listening verbal acts in this category included those speech acts in which the doctor actively sought out positive or negative confirmation of some aspect of the patient's commentary or otherwise obtained affirmation of a particular comment made by the patient, in the interrogative form. An example of this might include the doctor stating, "you drank 12 bottles of beer a day- is that correct?".

The second form of active listening verbal, "accurately summarizes patient's commentary", was understood as any speech act in which the doctor correctly synopsised something that the patient had just said or had recently said. However, it did not include exact repetitions of the patient's commentary. Although they can indicate some degree of active listening, exact repetitions were not interpreted to be as strong of indicators as the ability to paraphrase, which requires a greater degree of thought than repetition. Furthermore, the summary had to be free of judgement and solely a reflection of the patient's words and intentions. Thus, a simplified example of this may be a patient stating, "I only feel sad at certain times" with the doctor's response being "not all the time then" (with intonation indicating a statement rather than a question). Also, to qualify for this category a doctor may elaborate upon a patient's response without making incorrect statements regarding the patient's illness experience. For example, the doctor may prod a muscle and ask the patient if that prodding hurts. The patient may respond "no", to which the doctor follows with "no effect there". Again, this is not a direct repetition of the patient's words and is also not posed as a question.

The third form of active listening verbal, “gives reflective question” is similar to the first, “requests clarification or confirmation”. However, this act may refer directly to a specific element of the patient’s commentary that occurred in the speech act occurring directly before the doctor’s response. For example, the patient may state that they drink alcohol regularly, to which the doctor may inquire approximately how many bottles of alcohol they drink on weekdays versus weekends. It may also refer to the doctor posing an inquiry that references some previous commentary presented by the patient or it may logically follow from the patient’s immediately previous speech act without being a direct reference to it. For example, the patient may state that their weight is stable to which the doctor may ask “so your appetite is good?”. This question references the previous statement made by the patient but does not seek out a confirmation of the patient’s comment by exactly restating it in question form. It should be noted that this category was originally divided according to these distinctions; however, in some instances the distinction between whether the doctor was asking for specifications of a previous comment as opposed to simply giving a reflecting inquiry generally based upon the previous commentary was unclear. Furthermore, making this minute distinction in the pilot study was not found to drastically impact the overall findings either positively or negatively.

“Gives reflective comment”, the fourth form of active listening verbal, is defined as an instance where the doctor makes a statement pertaining to the patient’s situation, often gathered from the non-verbal attitudes expressed by the patient, without actually paraphrasing or directly repeating a comment that the patient has spoken aloud. This would appear in a scenario such as a patient describing a number of high-stress activities that they are involved in followed by a heavy sigh and slumping of the shoulders and the doctor replying, “it sounds like life is pretty rough for you right now”. This comment is not an imposition of an untrue fact, as was observed

within the category “accurately summarizes patient’s commentary”. However, it goes beyond the previously described category because it is based on an active observation of the non-verbal behaviours expressed by the patient in combination with a mental summarization of the patient’s commentary.

The category “affirms patient’s statements” included two subcategories: single lexical items of affirmation (which may include “no”, depending on the positive or negative quality of the patient’s commentary) and exact repetitions of the whole or part of the patient’s immediately previous statement, which would also indicate an affirmation of what the patient is saying. The former includes items such as “okay”, “alright” and vocalized, non-vocabulary items such as the hum-like sound “mmhmm”. An example of the latter could include the patient stating “It happened when I was 8 years old” to which the doctor states “8” before proceeding with another question or statement and/or indicating the age on the patient’s file.

The final active listening verbal category is “uses appropriate language”. As the definition of “appropriate” can be highly subjective and largely based on the patient’s perception of the doctor’s attitude, individual lexical items or small phrases that may be suggestive of “inclusive” or “politically-correct” language were not analysed for this study. Thus, what were deemed as uses of “appropriate language” were instances in which the doctor had initially used medical jargon or “medicalese” (Norman, Arfai, Brooks, & Eva, 2003) but was met with a confused response from the patient, and so the doctor made the decision to further explain the jargon before moving on with the clinical interview. For example, in one interview the doctor asked, “so you came in today to discuss your hematuria?”. The patient responded, “what’s that?”. Instead of choosing to ignore the patient’s confusion, the doctor explained in lay terms what hematuria was before proceeding with the information intake. Another example of this category

might include the doctor proactively stating the medical terminology or description and immediately describing what is meant by that jargon even before the patient became confused.

In contrast, the following anti-active listening verbal subcategorizations were noted: interrupts patient, questions do not reflect patient's commentary, superimposes facts, and uses inappropriate language. Interrupting the patient was qualified on the basis of the doctor audibly beginning a verbal speech act before the patient had finished their commentary or, more generally, overtaking the patient's turn in the communicative event. "Cutting off" the patient, so to speak, implied that the doctor was perhaps short on time or found that further descriptions from the patient were not necessary; thus, this act removed the patient's agency and belittled their power in the clinical interview context, which qualified interrupting as an anti-active listening trait.

The second anti-active listening verbal category, "questions do not reflect the patient's commentary", can be viewed in direct opposition to their active listening verbal form: "gives reflective questions". These included questions posed by the doctor that had no direct reference to an immediate or recent comment made by the patient nor to the general topic of the conversation in that moment. These questions often appeared as inquiries made in quick succession in order to fulfill a standardized questionnaire during the history-taking phase of the interview without any evident, deeper thought being given to the patient's previous responses. Similar to interruptions, this form of anti-active listening suggests that the doctor is eager to be finished with the clinical interview and apathetic, unwilling to "dig deeper" in order to seek out other underlying symptoms or causes regarding the patient's case. For example, as a doctor was examining a patient's liver problems, he asked, within a 20 second timeframe, if the patient had skin problems, urine problems, or had a stable weight. While all of these questions are necessary

in order to determine the problem, they did not allow for insights from the patient nor did they examine the patient's perspective of their situation. In other cases, this anti-active listening category also included examples of the doctor asking questions in clear ignorance of the patient's previous comments or non-verbal cues. For example, one patient stated that they could not get insurance because of their liver problems but believed that this was a mistake and that they were in good health. This was stated while the patient was sitting with their arms crossed and projecting nervous facial expressions. The doctor responded by asking if the patient was worried about the results of the recent testing.

“Superimposes facts” is a common strategy used by doctors in order to move through clinical interviews more efficiently. Unlike its opposite, active listening category “accurately summarizes patient's commentary”, superimposing facts involves providing a generalized opinion of the commentary that may or may not be a precise representation of the patient's experience. This strategy, although often needed in order to quickly address numerous cases throughout a doctor's shift, reinforces the perceived power held by the doctor and suggests to the patient that their opinion of their own health and illness are invalid. Again, it introduces the perception of the doctor as apathetic towards the patient's needs. For example, a patient was describing a recent break up and stated that it simply “wasn't working out anymore”. The doctor closed the discussion in the following communicative turn by stating “so it was a mutual understanding” before moving onto a new inquiry.

The final anti-active listening verbal category, “uses inappropriate language”, is another category that can be observed as a direct opposite to its active listening verbal form “uses appropriate language”. Here, as in the above active listening description, no attempt to make subjective judgements about the “appropriateness” of speech items will be made. Instead, the

category refers to instances in which the doctor chose to utilise medical jargon that was inaccessible to the patient. However, this category also included instances where the doctor spent extended periods of time explaining symptoms, treatments, etc. often with the incorporation of medical terminology without allowing a break, verbally or non-verbally, for the patient to take a speaking turn. For example, in an analysis of a patient's back, one doctor stated "that would be consistent with where the tenderness was at a lower lumbar region. I'm guessing L4 L5. And the sensation, probably don't need to check for pinprick because I've already seen the sensory deficit on the left L5". In these extended commentaries, the conversational inaccessibility is highly debilitating for the patient and further lowers their social position and agency within the dialogue. In another example, a doctor, in addressing a suicidal patient, stated "It's nice to be able to talk to somebody that is not a friend or a relative so that way you can tell whatever you want to that person. The other part is I want to make sure that you're safe, okay? It sounds like you're going deeper and deeper on your cuts and that can be dangerous. Would you agree that we can have an understanding that you will not cut until I see you back next week? And every week we're going to take it a week at a time that way you have that deadline to say, "okay it's been a week, I can try to control the pain". And if you feel the emotional pain is too strong and you want to cut again, grab the phone. We have phone numbers, distress lines that you can call, you can call my office, my staff. Even a few minutes of talking to somebody sometimes will diffuse that pain". In a sense, this is an extension of the "superimposes facts" category as it also illustrates the doctor's medical knowledge while at the same time removing the patient's voice and agency in the clinical encounter. In this example particularly, the doctor does not wait to see if the patient herself understands that what she's doing is dangerous, and this knowledge is not made clear in other parts of the video recording, nor does the doctor ask if the patient feels

comfortable calling the distress lines. In this case, she is using her medical position to tell the patient how to react, specifically how to react by removing the emotion behind the reaction. In summary, the description of each annotation was coded as depicted in the following table:

Table 1

Coding schema for ALNV, AALNV, ALV and AALV tiers

	Verbal	Non-Verbal
Active Listening Behaviours [ALV] / [ALNV]	<ul style="list-style-type: none"> ▪ Gives reflective commentary [com] ▪ Accurately summarizes patient's commentary [sum] ▪ Gives reflective questions [q] ▪ Affirms patient's statements [aff] ▪ Uses appropriate language (i.e. avoids exclusive jargon, culturally appropriate, and non-judgemental) [al] 	<ul style="list-style-type: none"> ▪ Nods [nod] ▪ Maintains eye contact* [eyes]
Anti-Active Listening Behaviours [AALV] / [AALNV]	<ul style="list-style-type: none"> ▪ Interrupts patient [inter] ▪ Questions do not reflect patient's commentary [nrq] ▪ Superimposes facts [nrsum] ▪ Uses inappropriate language (i.e. makes judgement statements, uses culturally inappropriate language and/or uses exclusive jargon) [aal] 	<ul style="list-style-type: none"> ▪ Avoids extended eye contact* [noeyes] ▪ Does not face patient or chooses to physically turn head away from patient when patient is speaking* [nohead] ▪ Appears distracted (i.e. checks watch or phone) [distract:watch] [distract:phone]
<p><i>Notes.</i> For physical examinations, non-verbal communicative behaviours marked with an asterisk (*) may not be possible.</p>		



Figure 9. Illustration of an OSCE retrieved from a pilot study video (seated from left to right: medical candidate, acting patient, and examiner).

It should also be noted that not all discourse could be codified using the above described schemas as some verbal acts were deemed “neutral”. Neutral verbal acts were those that could be regarded as merely performative or otherwise required as part of the clinical examination. For example, the medical candidate may need to examine a client’s feet and may request that the client point their toes in a certain direction. In this case, the question is neither an act of active or anti-active listening based on the client’s previous verbal act. It is simply a request that must be made in order to progress through the clinical interview. However, if it was apparent that the patient could not access this form of commentary or, conversely, the patient was made aware that they could access the doctor’s speaking turn at any time, the corresponding annotations were marked within the anti-active and active non-verbal listening categories respectively.

3.5.3 Methodology phase 2, part C: Challenges with and revisions of the analytical framework

The initial setup of the ELAN template posed a number of challenges. For organizational purposes, it had initially seemed prudent to create symbolic relationships through the development of three parent tiers: “doctor”, “patient” and “background notes”. The “patient” tier would have only one daughter tier, namely “transcription”. This would allow the coder to keep a record of the patient’s interaction with the doctor strictly on a verbal level, as the scope of this thesis would become too broad if a full analysis of the patient was also conducted. The “doctor” tier would have several daughter tiers, including the following: “transcription”, “active listening verbal” (ALV), “active listening non-verbal” (ALNV), “anti-active listening verbal” (AALV) and “anti-active listening non-verbal” (AALNV). The active and anti-active listening components would allow the coder to record both the details of the speech acts beyond what is stated in the “transcription” line as well as the occurrence and timespan of the non-verbal communicative events, such as eye contact and hand gestures. However, upon creating these tiers and dependencies, it was found that daughter tiers could not be given individual annotations; instead, each daughter tier was assigned the same annotation as its parent tier. As the non-verbal components of the doctor’s interaction would not always align perfectly with their speech, the non-verbal tiers (i.e. “active listening non-verbal” and “anti-active listening non-verbal”) had to become dissociated from their parent tier “doctor” in order to form standalone tiers. However, all tiers were still marked as being associated with the doctor by receiving the code “DOC” before the code assignment associated with the type of listening and type of behaviour.

In addition, the PC used by the coder did not allow for the use of simplified annotation shortcuts. For example, to make a new shortcut would require the coder to first arbitrarily

highlight a section of time then retrieve the “create new annotation here” option from the drop-down menu above. The annotation would then have to be adjusted so that it would more accurately account for the time at which the verbal or non-verbal event occurred. No amount of preferences editing would overcome this technical barrier. Yet, when the coder switched to a Mac in order to cross-examine the ELAN software, the shortcuts were readily available and allowed for faster and more accurate coding. This switch posed a threat on the confidentiality of the videos, as the Mac was not in the personal possession of the coder. Thus, extra care was taken to only save ELAN files on an external hard drive rather than on the MacBook’s internal hard drive. Additionally, following the completion of the coding, all ELAN software was permanently and securely erased from the MacBook to avoid potential confidentiality breaches.

3.5.4 Methodology phase 2, part D: Third-party feedback regarding the piloted analytical framework

Once the framework had been tested and revised by the coder, it could then undergo further approval from third party reviewers. The framework review committee consisted of four randomly selected individuals from outside of the researcher’s field of study and who had no previous knowledge of medical candidate training or assessment so that they would be able to provide relatively unbiased, neutral feedback regarding classification of verbal and non-verbal cues performed by the candidates in the video recordings. Each reviewer was shown a series of archetypal textual entries and images in association with their pre-assigned categorical references. Then, reviewers were asked to indicate whether the category assigned to the text or image was appropriate or not. If the reviewer indicated that the category was not appropriate, they were also asked to explain their reasoning for their observation as well as provide suggestions as to what category label would be more appropriate for the given text or image. The

decision to have reviewers rate the categorization of archetypal texts and images as opposed to having them code the videos using the analysis framework itself was twofold: firstly, it was not practical to seek out reviewers who had an extensive knowledge of ELAN software such that they would be able to proficiently create and label annotations; secondly, the time commitment that would have been required for reviewers to code even a single video would not have been feasible given the time constraints already posited for the current study, as coding a single video could take a minimum of two hours. Overall, feedback regarding the framework was positive. Some suggestions obtained through these informal discussions were considered in terms of the future interpretation of study findings; however, the framework itself required no other revisions beyond those that had already been made by the researcher.

3.6 Methodology phase 3: Implementing the revised framework and data collection workflow

Moving beyond the pilot study, two video sets were used for the main thesis study: supplemental training videos provided by Physicians Apply (an association affiliated with the MCC), which would serve as the control or “baseline” video set for the study; and mock OSCE videos provided by Deidre Lake collected through programming conducted by AIMGA and Communicate2Integrate. A total of 21 videos were analysed for the control group. The control group consisted of native-English speaking medical doctors who have already passed the MCCQE Part II. The videos had been scripted, which restricted the verbal mode of these participants; however, the non-verbal mode was unscripted. The control group was also split into two categories, “illustrative” and “non-illustrative”. The categorization of the videos according to these labels was based on the corresponding feedback provided by the program designers. Of the videos, 10 were purposefully portrayed to demonstrate excellent communication skills, both

in terms of clarity of speaking and listening to the patient. These were accordingly designated as “illustrative” baseline videos. The remaining 11 videos, which were noted to exhibit poor listening and speaking qualities, were then assigned to the category of “non-illustrative” baseline videos. It should be noted that while the videos exemplified good and bad communication skills, the doctors in each video variant were all able to arrive at the same generalized diagnoses of patients’ complaints. However, the extent to which the doctors were able to understand the patient’s individual situation and thus apply a more individualistic treatment plan, was largely based on the doctor’s communication abilities. Thus, communication proficiency was the key focus of each of these videos. In addition, these videos ranged in time duration from 30 seconds to over four minutes, depending, to some extent, on the “illustrative” or “non-illustrative” nature of the video (i.e. “non-illustrative” videos tended to be shorter than their “illustrative” counterparts).

The test group, on the other hand, was comprised of the videos received from the project’s partner C2I and AIMGA. A total of 5 participants, each of whom spoke English as a second language and had not yet passed the MCCQE Part II, consented to having their videos analysed. Each participant had 5, 10-minute long, clinical scenarios or mock OSCEs recorded as part of their preparation for the MCCQE Part II. The 5 scenarios were identical for the 5 candidates. The scenarios included: one consultation regarding a physical ailment, two history-taking and physical assessment situations, one consultation and treatment plan for a mental ailment, and one consultation and treatment plan for a terminal illness (otherwise known as “breaking bad news”).

With the revised and finalized tier framework, the next step in the workflow was to create ELAN files using the WAV and video files for the videos in both sets, as was done in the pilot

study. The established framework allowed for automatic tier framework assignment in each file, which was uploaded following the insertion of the WAV and video files during the initial creation of the ELAN file. ELAN files were saved based upon the original source of the video, the medical candidate participating in the video (under the guise of their pre-assigned identifier code or, in the case of the Physicians Apply videos, the code DOCR followed by a number was assigned as the doctors were already anonymous) and which video it was out of the total number of videos pertaining to the candidate (if they participated in more than one video) (e.g. Lake-Doc1(Spanish)-1of5). In addition, the files were assigned the timestamp submission of 1969:07:20 00:00, as described within the ethics application.

Initially, the workflow was outlined to be that all videos in both subgroups of the control set were analysed one at a time according to the behaviour under review. For example, all instances of eye contact were annotated across all videos prior to moving forward with annotations of all hand gestures. This would largely eliminate the possibility for confusion regarding whether a situation required an annotation or could be left unmarked given the goals of the analysis. However, the videos from the project partner at AIMGA and C2I were received much later than the Physicians Apply videos. As a result, the annotations and subsequent basic transcriptions were established for the control group videos from Physicians Apply before the same workflow was applied to the test group videos from the project's partner. While coding the test group videos however, the Physicians Apply videos were continuously cross-checked in order to maintain the highest level of consistency across all videos. Although all coding for the main study was conducted by the researcher, blind randomization of videos was used as frequently as possible throughout creation of annotations and transcriptions in order to lessen the

potential for reviewer bias. However, bias may still negatively impact the results of the study and thus will be further acknowledged as a study limitation in the concluding chapter of this project.

Once both videos sets were at the same stage of analysis, it was possible to begin assigning more detailed transcriptions to the annotations. Again, for the sake of consistency, each trait was coded across all videos before a new trait was analysed. All non-verbal features were coded first, ending with hand gestures. Hand gestures were analysed last so that the coder was able to become extremely familiar with the non-verbal idiosyncrasies of each doctor and thus most accurately determine whether the hand gesture was symbiotic (i.e. categorized as “iconic”, “metaphoric”, “deictic” or “interactive”) or was merely acting as a placeholder or unmeaningful, rhythmic linguistic marker (i.e. a “beat”). The non-verbal coding was followed by the verbal coding. Although initially it was determined that this stage could be assessed on a single computer screen, hard copies of all video transcripts were developed instead. This was done in part as a break from the computer screen for the researcher in order to not make unnecessary errors as well as for the ease that flipping through pages of text and making changes provided in comparison to scrolling through text on a screen. Thus, the texts were analysed for all verbal codes occurring in one video transcript at a time. Unlike the non-verbal cues, the verbal codes were much more nuanced in that the interpretation of a comment or question had to be ascertained from the text preceding it. It was not possible to simply analyse one single feature at a time across all videos as had been done in the non-verbal coding stage.

Chapter 4: Results and Analysis

4.1 Introduction

This chapter will provide an overview of the analysis and results as they pertain to existing literature regarding multimodal active listening and the assessment of patient-centred care in Canadian medical clinics. As previously stated, the goal of this study was to reconnoitre the presence or absence of active listening traits performed by native and non-native English-speaking medical candidates undergoing training programs for the second portion of the Canadian medical qualifying exam, the MCCQE Part II, through the lens of multimodality. In addition, the study aimed to realize the prevalence of individual traits in relation to the medical candidate's overall appearance of active or anti-active listening and determine whether an increased display of active listening traits correlated positively with increased collaboration with the patient and subsequent ability to create individualized treatment plans and diagnoses. Thus, this chapter is organized according to the guiding research questions previously outlined, beginning with an initial discussion of the micro and macro depictions of active listening followed by an application of this characterization to MCCQE Part II success.

4.2 Research question 1: Analysing NS medical candidates' and NNS medical candidates' performances of individual active and anti-active listening traits

4.2.1 Establishing assumptions and statistical approach for research question 1

The first guiding research question was given as follows: how are active and anti-active listening traits realized and used on a micro level by native English-speaking (NS) and non-native English-speaking (NNS) medical doctors in order to create a collaborative, empathetic relationship with patients during Canadian Objective Structured Clinical Examinations?

Based on the study's literature review findings, the following assumptions were made in relation to the first research question:

- 1) Active listening is characterized by the presence of some baseline quantity of positive verbal and non-verbal traits, while anti-active listening is defined by a baseline quantity of negative verbal and non-verbal traits.
- 2) From Assumption 1, it follows that the patient's evaluation of whether a doctor is actively listening or not, as evidenced by their willingness to share details regarding their illness experience, will depend upon the overall quantity of positive versus negative traits, fluctuating based upon the number of occurrences of those polarized traits in a given conversation.
- 3) Non-native English-speaking (NNS) doctors will demonstrate a greater quantity of negative verbal traits than native English-speaking (NS) doctors as a result of linguistic barriers (e.g. the increased need for clarification may result in more frequent interruptions) and sociocultural artifacts from their first languages (e.g. overlaps in speaking turns not resulting from a need for linguistic clarification).
- 4) As a result of Assumption 3, NNS doctors will overcompensate in regard to frequency of positive, non-verbal behaviours while NS doctors will maintain a lower, but still socially acceptable, display of positive non-verbal behaviours.

In order to begin considering these assumptions, the data was tabulated and compared for statistical differences. In terms of tabulating the findings in this initial analytical phase, only two categories were recognized in terms of both time duration and frequency of occurrence, while the remaining categories were only tabulated for frequency of occurrence, despite having the ability to timestamp these annotations within ELAN. The reason for this was that some items (e.g. hand

gestures) were both extremely difficult to capture accurately in terms of the timing of the trait's onset and finish and not useful for the objective of the present study. Thus, it was more pertinent to only analyse the number of times the candidate performed a given trait in a given speaking turn or, more critically, outside of their speaking turn. However, for eye contact and the lack of eye contact (when required by the speaker), both time duration and frequency were important because of their interrelated nature, which will be explored further in the subsequent Discussion chapter. Some candidates might only maintain eye contact for small periods of time but frequently attempt to make eye contact, while others may seek eye contact less frequently yet maintain contact for a longer period of time. The frequency and duration of eye contact will also fluctuate based upon the conversational topic of the clinical interaction. For example, a physical exam may necessarily elicit shorter, less frequent eye contact while a scenario in which bad news is being delivered may warrant more frequent, sustained eye contact. Thus, all verbal and non-verbal cues, with the exception of eye contact, were tabulated on the basis of frequency.

Although the test group videos were consistent in regard to time duration, being 10 minutes in length each, the control group video sets were highly variable, and none were close to 10 minutes in length. As a result, it was necessary to stabilize the tabulations of all items across all videos. For the items that were coded according to occurrence, this was initially done by dividing the number of occurrences (i.e. frequency) by the time duration in seconds in order to arrive at a ratio of occurrence per second. Unfortunately, this was not an ideal method of achieving numerical stabilization, as the resultant numbers were extremely small, differing only by decimal points. The alternative to creating these ratios was to select an arbitrary time duration such as 120 seconds and extrapolate the number of item occurrences that would be predicted to occur based on the actual occurrences in a given timeframe. However, this was determined to be even

more statistically inconsistent and unreliable than the previous stabilization approach. Thus, the creation of occurrence per second ratios was selected as the most accurate comparative format and all values were rounded to three decimal points in order to best reflect the relationship between the three groups (as opposed to rounding to the nearest whole number in an attempt to reflect the impossibility of having only a fraction of a behaviour, such as a hand gesture). However, ideally, all videos would be of the same time duration in order to compare the behaviours with greater accuracy, which will be highlighted as a study limitation in the Conclusions chapter. For the time durations of eye contact or lack thereof, this problem was eliminated, as a simple division of the item's time duration by the total duration of the video resulted in a percentage of time rather than a ratio of different categories and provided consistent ratio comparisons across all videos.

Once the frequencies of individual traits were relatively stabilized and compiled within each video set (e.g. all occurrences of the deictic hand gesture were stabilized and tabulated within the illustrative control group, non-illustrative control group and test group video sets separately), they underwent an initial descriptive statistical analysis in order to better understand which traits were more salient than others, both within videos sets and when video sets were compared. These traits were analysed within the four broad categories of non-verbal active listening, verbal active listening, non-verbal anti-active listening and verbal anti-active listening and then compared based upon mode type. After the initial descriptive statistics, a nonparametric test was applied which assessed the presence or absence of significant differences of individual traits between video sets. Using the non-verbal active listening quality of "head nods" as an example, this resulted in the following three comparative pairings:

- 1) All illustrative head nods in comparison to all non-illustrative head nods

- 2) All illustrative head nods in comparison to all test group head nods
- 3) All non-illustrative head nods in comparison to all test group head nods

4.2.2 Performances of non-verbal active and anti-active listening traits

As stated in the previous section, an initial descriptive statistical analysis was used to highlight key trends in data collected across the verbal and non-verbal modes within the active and anti-active listening categories. In the data collected from the 25 OSCEs performed by the 5 participants in the test group, an average total of 27.148 non-verbal active listening traits were identified as being performed per video (as opposed to a summary of all non-verbal active listening traits performed by a participant in all 5 OSCE video recordings). Within this set, the category of head nods occurred most frequently, comprising 11.050 of the total non-verbal active listening traits. In terms of most salient categories, this was followed by frequent eye contact occurrences, which were displayed 6.86 times per video on average, and then deictic hand gestures, which occurred an average of 2.300 times per video. In the average 29.563 non-verbal active listening items performed per video by the 11 participants in the 11 videos of the illustrative control group, eye contact occurred most frequently, appearing 11.115 times. Following this category, head nods was found to be the second most frequently produced category (8.669 times on average). Finally, in the data collected from the 10 OSCEs performed by 10 participants in the non-illustrative control group, an average total of 26.333 non-verbal active listening traits were identified per video. Eye contact was also the most frequently performed category (13.837 times on average) followed by head nods (3.229 times per video on average) and then deictic hand gestures (2.713 times per video on average). See Table 2 for a complete list of non-verbal active listening categories and corresponding number of occurrences found in the test group, illustrative control group and non-illustrative control group.

Table 2

Non-verbal active listening traits identified in the test group, illustrative control group and non-illustrative control group

Non-verbal active listening trait type	Average # of occurrences within a single test group video	Average # of occurrences within a single illustrative control group video	Average # of occurrences within a single non-illustrative control group video
Head nods	11.05	8.669	3.229
Eye contact (average duration)	3.380	5.870	4.630
Eye contact (average # of occurrences)	6.860	11.115	13.837
Iconic hand gestures	1.446	0.230	0.259
Metaphoric hand gestures	1.259	1.267	1.098
Deictic hand gestures	2.300	2.039	2.713
Interactive hand gestures	0.853	0.373	0.567
<i>Notes.</i> All non-verbal active listening trait types were coded according to frequency of performance with the exception of eye contact, which was coded for duration in addition to frequency (as indicated above).			

In the non-verbal anti-active listening data collected from the 25 videos analysed in the test group, an average total of 2.547 traits per video were identified. The category of “hands demonstrating distraction” (e.g. fidgeting with hair) occurred most frequently, comprising 11.050 of the total non-verbal anti-active listening traits. In terms of most salient categories, this was followed by a lack of eye contact when required, for which each video averaged a total of 0.563. In the average 3.656 non-verbal anti-active listening items found in the 11 videos comprising the illustrative control group, the category of “hands demonstrating distraction” also occurred most frequently, appearing an average of 1.784 times per video. Following this category, a lack of eye contact when required was also found to be the second most frequently produced category, being produced 1.117 times on average per video. Finally, in the data collected from the non-

illustrative control group, which was composed of 10 OSCE video recordings, an average total of 11.524 non-verbal anti-active listening traits were identified per video. For this group, a lack of eye contact when required was the most frequently performed category (4.347 times on average) followed by a visible turn of the head away from the participant (3.150 times per video on average). See Table 3 for a complete list of non-verbal anti-active listening categories and corresponding number of occurrences found in the test group, illustrative control group and non-illustrative control group.

Table 3

Non-verbal anti-active listening traits identified in the test group, illustrative control group and non-illustrative control group

Non-verbal anti-active listening trait type	Average # of occurrences within a single test group video	Average # of occurrences within a single illustrative control group video	Average # of occurrences within a single non-illustrative control group video
Head visibly turned away from patient	0.4169	0.102	3.15
Lack of eye contact when required (duration)	0.3145	0.6525	1.932
Lack of eye contact when required (occurrences)	0.563	1.117	4.347
Hands demonstrate distraction	1.253	1.784	2.095
<i>Notes.</i> All non-verbal anti-active listening trait types were coded according to frequency of performance with the exception of lack of eye contact, which was coded for duration in addition to frequency (as indicated above).			

In summary, head nods occurred most frequently in regard to active listening traits within the illustrative control group-test group comparative pairing. Still, the test group did not produce high frequencies of eye contact duration and frequency, which accompanied salient non-verbal, anti-active listening traits of hand distractions and lacking eye contact when eye contact was

required. These traits overall reflect the test group participants' overreliance on taking notes during patient storytelling, perhaps as a means of overcoming linguistic barriers. Although the present study focuses on medical interactions, these findings align with previous research that has analyzed the performance of such positive listening traits in similarly formal interactional settings. For example, Bodie et al. (2015) observed the impact of individual verbal and nonverbal listening traits on speakers' perceptions of listeners' helpfulness, sensitivity and supportiveness. In this study, 171 undergraduate students were asked to disclose an upsetting problem to either a trained active listener or an untrained listener. It was found that nonverbal behaviours including head nods, smiling, body orientation, animation, voice and open body posture had little to no impact, individually, on the positive reception of the listener by the speaker. However, facial expressions generally and eye contact significantly contributed to the perception of listeners' affect improvement. Additionally, previous research has recognized the negative impact of eye contact avoidance and the overbearing nature of the overproduction of hand gestures within North American culture, particularly when facts regarding a specific scenario are urgently being sought out (cf. Sue & Sue, 1977; McCarthy, Lee, Itakura, & Muir, 2006; Kita, 2009). Therefore, the findings of the present study support the argument that the non-verbal component of active listening did not necessitate a large production of positive cues but more critically the avoidance of particularly noticeable negative cues which imply that the listener is distracted or otherwise not engaged in the dialogue.

4.2.3 Performances of verbal active and anti-active listening traits

In terms of verbal active listening traits, the data collected from the 5 participants across 25 videos in the test group showed an average total of 17.559 items per video. The category of affirmations occurred most frequently, comprising 10.378 of the total verbal active listening

traits. In terms of most salient categories, this was followed by frequent, mindful questions, for which participants in each video averaged a total of 3.840 in a given interaction. In the average 12.631 verbal active listening items performed by the 11 participants in the illustrative control group, mindful questions occurred most frequently, appearing 4.210 times per interaction. Subsequently, affirmations were found to be the second most frequently produced item, being produced 3.961 times per interaction on average. Finally, in the data collected from the 10 participants in the non-illustrative control group, an average total of 10.806 verbal active listening traits were identified per interaction. Affirmation was also the most frequently performed item in this control group (3.479 times on average), followed in terms of saliency by mindful questions (2.960 times per participant on average). See Table 4 for a complete list of verbal active listening categories and corresponding number of occurrences found in the test group, illustrative control group and non-illustrative control group.

Table 4

Verbal active listening traits identified in the test group, illustrative control group and non-illustrative control group

Verbal active listening trait type	Average # of occurrences within a single test group video	Average # of occurrences within a single illustrative control group video	Average # of occurrences within a single non-illustrative control group video
Reflective question	3.840	4.210	2.960
Reflective comment	0.960	1.160	0.410
Accurate summary	0.427	0.756	0.784
Repetition	0.485	0.573	0.973
Affirmation	10.378	3.961	3.479
Accessible general medical commentary	1.469	1.971	2.200

It should be noted that some key verbal acts either could not be classified within one of these broader categories or were particularly marked within an existing category and required

further discussion. First, some doctors demonstrated the ability to interpret patients' underlying emotions in relation to specific aspects of their illness narrative strictly through individual non-verbal manifestations of that emotion performed by the patient. For example, one doctor noticed a patient's avoidance of eye contact while discussing the patient's recent mood swings and was able to further expand upon the topic in order to realize that the patient was suffering from depression. A second strategy for building an empathetic relationship with the patient was accomplished through the repeated use of the patient's name throughout the dialogue, which not only provided a conversational opening for the patient to provide feedback but also recognized the role of the patient within their own illness experience. Finally, one doctor recognized how long periods of note-taking could lead patients to stop speaking. In order to combat this anti-active listening cue, the doctor verbally acknowledged that he would be taking notes throughout the narrative and confirmed with the patient whether or not his note-taking would be acceptable.

In terms of verbal anti-active listening traits, the data collected from the 5 participants across 25 videos in the test group showed an average total of 1.985 items per video. The category of rapid-fire and/or unrelated questions occurred most frequently (1.160 times on average per video). This category was followed in terms of saliency by frequent interruptions of the patient's storytelling, for which each participant averaged a total of 0.338 interruptions. Conversely, in the average 0.338 verbal anti-active listening items performed by the 11 participants in the illustrative control group, interruptions appeared most frequently (3.040 times). Following this category, inaccurate summaries of patients' stories were the only other verbal anti-active listening category produced, being produced 0.034 times on average. Finally, in the data collected from the 10 participants in the non-illustrative control group, an average total of 4.396 verbal anti-active listening traits were identified per participant. Interruptions were also the most

frequently performed item (1.321 times on average) followed by inaccurate summaries of patients' stories (1.295 times per participant on average). See Table 5 for a complete list of verbal anti-active listening categories and corresponding number of occurrences found in the test group, illustrative control group and non-illustrative control group.

Table 5

Verbal anti-active listening traits identified in the test group, illustrative control group and non-illustrative control group

Verbal anti-active listening trait type	Average # of occurrences within a single test group video	Average # of occurrences within a single illustrative control group video	Average # of occurrences within a single non-illustrative control group video
Rapid-fire and/or unrelated question	1.160	0	0.54
Unrelated and/or unempathetic comment	0.088	0	1.24
Inaccurate summary	0.081	0.034	1.295
Interruption	0.338	0.304	1.321
Inaccessible general medical commentary	0.160	0	0
Inappropriate language use	0.158	0	0

Similar to their verbal active listening counterparts, some behaviours within the anti-active listening verbal analysis were particularly outstanding within their category or could not otherwise be included within the confines of the established categories. For example, one NNS doctor repeatedly used the phrase “I know” as a means of connecting with the patient’s narrative. Although his use of this lexical bundle was, from a third-party viewpoint, a result of attempting to be empathetic, the Canadian interpretation of this phrase was not one of agreement but rather an opening to a conversational turn pertaining to facts known by the speaker. As such, the phrase was often misinterpreted by the patients as the doctor being condescending, evidenced by an

abrupt end to their speaking turn and often accompanied by non-verbal behaviours indicating attitudes of mild disbelief or shock. Additionally, a NNS doctor demonstrated blatant anti-active listening by verbally avoiding a critical comment posed by a patient, which led him to misdiagnose and mistreat the patient at the end of the clinical interview.

In summary, the NNS test group was relatively unsuccessful in establishing common ground with the patients largely as a result of their inability to summarize or repeat key elements of patients' stories while at the same time verbally barring access to further conversation, especially when compared to the illustrative control group. Several studies in the medical domain have examined the impact of verbal positive and negative listening traits performed by doctors in relation to doctors' ability to collaborate with patients. For example, Fassaert et al. (2007) conducted a study of 524 videotaped consultations in which third-party viewers evaluated the physicians' active listening skills using a newly developed, Likert-scale style questionnaire known as the ALOS-global. In addition, patients were asked to fill out questionnaires before and immediately after their consultations in order to assess their anxiety levels and the physicians' performance in general. Although the study focused strictly on consultations pertaining to minor physical ailments, the findings showed a significant correlation between active listening, as measured by the ALOS-global analytics, and the category of verbal attention, which specifically measured statements reflecting attention and a recognition of the collaboration between the doctor and the patient. In contrast, Henzel (1989), after a longitudinal study of physicians' interactions with patients in California, found that a number of negative verbal listening cues are used to create power disparity between the doctor and the patient. The study specifically examined the following negative listening categories: manipulating the use of professional register, specifically lexicon, which prevents patients from fully understanding their diagnoses;

eliminating patient agency, particularly in regard to their body parts and medical actions; preference for yes/no and tag questions as opposed to open-ended questions; and other imbalances of turn-taking within the dialogue that restrict the patient from speaking and having access to choice of response. Each of these linguistic techniques were found to subordinate the agency of the patient in the conversation and lessen the overall demonstration of active listening on the part of the doctor. Therefore, the findings of the present study support the literature pertaining to the role of positive and negative listening cues in relation to finding common ground with patients and facilitating the elicitation of further information through the creation of a power-balanced relationship.

As described above, following this initial descriptive analysis, a Kruskal-Wallis analysis was performed across all pairings due to the nonparametric nature of the study, with particular attention given to those traits that were most marked within the descriptive statistics findings. However, as Appendix G illustrates, no statistically significant relationships (i.e. $< .05$) appeared when analysing single trait frequencies and/or durations across the various video sets, with the exception of beats and eye contact duration and frequency, which will be further assessed in the subsequent section. This lack of statistical significance was likely due in part to the small sample size used in this study. Furthermore, the repeated Kruskal-Wallis tests across all of these measures were not an effective form of statistical analysis and it is likely that any occurrence of statistical significance by this approach occurred merely by chance, introducing a type II statistical error as well as an inflated family-wise error rate. Thus, the values presented in Appendix G should be viewed primarily as exploratory in nature.

4.3 Research question 2: A macro level analysis of NS medical doctors' and NNS medical doctors' performances of active and anti-active listening

4.3.1 Establishing assumptions and the statistical approach for research question 2

The second guiding research question was given as follows: how are active and anti-active listening traits realized and used on a macro level by native and non-native speaking medical candidates in order to create a collaborative, empathetic relationship with patients during Objective Structured Clinical Examinations?

Overall, the inaccuracies posed by the initial phase of statistical analysis discussed in the previous section suggests that positive active listening traits on their own do not account for the quality of listening demonstrated. For example, nodding frequently throughout the dialogue was not the sole causal determinant of doctors appearing to be listening actively. On the other hand, some negative traits, although still not appearing frequently enough to be statistically significant, were highly marked in the data and did have an impact on the ability of the doctor to accurately prescribe a treatment plan or make a diagnosis. Thus, from these findings, the assumptions could be further modified and re-tested for accuracy. Overall, in the comparison of individual traits, the biggest differences appeared to occur not when the same trait was analysed across all data sets but rather when positive traits as a whole were compared to negative traits collectively. Thus, the assumptions were reduced and refined in the following manner:

- 1) The greater quality of active listening is not necessarily defined on the basis of an overall greater quantity of individual positive traits in either mode but rather on the salient presence and quantity of negative listening traits. That is, a relative absence of negative verbal and non-verbal listening traits would suggest a high-quality demonstration of active listening.

- 2) NNS doctors will more frequently demonstrate these negative qualities than NS doctors as a result of consciously or subconsciously attempting to overcome cultural and linguistic barriers within the multilingual, multicultural clinical interview setting.

In this phase of analysis then, all of the verbal and non-verbal components of active listening as well as those within the anti-active listening category found in each video were compiled as two distinct sets within each video set (e.g. all of the positive verbal and non-verbal traits in the illustrative control group were compared to those within the non-illustrative control group and the test group, followed by the negative traits, etc.) and then compared using descriptive statistics, Kruskal-Wallis tests and Dunn's post-hoc tests.

4.3.2 Macro level active and anti-active listening results and analysis

In regard to the initial descriptive statistics, when nearly all positive active listening traits were taken as a collective and stabilized for video length differences, as demonstrated in Table 6, the NNS test group performed 340.584 traits, while the illustrative and non-illustrative control groups performed 246.577 and 185.791 traits, respectively. Yet the test group was still not always successful in their ability to gather key information from patients in order to arrive at appropriate, non-generic diagnoses and treatment plans, as will be further discussed in the subsequent section. This suggests that the barriers to collaborative dialogue lay not in the active listening performance of doctors but rather in the saliency of their anti-active listening trait performance. Indeed, as seen in Table 7, the results showed that the illustrative control group, which was expected to be the standard of excellent communication skills, demonstrated only 19.200 anti-active listening traits (following a stabilization of the total trait performance in order to account for variance in the video length), while the test group and non-illustrative control group performed 36.067 and 91.391 traits, respectively (similarly following a process of

stabilization for video length variance). Although the test group results are closer in number to the illustrative control group than to the non-illustrative control group, their increased presence and saliency appears to be impacting their overall OSCE performance. Further depiction of the positive and negative listening set descriptive statistics may be seen in Appendices H and I.

Table 6

A summary of positive (active) listening traits across both control groups and the test group

Positive (active) listening traits in test group		Positive (active) listening traits in illustrative control group		Positive (active) listening traits in non-illustrative control group	
Total # of occurrences	Total # of occurrences stabilized for comparison	Total # of occurrences	Total # of occurrences stabilized for comparison	Total # of occurrences	Total # of occurrences stabilized for comparison
817.402	340.584	246.577	246.577	203.807	185.791
<i>Notes.</i> “Stabilized for comparison” indicates that the total # of occurrences have been modified in order to represent the number of occurrences that statistically would have appeared across 11 videos. The modification does not appear to impact the illustrative control group as it was already comprised of 11 videos.					

Table 7

A summary of negative (anti-active) listening traits across both control groups and the test group

	Negative (anti-active) listening traits in test group		Negative (anti-active) illustrative control group		Negative (anti-active) non-illustrative control group	
	Total # of occurrences	Total # of occurrences stabilized for comparison	Total # of occurrences	Total # of occurrences stabilized for comparison	Total # of occurrences	Total # of occurrences stabilized for comparison
Anti-active listening traits	86.560	36.067	19.200	19.200	100.530	91.391
<i>Notes.</i> “Stabilized for comparison” indicates that the total # of occurrences have been modified in order to represent the number of occurrences that statistically would have appeared across 11 videos. The modification does not appear to impact the illustrative control group as it was already comprised of 11 videos.						

Following this descriptive statistical analysis, differences between the groups in these two larger data sets were compared through the use of a Kruskal-Wallis test because of the non-parametric nature of the data. The Kruskal-Wallis test showed that there was a statistically significant difference in the number of positive traits between the three groups, $\chi^2(2) = 18.05$, $p < .001$, with a mean rank score of 24.66 for the illustrative control group, 18.53 for the non-illustrative control group and 34.06 for the test group. Additional post-hoc comparisons using Dunn's test indicated that there were significant differences ($p < .05$) in the means between the illustrative control group and the test group ($p_{\text{bonf}} = .021$) and in the means between the non-illustrative control group and the test group ($p_{\text{bonf}} < .001$) (see Table 8 for details).

Table 8

Dunn's post hoc test results for positive set

Source 1	Source 2	z	W_i	W_j	p	p_{bonf}
Illustrative control group	Non-illustrative control group	1.248	18.25	11.09	0.106	0.318
Illustrative control group	Test group	-2.466	18.25	30.44	0.007	0.021
Non-illustrative control group	Test group	-4.046	11.09	30.44	< .001	< .001

In terms of the relationship between negative trait sets between the control groups and test group, the Kruskal-Wallis showed that there was a statistically significant difference between the three groups, $\chi^2(2) = 14.31$, $p < .001$, with a mean rank score of 1.920 for the illustrative control group, 9.139 for the non-illustrative control group and 3.607 for the test group. Furthermore, post-hoc comparisons using Dunn's test indicated that there was a significant difference ($p < .05$)

in the means between the illustrative control group and the non-illustrative control group ($p_{\text{bonf}} < .001$) as well as in the means between the non-illustrative control group and the test group ($p_{\text{bonf}} = .005$) (see Table 9 for details).

Table 9

Dunn's post hoc test results for negative set

Source 1	Source 2	z	W_i	W_j	p	p_{bonf}
Illustrative control group	Non-illustrative control group	-3.641	14.20	35.09	< .001	< .001
Illustrative control group	Test group	-1.401	14.20	21.13	0.081	0.242
Non-illustrative control group	Test group	2.921	35.09	21.13	0.002	0.005

Although these “positive” and “negative” categories were comprised of the majority of individual traits analysed across modes within this study, there were three exceptions to the formulation of this diametric comparison: beats hand gestures, eye contact duration and eye contact frequency. Beats were not initially included in the positive or negative data set as they were not believed to have contributed meaningfully to the content of the conversation. However, while their primary role includes acting as conversational tempo markers and placeholders, beats were also observed as a means of interrupting a speaking turn, which was notably distracting to patients. Thus, it could be argued that beats can in fact be categorized as negative attributes of listening. However, the practice of using beats as an interruption of speech was utilised solely by the test group and not in the illustrative or non-illustrative control group videos. This appeared to suggest that the overuse of beats was related to the listener’s language proficiency and perhaps

cultural background rather than being attributable to the category of anti-active listening and were thus analysed separately from the amassed positive (active) and negative (anti-active) listening traits.

Eye contact duration and frequency were also analysed apart from the larger positive and negative data sets. These two traits were negatively correlated with each other and, as a result, were more interrelated than any other traits under review. In some cases, making more frequent eye contact but for shorter time durations was appropriate, while in other interactions, making less frequent eye contact but maintaining those eye contact moments for longer periods of time was appropriate. The selection of one form of eye contact maintenance and duration over another was largely based on the context of the medical interaction. Thus, this connection between duration and frequency needed to be accounted for in the statistics phase of the analysis through direct, distinct comparison rather than being compiled into the larger data sets. The descriptive statistics results for beats as well as eye contact duration and frequency are further depicted in Appendices J-L.

As with the larger positive and negative data sets, differences between video groups in terms of the hand gesture “beats” were compared through the use of a Kruskal-Wallis test because of the non-parametric nature of the data. The Kruskal-Wallis test showed that there was a statistically significant difference in the number of beat performances between the three groups, $\chi^2(2) = 9.249$, $p = .010$, with a mean rank score of 4.739 for the illustrative control group, 5.547 for the non-illustrative control group and 8.477 for the test group. Additionally, post-hoc comparisons using Dunn’s test indicated that there were significant differences ($p < .05$) in the means between the illustrative control group and the test group ($p_{\text{bonf}} = .010$) as well as in the

means between the non-illustrative control group and the test group ($p_{\text{bonf}} = .048$) (see Table 10 for details).

Table 10

Dunn's post hoc test results for beats hand gestures

Source 1	Source 2	z	W_i	W_j	p	p_{bonf}
Illustrative control group	Non-illustrative control group	-0.545	15.10	18.23	0.293	0.879
Illustrative control group	Test group	-2.707	15.10	28.48	0.003	0.010
Non-illustrative control group	Test group	-2.144	18.23	28.48	0.016	0.048

Differences between the groups in terms of eye contact frequency were also compared through the use of a Kruskal-Wallis test. The test showed that there was a statistically significant difference in the eye contact frequency between the three groups, $\chi^2(2) = 10.48$, $p = .005$, with a mean rank score of 11.12 for the illustrative control group, 13.84 for the non-illustrative control group and 6.67 for the test group. Post-hoc comparisons using Dunn's test also indicated that there was a significant difference ($p < .05$) in the means between the non-illustrative control group and the test group ($p_{\text{bonf}} = .004$) (see Table 11 for details).

Table 11

Dunn's post hoc test results for eye contact frequency

Source 1	Source 2	z	W_i	W_j	p	p_{bonf}
Illustrative control group	Non-illustrative control group	-0.789	27.20	31.73	0.215	0.645
Illustrative control group	Test group	2.013	27.20	17.25	0.022	0.066
Non-illustrative control group	Test group	3.027	31.73	17.25	0.001	0.004

Finally, a Kruskal-Wallis test was also used to examine the differences between the groups in terms of eye contact duration. The Kruskal-Wallis test showed that there was a statistically significant difference in the duration of eye contact between the three groups, $\chi^2(2) = 9.642$, $p = .008$, with a mean rank score of 5.87 for the illustrative control group, 4.63 for the non-illustrative control group and 3.46 for the test group. Additionally, post-hoc comparisons using Dunn's test indicated that there was a significant difference ($p < .05$) in the means between the illustrative control group and the test group ($p_{\text{bonf}} = .004$) (see Table 12 for details).

Table 12

Eye contact duration Dunn's post hoc test results

Source 1	Source 2	z	W_i	W_j	p	p_{bonf}
Illustrative control group	Non-illustrative control group	1.247	32.70	25.55	0.106	0.319
Illustrative control group	Test group	3.016	32.70	17.79	0.001	0.004
Non-illustrative control group	Test group	1.621	25.55	17.79	0.052	0.157

From these findings, it can be seen that significant differences, or more specifically the noticeability of the quality of listening, are determined on the basis of an absence of negative traits rather than a significant quantity of positive traits. In particular, when these findings are related to the findings for the first research question, inaccurate summaries and insensitive comments within the verbal category as well as the lack of eye contact, primarily as a result of focusing too much on writing, within the non-verbal category were highly salient, evidenced by their greater contribution to these findings. In contrast, provided that the traits were performed consistently by each doctor, there was no upper limit as to how often the positive qualities had to be performed in order to be classified as listening actively. Therefore, it was important that the doctors avoided performing salient, negative behaviours that were associated with anti-active listening in order to indicate that they were engaged with the speaker's prose.

These findings of the performance of active listening offer a different but complimentary approach to the active listening literature found both within and outside of the medical domain. For example, Bodie and Jones (2012) asked 383 participants to assess 5-minute conversations in

which one person disclosed an emotionally upsetting event to another person, who provided varying degrees of verbal and nonverbal support. Although both verbal and nonverbal support, as individual metrics, were found to correlate with viewers' judgements of active listening, when high levels of both verbal and nonverbal cues were demonstrated, the supporters were rated more highly as active listeners by the viewers. While this study focuses on positive listening cues, the findings still imply that a multimodal demonstration of cues, rather than a demonstration of cues within a single mode, is key to construing excellent active listening qualities. Likewise, Vogel, Meyer and Harendza (2018) analysed final-year undergraduate medical students' verbal and non-verbal communicative proficiency during history-taking interviews, particularly in relation to perceptions of empathy and correlation to gender. Video recordings of 30 students in 5 simulated clinical scenarios were analysed in two ways: external ratings using a newly developed observation scale of verbal and non-verbal communication and CARE-questionnaires to garner information pertaining to perceptions of empathy. The results suggested that verbal and non-verbal communicative factors contribute meaningfully but in differing ways to the overall impression of active listening as well as perceptions of empathy between male and female students. Both studies, reflective of the majority of active listening literature, approached the definition of active listening as strictly a presentation of positive listening cues without acknowledging the absence of negative cues as having an equal if not greater impact on the speaker's impression of the listener's active listening ability. However, these studies still compliment the present research findings as well as a smaller distribution of active listening studies (cf. Henzel, 1989) in that they do not suggest that there is some baseline quantity of positive cues to be performed as an active listener, but rather implicitly show an absence of negative cues as key to active listening.

4.4 Research question 3: Realizing the relationship between active listening performance and OSCE success

The third guiding research question was given as follows: how does a cross-cultural understanding of active listening impact OSCE success and inform subsequent training programs in Canada?

Although results regarding the test group participants' official mock OSCE scores were not made available to the researcher, an analysis of the commentary between the doctor and the reviewer recorded following each mock OSCE scenario gave some indication as to whether the physician was successful in identifying key problems posed by the patient and thus able to provide appropriate, corresponding diagnosis, or if the doctor was only able to provide a generalized or inaccurate report of the patient's situation. As seen in Table 16, the 5 test group participants were able to arrive at generalized diagnoses in all 25 OSCEs. However, in 36% of the scenarios, the reviewer suggested that the doctor needed to further probe the patient's scenario, which implies that the doctor was unable to completely derive case-specific diagnoses. In addition, 12% of the doctors proposed additional diagnoses that were observed to be incorrect based upon an examination of the patients' descriptions of their symptoms. Finally, 8% of the videos showed reviewers giving the doctor clear, verbal praise regarding their communicative proficiency. A full breakdown of the commentary findings pertaining to each test group participant is provided in Appendix M.

Table 13

Descriptive results of reviewers' commentary regarding test group participants' ability to communicate effectively with patients and provide accurate diagnoses

Commentary feedback from reviewer	Further general examination or diagnosis required of doctor	Further probing of specific items required of doctor	Doctor is able to correctly state general diagnosis	Additional, incorrect diagnoses given by doctor	Positive acknowledgment of doctor's communication skills (from reviewer)
Average # of occurrences (%)	36	-	100	12	8
<i>Notes.</i> The number of specific items needed to be further probed by the doctor was not always made clear by the reviewer. See Appendix M for further details.					

These findings align with the report provided by the MCC which showed that international medical graduates have a significantly lower pass rate for the MCCQE Part II than their Canadian-trained counterparts. Although this pass rate has improved in recent years, increasing from 59.5% first attempt pass rate in 2013 to 74.4% pass rate in 2017 (Medical Council of Canada, 2017a, p. 34), these rates do not match those of domestically-trained physicians, which increased from 95.4% in 2013 to 97.2% in 2017 (Medical Council of Canada, 2017a, p. 34). The pedagogical implications of these findings will be further explored in the following chapter.

Chapter 5: Discussion

5.1 Introduction

This chapter provides a discussion of the results as they pertain to a wider dialogue of power imbalance and agency within clinical interactions. Thus, the following chapter is organized according to the guiding research questions previously outlined, beginning with an initial overview of the overarching trends found in the previous chapter followed by a critical discussion of the impact of these active and anti-active listening performances in relation to perceptions of doctors' soft skills and ability to collaborate with patients, as well as a final exploration of the pedagogical implications of these findings. Policy recommendations, study limitations and recommendations for future research will be addressed in the final chapter hereafter.

5.2 Responding to research question 1: Addressing the character of active listening on a micro level

5.2.1 Non-verbal modes of active and anti-active listening

The initial research question posed in this study inquired about the impact of individual verbal and non-verbal traits on doctors' overall performance of active or anti-active listening. Thus, as was discussed in the previous chapter, it was critical to compare each verbal and non-verbal, active and anti-active item in the analytical framework performed by the NNS test group participants to the manifestation of all corresponding items by the illustrative and non-illustrative control group population. Although these comparative sets did not warrant statistical differences using Kruskal-Wallis tests, largely as a result of the small sample size, the overall trends in the data provided valuable feedback in regard to which items contribute most significantly to the realization of active or anti-active listening.

Overall, as was observed in the previous chapter, few non-verbal active listening traits were underperformed by the NNS test group participants. Specifically, this group made eye contact for an average duration of 3.38 seconds while eye contact was maintained for nearly twice as long within the illustrative control group. The non-illustrative control group also maintained eye contact for over one minute longer than the test group. Similarly, both the illustrative and non-illustrative group made eye contact approximately two times more frequently than the test group. In regard to deictic, metaphoric and interactive hand gestures, all three groups performed approximately the same ratio of gestures per second. However, the NNS test group overperformed iconic hand gestures in comparison to their NS control group counterparts. The test group also appeared to overperform head nods. However, the greater presence of head nods mirrored the finding that the illustrative control group also performed a large number of head nods, especially in comparison to the non-illustrative group, who performed comparatively fewer nods.

In contrast, the test group consistently underperformed non-verbal anti-active listening traits in comparison to both the illustrative and non-illustrative control groups with some exceptions. One exception to this trend was the action “visible head turns away from the patient”, whereby the test group performed slightly more turns away than the illustrative control group. Additionally, the categories of “lack of eye contact when required” as well as “demonstrates distractions with the hands” were especially salient across all three groups but particularly so in the test group. The saliency of these traits was most commonly a result of doctors taking extensive notes while the patient was speaking. Thus, a broad analysis of the trends in the descriptive statistics of individual non-verbal traits suggested that the NNS test group participants demonstrated a greater reliance on certain anti-active non-verbal traits overall,

varying from the trait performances of the illustrative and non-illustrative NS control groups while at the same time was found wanting in terms of key active non-verbal traits such as eye contact, especially when compared to NS doctors.

As the role of active listening is to build an empathetic relationship between listeners and speakers on the basis of shared knowledge and accessibility in the dialogue, it is critical to explore the shifts in power and interlocutor agency as it is altered across individual components of various communicative modes. Thus, the following section will examine the spectrum of non-verbal listening performances of hand, eye and head movements in relation to the doctor's broader demonstration of power and agency in medical interactions. Firstly, hand gestures offer a number of unexpected findings regarding how doctors and patients negotiate power and build a working relationship. The majority of hand gestures in this study acted as compliments to the verbal content of the conversation. That is, deictic, metaphoric and iconic hand gestures had little to no effect on the expression of power conveyed by the doctors. Rather, they served to visually manifest focal points in the doctors' questions or feedback. Furthermore, it appeared that these hand gestures had little bearing on the overall display of empathy, enabled by active listening, performed by the doctor. For the NNS doctors, gestures were used frequently as a means of overcoming linguistic barriers in the conversation, which resulted particularly in the overproduction of iconic hand gestures, whereas for NS doctors, who commanded a wider vocabulary for portraying difficult subjects, fewer hand gestures of this nature were used.

However, two hand gesture types appeared to greatly impact the positive or negative quality of listening showed by the doctor. The first, interactive hand gestures, acted as a way of tangibly "bridging the gap" in power between the doctor and the patient, thereby positively impacting the quality of listening demonstrated by the doctor. These were primarily apparent in

introductory and closing handshakes whereby conversation-initial and -final demonstrations of respect were conveyed. As a handshake involves bodily contact, it increases “immediacy” between the patient and the doctor, or the sense of intimacy between these interlocutors as a result of shared space (Mehrabian, 1971, p. 7), and thereby facilitates the development of a balanced sense of power in the doctor-patient relationship. Other interactive gestures such as touching the patient’s shoulder or hand were frequently used during conversations involving difficult medical topics. This was a very clear expression of active listening in that the patient was in a highly vulnerable position and the doctor was expressing tangible empathy and care through touch and was often accompanied by other gentle, unobtrusive non-verbal gestures such as sustained eye contact or minimal head nodding.

On the other hand², the gestural category of beats was found to negatively impact the quality of listening portrayed by the doctor. Although the primary use of beats is to mark speaking tempo, in some cases beats were used to extend a speaking turn length in order to block the patient from re-entering the conversation. In this way, beats were used as visible extensions of turns. In other cases, beats could be used alongside verbal output as a means of interrupting the patient, which again reduces the patient’s agency in and contribution to the conversation. This was particularly marked in the NNS doctors’ conversations, potentially as an artifact of the non-verbal features that were characteristic of their first languages. These obtrusive beats were often accompanied by other anti-active listening traits such as when doctors showed distraction, such as playing with their lab coats or hair, or focused on writing notes while the patient was speaking. Although the number of unmarked beats and linguistically semantic gestures (i.e. deictic, iconic and metaphoric gestures) were higher in frequency, marked beats which acted as

² Pun intended.

interruptions or extensions of turns, in addition to other disruptive hand movements, were very noticeable, despite being fewer in number. They also tended to overshadow the positive effects of interactive hand gestures and the overall good intentions of the doctor to express empathy towards the patient.

The appropriate performance of head nods and head positioning are also indicative of a listener's engagement in the conversation and can be used as a means of allowing the speaker to be given power through speaking or as a means of removing agency and subsequently power through visible changes in head movements. Nodding is used to both insert the listener within the speaker's prose and to indicate the listener's viewpoint regarding the storyline without verbally intruding upon the speaker's turn (Stivers, 2008, p. 52). However, they must be performed at particular moments during the storytelling. Thus, nods were expected to occur during the middle of the storytelling in order to demonstrate affiliation to the patient and their story while at the same time remaining unobtrusive. Indeed, if the doctor upgraded from a nod to a vocal affirmation or other verbal recognition in the middle of the patient's turn, the listener was performing a disaffiliating act because, at this point, the patient is searching for some form of confirmation that the doctor was engaged with their story but is not yet prepared to end their speaking turn or receive a more detailed response to their disclosure of information. In this way, once a doctor has established their non-verbal idiosyncrasies, it is difficult to determine what counts as excessive head nodding, so long as it is performed mid-telling. This reflects the trends in the data which showed that, although NNS doctors overperformed head nods in comparison to NS doctors, the number of head nods in a given conversation did not impact the overall listening quality of the doctor. On the other hand, a complete lack of nods mid-speaking turn or even a doctor's obvious turn of the head away from the patient was cause for confusion and/or worry on

the part of the patient. As the production of storytelling is achieved through a cooperative banter of non-verbal cues alongside verbal dialogue, missing a single cue such as nodding prohibited patients from inferring whether the doctor was still following along with the story or if modifications to the story were needed in order to re-establish a middle ground with the doctor.

Finally, eye contact duration and frequency were also explored in terms of their contribution to the relationship between the doctor and the patient and the broader balance of power in the medical interaction. According to Streeck (2009), eye contact “is the primary method by which the interactional axis between speaker and listener is sustained; speakers judge from the interlocutor’s gaze-direction whether their talk is being received by them” (p. 86). However, interlocutors do not maintain consistent eye contact. Instead, eye contact is utilized at various key points in the conversation in order to achieve “mutual gaze” or an agreement between the interlocutors regarding the beginning and end of speaking turns (Streeck, 2009, p. 86). In terms of the present study, this mutual gaze needed to be initialized just prior to the patient’s speaking turn, otherwise the doctor violated the speaker’s gaze expectancy (Streeck, 2009, p. 86). Following this initial gaze, the components of eye contact duration and frequency are able to be more variable provided that the appropriate eye contact sequencing is maintained (Streeck, 2009, p. 87). Eye contact sequencing entails making eye contact at turn beginnings and again at turn endings, while the middle of speaking turns is established by withdrawing direct eye contact in favour of a nearby target (Streeck, 2009, pp. 86-87). Indeed, this was noted in the variance across both the control videos and the test group videos in terms of the frequency and duration of eye contact. For some doctors, eye contact frequency was reduced, but this was a result of maintaining eye contact for longer periods of time. For others, the duration of their eye contact was shorter, but eye contact was sought out more frequently. Thus, as was found in the

nodding discussion, there did not appear to be a preference from the patients in regard to which approach was “better” or more specifically which approach was more suggestive of active listening. However, it was extremely noticeable when a doctor did not return a patient’s gaze at key moments in the patient’s storytelling, such as at turn beginnings or closures. In these instances, the patient would either make “repair-initiators” (Streeck, 2009, p. 86), such as changing their prosody or making syntactic breaks, in order to force the listener’s gaze back to them when the doctor was already beginning to return their gaze to the patient, or stop speaking completely in attempt to receive the doctor’s gaze when it had completely strayed elsewhere. In this study, these absences of eye contact were particularly notable when the doctor was making extensive notes while the speaker was talking, without pausing to look up at the patient in order to make them aware that they were still being listened to. Thus, the doctor’s return of the patient’s gaze allowed both to create a sense of shared understanding and connection to each other and the story while a lack of eye contact physically denied this connection, which further hindered relationships between the patients and doctors.

Overall, although further study with more robust sample sizes is needed, the trends found in terms of the non-verbal modes of active and anti-active listening suggest that individual traits alone can contribute at least minimally to the greater perception of the quality of listening being performed. However, it is the combination of each item, alongside careful timing of each item’s performance within an interaction, that may in fact have the greatest impact on the speaker’s observation of the listener’s ability to engage actively or not.

5.2.2 Verbal modes of active and anti-active listening

Similar to the non-verbal trait analysis described above, the verbal active and anti-active modes showed no significant differences statistically but still featured notable trends in the

descriptive statistical analysis that contribute to the overall quality of listening performed by the participants in this study. Thus, an awareness of the patient's agency in the conversation as well as a recognition of the patient's emotional relationship to their illness experience, or lack thereof, was also demonstrated within the verbal mode. As was found in the previous chapter, in terms of the verbal active listening traits, the NNS test group varied in their performance of individual traits when compared to the control groups. In regard to reflective questions, reflective comments and accessible, general medical commentary, the test group participants were found to have underperformed, but were still more similar in these performances to the illustrative control group than to the non-illustrative control group. However, the test group did underperform compared to both control groups in terms of their ability to accurately summarize patients' commentaries and in their repetition of key facts of the patients' stories. Additionally, the test group actually overperformed in terms of giving affirmations when compared to the control groups.

In terms of the verbal anti-active listening traits performed, there was greater variance amongst the three groups. For inaccurate summaries and interruptions, the NNS test group results mirrored those of the illustrative group more than the non-illustrative group, which showed the most occurrences in these categories. However, the illustrative control group had zero occurrences of the trait "unrelated and/or unempathetic comments"; as a result, although the NNS test group performed this verbal item relatively few times, their performance of this anti-active verbal category was more aligned with the non-illustrative control group's performances of this trait. Most notable however was that the NNS test group greatly overperformed the verbal anti-active listening trait of asking rapid fire and/or unrelated questions, even compared to the non-illustrative control group. Furthermore, the NNS test group's use of inaccessible general

medical commentary as well as inaccessible language greatly exceeded the corresponding categories in both control groups, as neither control group performed these verbal categories whatsoever.

As a conversation that demonstrates positive active listening attributes results in feelings of shared empathy and the development of stronger social bonds between interlocutors, a listener's ability to demonstrate keen interest in the speaking subject is the first component of being a positive active listener from the verbal standpoint. In the medical context, this is demonstrated through the use of thoughtful questioning and commentary as well as ongoing, non-linguistic affirmations or vocalized feedback cues. These verbal actions show that the listener was mindfully engaged with the speaker's prose as they were able to receive and memorize, at least for the duration of the conversation, key facts. Even if the listener was not fully interested in the subject internally, their external, verbal behaviour maintained an attitude of diligent tracking of the conversation.

In contrast, those listeners who, even if they were interested in the subject, were unable to make mental notes of the speaker's prose and engage with the conversation enough to generate reflective questions and comments demonstrated negative active listening qualities, or perhaps more specifically anti-active listening qualities, which could be further compounded by redundancies in their verbal feedback such as asking a question that the speaker had already addressed without a lead-in statement to offset the listener's apparent lack of engagement. According to active listening dogma, these features of dialogue would be viewed as anti-active listening behaviours (i.e. those behaviours which demonstrate that the physician is in fact engaging in the conversation but is not actually demonstrating mindfulness or care towards the speaker's point of view). Each of these acts variously reduce the time given to the patient to

express their needs and views. They also remove the power and authority that patients hold over their health. Thus, when NNS doctors overused a rapid-fire questioning style which was notably absent of reflective interrogatives, in addition to other, negative verbal cues, the patient appeared to be unwilling to share key details about their illness experience. As a result, while these conversation strategies may allow the physician to keep visits short and “to the point”, they detract from the doctor-patient relationship. This is reflected in the literature, in which case upon case of patient feedback stresses that when a physician appears to be in a hurry or otherwise not mentally engaged in or present for the visit, the patient feels less willing to share the facts necessary for producing effective treatments and are more inclined to believe that the physician is cold and uncaring (cf. Jaspars et al., 1983).

A second critical component for determining the quality of active listening verbally is the degree of accessibility granted by the doctor to allow the patient the space to contribute freely to the conversation. This gives the patient agency and allows them to hold power in the conversation in the sense that they can access the speaking floor when desired and when socially appropriate. Affirmations, for example, can act as modes of accessing the conversation as well as denying access to the conversation depending on their timing and frequency in the speaking turn. According to Stivers (2008), the shift up from nodding to vocal affirmations in the middle of a story disturb the flow of prose and may cause the speaker to end their speaking turn altogether (p.53) and denies them access to the conversation. Therefore, this form of affirmation is classified as a disaffiliative or anti-active listening act (Stiver, 2008, p. 53). In this study, NNS doctors tended to overuse these mid-turn affirmations as a means of demonstrating their support of the patient’s case; however, as a vocalized act on the part of the listener mid-speaking turn indicates the abrupt end of a turn (i.e. an interruption) in normative Canadian dialogue, these

NNS doctors were misunderstood by the patients as being anti-active listeners, whereas NS doctors, who used affirmations as a vocal agreement with the patient at the end of the patient's speaking turn, could also produce large amounts of affirmations but not be viewed as anti-active listeners because of their socially appropriate choice of affirmative timing.

Another more explicit way for the listener to deny access verbally was to extend their question or commentary such that the speaker was unable to retrieve the speaking floor without violating the social norms of the conversation. Similar to mistimed affirmations, these extensive turn-holdings impeded upon the patient's ability to participate in the conversation and relinquishes their sense of power in the conversation, which in turn leads them to believe that the doctor does not care to hear their opinion and/or is rushed for time. Similarly, language choice was a critical barrier to accessibility in these clinical interactions. Several of the NNS doctors chose to use medicalese (medical jargon) that was not familiar to the patient and were therefore at risk of preventing patients from re-entering the dialogue as the primary speaker (Norman et al., 2003). This is a result of the fact that doctors are already perceived as "experts" on the patient's condition and infringing upon this knowledge by asking questions could further decrease the patient's power in the conversation. However, if the NNS doctors clarified the medicalese, they were able to invite the patient to share in this expert knowledge and thus were able to restore the power imbalance and the patient's willingness to share their story.

In summary, determining the active listening quality of the doctor from a verbal standpoint depended upon not only their ability to mindfully engage with the speaker's interests but also their ability to facilitate accessibility or participation within the conversation in order to neutralize the power imbalance caused by the perception that the doctor holds higher status than the patient. On a broader level, these components maintain a focus on the patient's experience of

health and illness, which is the goal of the remodelled healthcare approach in Canada in both testing and practice settings.

5.3 Responding to research question 2: Addressing the character of active listening on a macro level

As the micro level analysis of the verbal and non-verbal listening modes illustrate, some items within each mode appear to contribute somewhat more significantly than others within clinical interactions depending upon the timing of the item, its placement in relationship to other listening items and even the existing power balance at play between the doctor and the patient. Thus, the gestalt of items within the verbal mode as well as those of the non-verbal mode are critical to the overall interpretation of the doctor's listening ability. Accordingly, this research shifts not only the way that active listening should be defined, specifically in relation to the expression of empathy and the ability to create meaningful relationships, but also the way that culturally bound traits pertaining to active listening need to be taught to NNS medical candidates, which will be addressed in the relation to the third research question guiding this study. Despite their differences non-verbally, non-native speakers of English in this study were still able to convey their genuine interest in their patient's experiences, specifically through their use of head nods and iconic hand gestures. This was also found when the NNS doctors, despite some linguistic challenges, were able to express interest in their patients' stories through the use of affirmations, albeit mistimed in many cases, and permit a degree of accessibility to the verbal component of their interviews. This seems to suggest that there may not be a minimum or baseline frequency that these verbal and non-verbal behaviours need in order to demonstrate active listening. This was notable in the statistical analysis pertaining to this second research question. Positive indicators of active listening, as a combined total of non-verbal and verbal

behavioural scores, found in the illustrative and non-illustrative control groups negatively correlated with those of the test group as a result of the overproduction of positive traits by the NNS doctors in comparison to the NS doctors in the control groups. However, occurrences of anti-active listening, or the set of negative verbal and non-verbal listening cues were much more salient, both for NNS doctors in the test group and for NS doctors in the non-illustrative control group. The illustrative control group was found to negatively correlate with the non-illustrative control group in terms of negative listening cues, as expected. However, in addition, the test group positively correlated with the non-illustrative control group, which suggests that in fact their production of negative listening cues is more aligned with NS doctors who demonstrated an overall lack of soft skills and an inability to build a collaborative relationship with their patients. Thus, these findings suggest that the threshold for non-engagement, as indicated by the presence of negative listening cues, in medical interactions is lower and much easier to achieve, which is to the detriment of the doctor's perceived personal qualities as rated by patients. This is also reflected in literature pertaining to first impressions and the halo effect. In these studies, doctors who began their interviews poorly had a much more difficult time "making up" for their poor start than those who started their interview well but then made equally problematic errors later on in the interview. For example, Wood et al. (2017) created video recordings of 6 variations of a single history-taking OSCE station and had 23 reviewers score the doctors' clinical abilities and their confidence in the accuracy of the doctors' assessments after the first 60 seconds of the video and then again at the end of the video. The findings suggested that first impressions may impact judgement variability to the detriment of those doctors who were unable to positively establish their clinical abilities from the beginning of the medical interaction.

From these observations then, two potential characterizations of the conceptualization of active listening arise. In the first, active listening and anti-active listening exist on a multi-dimensional continuum in which the positive or negative expression of individual traits move a metaphorical sliding scale towards one type of listening or the other. It is evident that a number of factors relating to the non-verbal aspect of listening must be considered when determining the degree of “active” or “anti-active” listening being performed. At first glance, this appears to be a contradiction. However, much of the evidence of the listener’s engagement in the speaker’s prose is illuminated through their response to the speaker. In this regard, verbal listening skills, like their non-verbal counterparts, must be considered on a continuum of holistic, multimodal listening rather than as sole determinants of one’s capability to actively “listen”. No factor can be used alone in the performance of proficient active listening. Overall, it may be the case that the term “active listening” requires further review, for this fact that evidence of listening is not just reflected in the non-verbal components of an individual but also in the artifacts of the speaker’s commentary in the listener’s response. However, for the purposes of this definition, it is enough to state that it seems that both the verbal and non-verbal factors must be considered as checkpoints on a continuum of listening.

But how do the verbal and non-verbal components work together to create an image of active listening according to this model? Logically, there would be three ways of deciding whether someone is listening or not. In the first scenario, the majority of cues observed in both the verbal and non-verbal categories are positive indicators of active listening, enabling the easy judgement that the listener is in fact listening actively. In the second scenario, both the verbal and non-verbal categories consist of a majority of anti-active listening traits, which again provide a clear judgement that the person is not actively listening. In the third judgement situation

however, the rating of an individual's active or anti-active listening performance is less clear. In this event, the verbal category may indicate active listening behaviours while the non-verbal category implies anti-active listening skills, or vice versa. On the other hand, the performances in both categories may be mediocre such that they are not strongly indicative of active listening nor are they indicative of anti-active listening. In these instances, it may be plausible to discuss the greater weight of non-verbal behaviours in comparison to verbal behaviours in a given conversation, which may further "tip the scale" to favour active listening over anti-active listening, or vice versa. Individual speakers' experiences in past conversations with the same listener or listeners of similar backgrounds (e.g. other doctors in previous clinical interviews) may further guide the speaker to perceive the listener as listening actively or anti-actively.

However, if the speaker does not leave the conversation feeling that the listener was particularly absorbed in the conversation or portrayed some degree of empathy towards the speaker, but neither did they feel that they were exhibiting clearly demonstrable signs of wanting to leave the conversation or that they wholeheartedly did not empathise with or were interested in the speaker's commentary, it may be the case that there is need for a middle ground on the continuum between anti-active and active listening. Passive listening, which raises thoughts of noncommittal presence in a conversation, provides a suitable label for this middle ground between active and anti-active listening without committing the listener to one extreme of the scale or the other. It allows for the "greyness" of personal perceptions and idiosyncrasies. Overall, this model of active listening accounts for the multiple variables that comprise the notion of "listening" or perhaps what should be deemed "communicative engagement", as not all expressions of listening require that the "listener" remain verbally silent (as discussed above). However, it still remains unclear where the balance is exactly shifted. That is, at what point does

a speaker believe that a listener is engaging actively with their conversation rather than passively? Furthermore, at what point does a speaker begin to feel that the listener is not actively engaged according to this metaphorical scale? It could be assumed that this continuum and the placement of the middle point is based not only on the cultural context of the conversation but also on the speaker's individual communicative preferences, which are a culmination of past communicative experiences linked to emotional experiences. Although the characterization reflects the fluidity of language and of communicative events, this model seems highly arbitrary and does not accurately capture the evidence for why third-party reviewers can evaluate the same video and determine whether a listener displays active listening or not.

Therefore, a second model of active listening must be proposed. Instead of characterizing "good" or "poor" listening skills according to a continuum consisting of a number of verbal and non-verbal features, perhaps a more streamlined definition of active listening is required. If a speaker subscribes to a particular sociolinguistic group, then they subscribe to what is acceptable and what is unacceptable for that particular linguistic identity. As was noted earlier in the discussion, the repetition of what is acceptable is relatively limitless; and yet what is unacceptable is immediately salient to the sociolinguistic group because it violates some sociolinguistic norm, however minute. By that same token then, it can be suggested that active listening is not a standalone concept composed of an arbitrary number of changing attributes; rather, active listening is the absence of those features that the sociolinguistic group deems to be acts of "anti-active listening", of which there is a far more finite number of presentations that may occur before the individual notices the negative presence and attributes it to additional negative, personal qualities of the listener. Although this model is more static, it also agrees with literature that claims that first impressions, halo effects, and other such biases towards an

individual are extremely hard to overcome on the part of the speaker if they are not mindful of such biases (cf. Wood, 2014; Wood, Chan, Humphrey-Murto, Pugh, & Touchie, 2017; 2018).

From a psycholinguistic perspective, this approach is also fairly intuitive. For children learning their first languages, correction is made in the form of what not to say or do, not in the presentation of a multitude of directions regarding acceptable ways to say or perform a specific linguistic act. When turning to theories of underlying representation, which portray how linguistic features are stored in the brain, this model is also supported. Speakers of any language will, subconsciously, always adopt the most basic, streamlined version of a linguistic feature rather than attempt to memorize or learn a number of different rules in order to arrive at a particular linguistic output (O'Grady & Archibald, 2005) For example, if a language uses reduplication as a means of intensifying an adjective, the speaker will not memorize every single version of the intensified adjectives; instead, they will simply memorize a rule such as “reduplicate the root to arrive at the intensified adjective”. Similarly, a speaker may in fact store the verbal and non-verbal variants of anti-active listening as the “root” form of listening and then assume that other presentations of that variant, no matter how many forms it appears in, must be a relatively positive act of active listening. This concept of an underlying representation of listening, in both the verbal mode and the non-verbal mode, is also supported in the literature. For example, Cassell, McNeill and McCullough (1999) found that listeners store some form of baseline hand gestures in their heads. In this study, the underlying representations of hand gestures was made evident when listeners were able to recognize incoming hand gestures that contradicted their mental constructs of acceptable hand gestures given the linguistic context. Thus, learning negative forms of active listening, or more specifically acquiring a working knowledge of anti-active listening over time, is what allows speakers to determine whether

displays of active listening are being produced because they variegate from their underlying representation of this “listening” construct.

5.4 Responding to research question 3: Active listening in medical pedagogy for NNS, foreign-trained doctors

The third guiding research question examined the application of active listening to OSCE success as well as the potential for active listening in medical pedagogy for NNS, foreign-trained doctors. In order to respond to this question, the commentary following each test group participant’s mock OSCE was reviewed in terms of diagnostic accuracy and overall communicative proficiency and then compared to the review conducted by the Medical Council of Canada regarding the success of Canadian-trained doctors in comparison to foreign-trained doctors on the MCCQE Part II. Overall, it was found that the imbalance of scores does not appear to be solely based upon the doctors’ abilities to demonstrate sufficient medical knowledge, as all doctors will have passed the MCCQE Part I, which focuses strictly on medical knowledge, prior to being allowed to take the MCCQE Part II. Therefore, it may be inferred that the barrier to passing does in fact lie in the doctor’s ability to create a collaborative, empathetic relationship with the patient in order to elicit information critical to formulating diagnoses and treatment plans (unlike in the MCCQE Part I where all necessary information is presented to the medical candidate without the need for elicitation from a patient). This is achieved primarily through the doctor’s proficiency in active listening, applied within the theoretical framework of the patient-centred care model (PCCM). As briefly introduced in the literature review, the current PCCM is composed of four core values. The first value involves exploring patients’ perceptions of and experiences with health, disease and illness. This allows the physician to understand the patient’s perspective and sociocultural background, which in turn enables them to

ask questions that will elicit more detailed histories than those gathered from routine checklists. In addition, it demonstrates respect for the patient's experience and knowledge of their own situation. This exploration is further summarized by the acronym FIFE, which encourages gathering of the patient's feelings, ideas about what is wrong, fears, effect of the illness on their well-being, and expectations of the physician (Stewart, Brown, McWilliam, Freeman & Weston, 2014, p. 39). Patient-centred care pedagogy suggests that this information can be garnered through any number of open-ended questions, which is also one of the key pillars of verbal active listening. Stewart et al. (2014) also note that physicians should observe and respond to the non-verbal cues of their patients such as tearing up or breathlessness in order to more fully understand their patients' lived experiences of disease and illness (p. 53). The second component of the PCCM recognizes achieving an understanding of the whole person, both in terms of their individual and familial history but also external, contextual factors such as financial security or strain, education, community support, etc. The third component examines the importance of finding common ground with patients. This core value of the PCCM particularly notes the need for clinical visits to be interactive dialogues based on a relatively balanced power structure. This not only facilitates stronger patient-doctor relationships particularly in regard to building trust, but also aids in the creation of a treatment that suits both parties. The fourth component of the PCCM underscores the continual enhancement of the patient-doctor relationship through a consistent demonstration of compassion, caring, empathy and trust. These "soft skills" are maintained through ongoing self-awareness activities and cognizance of the pitfalls of transference or the projection of past negative experiences or biases onto another person (Stewart, Brown, & Freeman, 2014, p. 155).

Studies have also shown that the presence of active listening within the PCCM correlates with increased perceptions of empathy and patient satisfaction. For example, in the study described above conducted by Fassaert et al. (2007), it was found that patients who were anxious prior to the medical consultation, particularly those who were visiting the doctor for the first time, elicited higher levels of active listening from the physician, who also used the opportunity to establish a comfortable atmosphere and set a precedent for follow-up appointments. In another study, Improta (2011) observed four videotaped medical interviews between native and non-native English speakers and had both doctors and patients complete interviews regarding their satisfaction with the interaction. It was found that a number of negative listening cues, particularly interruption and conversational redirection, strongly correlated with patients' reports that the doctor lacked empathy.

In addition, a number of studies have recognized the cross-cultural barriers that need to be recognized and taught to foreign-trained doctors in order to perform culturally appropriate active listening. For example, Lamiani et al. (2008) gave two groups of healthcare professionals, one from the United States and one from Italy, a medical scenario and asked each group to write a patient-centred dialogue on that topic. Each group then provided feedback on the other group's dialogue. Although both realized the importance of emphasizing the patient's illness experience and emotions, the American group achieved patient-centred care by maintaining patient autonomy while the Italian group utilised a paternalistic approach to providing care. Further discussion then recognized the importance of understanding cross-cultural perspectives in providing optimal patient care. Thus, it was recommended that patient-centred care pedagogy needs to begin with establishing an understanding of the many realizations of patient-centered care globally before culture-specific patient-centred practices can be taught.

However, there is still a gap between doctors' self-reported listening skills and patients' observations of doctors' listening abilities. For example, Kenny et al. (2010) conducted a cross-sectional study, collecting data from 91 doctors and their 1,749 patients across three Canadian provinces using post-consultation questionnaires regarding the doctor's communication skills. Although doctors consistently rated themselves highly regardless of the patient, patients of the same doctor gave extremely varied responses regarding the doctor's actual listening skills. In another study, Stewart (1995) explored 21 medical consultations that reported positive, negative or inconclusive diagnoses in relation to the quality of communication skills demonstrated by the doctor. It was determined that effective communication strategies strongly correlated with improved patient health outcomes. Furthermore, the study reinforced the need to bolster medical curriculum programs by including strategies for effective communication in the medical interview.

Thus, the implications of active listening pedagogy are extremely positive. Putnam, Stiles, Jacob and James (1988) studied the effects of teaching specific verbal listening cues on medical consultation outcomes. Using 268 interviews, 156 of which were in the trained test group while the remaining 112 made up the control group, it was found that teaching verbal active listening techniques increased both patient exposition and accessible physician explanations. Additionally, several of the studies recognized above either demonstrated the impact of active listening pedagogy in regard to improvements in collaboration and empathy within various interactional settings (cf. Bodie et al., 2015) or stressed the need for the inclusion of interdisciplinary listening training (cf. Stewart, 1995).

The ramifications of the definition of active listening within the second model described are also particularly critical for cross-cultural medical training programs. It is not enough to

teach students the “right” way to say something or, in this case of this study, the culturally acceptable way to actively listen. In fact, the struggle for learners to accurately obtain pragmatically acceptable forms of speech due to a dearth of curricula is highlighted throughout the literature (Bardovi-Harlig, 2017, p. 226) and the pragmatics of listening appear even less so throughout these limited curricula. However, if the students are simply given guidelines of what not to say or do in different scenarios, then they may more easily be able to recognize the potential for error within their own practice prior to making a sociopragmatic or pragmlinguistic mistake. This is facilitated by the fact that the number of “unacceptable” forms will be more finite for a broader range of scenarios than the “acceptable” forms, which often rely on a whole host of extralinguistic factors (Bardovi-Harlig, 2017, p. 228). Thus, similar to the recommendations given for second language acquisition (SLA) pedagogy, although it can be informative to delineate acceptable communication styles in training videos, an emphasis on unacceptable forms is also necessary. Moreover, these forms should be made explicit to medical candidates and feedback should be delivered immediately, if possible, in keeping with the findings of recent SLA feedback scholarship (e.g. Kartchava & Amar, 2014; Nassaji & Kartchava, 2017). This pedagogical recommendation is also aligned with the recent shift in the Medical Council of Canada structure towards providing feedback through video recordings of doctors’ practice OSCEs. More extensive pedagogical and policy recommendations however will be reinforced in the subsequent Concluding Remarks chapter.

Chapter 6: Concluding Remarks

The aim of this study was to examine the micro and macro level characterizations of active listening as demonstrated by NS and NNS doctors in relation to their performance of Objective Structured Clinical Examinations (OSCEs) and to cross-cultural medical training programs. Previous works have highlighted the fact that although there is paradigmatic shift towards a more patient-centred care model (PCCM) of healthcare, the realization of PCCM in practice by those who have not received training regarding the culturally appropriate demonstration of PCCM behaviours is notably absent. In Canada, this gap between theory and practice is further highlighted by the statistics which show that NNS, foreign-trained doctors perform poorly compared to their Canadian counterparts on the second installment of the national medical qualifying exam which specifically tests the candidate's ability to actively listen to and communicate with patients. Overall, the findings of this study suggest that active listening within the PCCM in medical practice is enacted differently cross-culturally and that an understanding of the salient anti-active listening traits within a given society's medical domain may be beneficial in improving doctor-patient relationships and also may allow doctors to create more individualized treatment plans for patients.

Overall, both NS and NNS doctors were able to demonstrate active listening traits within verbal and non-verbal modes and arrive at diagnoses. However, the data revealed that the two groups differed in the frequency that anti-active listening traits were performed. The exemplary NS doctors performed fewer salient anti-active listening traits when conversing with patients whereas the NNS doctors more frequently enacted salient anti-active listening traits across both modes. As a result, when doctors employed these anti-active listening traits, patients were less likely to provide critical details regarding their health circumstances, which resulted in more

generic or, in some cases, inaccurate diagnoses and treatment plans. Conversely, doctors who did not demonstrate these anti-active listening traits were able to successfully establish a collaborative, interactive dialogue with patients and arrive at individualized, accurate diagnoses and treatment plans.

Therefore, results from this study showed considerable variability in the cross-cultural demonstration of active listening and, by consequence, differences in the portrayal of power and authority within clinical interactions. In this respect, the findings call for an increased awareness of approaches to active listening cross-culturally as well as improved training programs for NNS, foreign-trained doctors in Canada that provide an opportunity for these new doctors to recognize and practice a Canadian approach to patient-centred care for the purposes of successfully completing the MCCQE Part II and for their future practices in medical clinics.

As a result of the project, the coordinator of AIMGA and C2I will be provided with feedback pertaining to how each participant enacted specific active or anti-active listening traits in comparison to the baseline findings in order to encourage a greater awareness of the patient-centred care model of practice in Canada. In addition, the findings will be brought forward to the MCC in order to facilitate a greater understanding of the cross-cultural gap that may exist in the MCCQE Part II assessment, specifically within the communication component, and to encourage further development with regard to cross-cultural training programs in this area of assessment.

6.1 Study limitations and areas of future research

Several limitations and corresponding areas of future research need to be acknowledged in regard to this research. First, as the researcher was the only coder for the main study, some bias may have appeared in the results. Although steps were taken to blind the coder's analysis, primarily through randomization of the videos and by revisiting videos to confirm consistency of

coding, future studies may wish to create smaller segments of videos and have additional reviewers code these segments to determine coding consistency. In addition, the small sample size may have led to some inflated family-wise error rates and/or Type II statistical errors. These errors may be compounded by the fact that the control group videos were scripted while the test group videos were unscripted. As such, the results in this study are primarily exploratory in nature, and future studies may require larger sample sizes in order to more accurately test the relationships between the control groups and test group.

Secondly, although some feedback regarding the accuracy of the doctors' diagnoses was provided in the question period following each OCSE, access to the NNS doctors' final OSCE examination scores was not provided. Additionally, the accuracy of the diagnoses given by the NS doctors was necessarily assumed by the researcher on the basis that the videos were used as models of exemplary or unexemplary OSCE performance. Thus, receiving more detailed feedback from external reviewers regarding participant performance would provide more accurate findings regarding the direct relationship between active listening performance and successful OSCE performance. In addition, this feedback would provide an excellent source of qualitative data pertaining to what reviewers actually notice about candidates' communication skills within the scope of the examination assessment, which could be held in comparison to the multitude of microlevel active listening traits being performed.

Furthermore, although the research findings are being provided to the research partner who coordinates OSCE training programs at AIMGA and C2I, the researcher does not have the ability to directly discuss the feedback with the participants nor determine whether the feedback was applied in future attempts of the MCCQE Part II or related, preparatory OSCEs. For example, it was found in this study that having an extensive knowledge of what not to do in

terms of active listening could be more useful than simply performing a number of acceptable active listening behaviours. Thus, it is recommended that future studies compare the OSCE success of groups who are provided with training specifically on anti-active listening as opposed to those provided with training only on positive active listening traits. An extension of this research would include whether participants who had in fact improved in terms of their cross-cultural communication skills, specifically in the area of listening, received the same or different feedback from external reviewers, as some studies have shown that individuals hold subconscious beliefs regarding the intelligibility of NNS individuals regardless of their overall language and cultural proficiency.

A final note regarding the limitations of this study was that it was a largely quantitative study focused on enumerating a wide variety of microlevel active and anti-active listening traits from the perspective of a single, third-party reviewer, which prohibited interrater reliability. Future studies should consider marrying these findings with more qualitative methods, such as those found in other active listening studies, which garner patient, doctor and additional third-party reviewer perspectives of the doctor's active listening skills alongside their personal and intellectual qualities.

References

- Allen, D., & Guy, R. F. (1974). *Conversation analysis*. The Hague: Mouton & Co. N.V., Publishers.
- Amundsen, D. W., & Ferngren, G. B. (1983). Evolution of the patient-physician relationship: Antiquity through the renaissance. In E. E. Shelp (Ed.), *The clinical encounter* (pp. 3-46). Dordrecht, Holland: D. Reidel Publishing Company.
- Argyle, M. (1983). Doctor-patient skills. In D. Pendleton & J. Hasler (Eds.), *Doctor-patient communication* (pp. 57-74). Orlando: Academic Press Inc.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Baile, W. F., Buckman, R., Lenzi, R., Glober, G., Beale, E. A., & Kudelka, A. P. (2000). SPIKES- A six-step protocol for delivering bad news: Application to the patient with cancer. *The Oncologist*, 5, 302-311
- Balint, M., Hunt, J., Joyce, D., Marinker, M., & Woodcock, J. (1970). *Treatment or diagnosis: A study of repeat prescriptions in general practice*. Philadelphia, PA: JB Lippincott.
- Bardovi-Harlig, K. (2017). Chapter 13: Acquisition of L2 pragmatics. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 224-245). New York; Oxon: Routledge.
- Barefoot, P. (2005). Buildings for health: Then and now. In H. King (Ed.), *Health in antiquity*, (pp. 205-215). London; New York: Routledge.
- Bavelas, J. B., Chovil, N., Lawrie, D. A., & Wade, A. (1992). Interactive gesture. *Discourse Processes*, 15(4), 469-489.

- Beltrán-Palanques, V. (2016). The distinctive multimodal nature of pragmatic competence: Bridging the gap between modes. In V. Bonsignori & B. C. Camiciottoli (Eds.), *Multimodality across communicative settings, discourse domains and genres* (pp. 93-115). Newcastle upon Tyne, UK: Cambridge Scholars Publishing.
- Beltrán-Planques, V., & Querol-Julián, M. (2018). English language learners' spoken interaction: What a multimodal perspective reveals about pragmatic competence. *System*, 77, 80-90.
- Bodie, G. D. (2011). The active-empathetic listening scale (AELS): Conceptualization and evidence of validity within the interpersonal domain. *Communication Quarterly*, 59(3), 277-295.
- Bodie, G. D., & Jones, S. M. (2012). The nature of supportive listening II: The role of verbal person centeredness and nonverbal immediacy. *Western Journal of Communication*, 76(3), 250-269.
- Bodie, G. D., Vickery, A. J., Cannava, K., & Jones S. M. (2015). The role of “active listening” in informal helping conversations: Impact on perceptions of listener helpfulness, sensitivity, and supportiveness and discloser emotional improvement. *Western Journal of Communication*, 79(2), 151-173.
- Boyle, D., Dwinnell, B., & Platt, F. (2005). Invite, listen, and summarize: A patient-centered communication technique. *Academic Medicine*, 80(1), 29–31.
- Bublitz, W. (1988). *Supportive fellow-speakers and cooperative conversations: Discourse topics and topical actions, participant roles and ‘recipient action’ in a particular type of everyday conversation*. Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Calloway-Thomas, C. (2010). *Empathy in the global world: An intercultural perspective*. Thousand Oaks, CA: Sage Publications, Inc.

- Calvo, M. G., Gutiérrez-García, A., Fernández-Martín, A., & Nummenmaa, L. (2014). Recognition of facial expressions of emotion is related to their frequency in everyday life. *Journal of Nonverbal Behaviour*, 38(4), 549-567.
- Camiciottoli, B. C., & Bonsignori, V. (2016). Introduction. In V. Bonsignori & B. C. Camiciottoli (Eds.), *Multimodality across communicative settings, discourse domains and genres* (pp. 1-13). Newcastle upon Tyne, UK: Cambridge Scholars Publishing.
- Campoy-Cubillo, M. C. (2016). Multimodal listening skills: Issues in assessment and implementation. In V. Bonsignori & B. C. Camiciottoli (Eds.), *Multimodality across communicative settings, discourse domains and genres* (pp. 14-36). Newcastle upon Tyne, UK: Cambridge Scholars Publishing.
- Caris-Verhallen, W. M. C. M., Kerkstra, A., & Bensing, J. M. (1999). Non-verbal behaviour in nurse-elderly patient communication. *Journal of Advanced Nursing*, 29(4), 808-818.
- Carpenter, R., & Reddi, B. (2012). Chapter 6: Hearing. In *Neurophysiology: A conceptual approach* (5th ed.) (pp. 108-125). London: Hodder Arnold.
- Cassell, J., McNeill, D., & McCullough, K. (1999). Speech-gesture mismatches: Evidence for one underlying representation of linguistic and nonlinguistic information. *Pragmatics & Cognition*, 7(1), 1-34.
- Castro, D. R., Kluger, A. N., & Itzhakov, G. (2015). Does avoidance-attachment style attenuate the benefits of being listened to? *European Journal of Social Psychology*, 46, 762-775.
- Celce-Murcia, M., Dornyei, Z., & Thurrell, S. (1995). Communicative competence: A pedagogically motivated model with content specifications. *Issues in Applied Linguistics*, 6(2), 5-35.

- Chugani, H. (2004). Fine-tuning the baby brain. *Cerebrum*. New York: A Dana Foundation Publication.
- Clementi, P. (2006). *Patient expectations during health care encounters theory: A grounded theory study* (Doctoral dissertation). Retrieved from ProQuest Dissertations Publishing. (AAT 3229775)
- Cohn, J. F., Ambadar, Z., & Ekman, P. (2007). Chapter 13: Observer-based measurement of facial expressions with the Facial Action Coding System. In J. A. Coan & J. J. B. Allen (Eds.), *Handbook of emotion elicitation and assessment* (pp. 203-221). New York: Oxford University Press.
- Condon, W. S., & Ogston, W. D. (1966). Sound film analysis of normal and pathological behavior patterns. *Journal of Nervous and Mental Disease*, 143(4), 338-347.
- Cox, C. (2014). Getting started with ELAN [PDF file]. Retrieved from https://culearn.carleton.ca/moodle/pluginfile.php/2479601/mod_resource/content/2/Cox-2014-Getting_started_with_ELAN.pdf
- Cozolino, L. (2006). The social brain. *Psychotherapy in Australia*, 12(2), 12-17.
- Depraz, N. (2001). The Husserlian theory of intersubjectivity as alterology: Emergent theories and wisdom traditions in the light of genetic phenomenology. *Journal of Consciousness Studies*, 8(5-6), 169-178.
- Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford, UK: Oxford University Press.
- Drollinger, T., Comer, L., & Warrington, P. (2006). Development and validation of the active empathetic listening scale. *Psychology & Marketing*, 23(2), 161–179.
- Efron, D. (1941). *Gesture and environment*. Oxford, England: King's Crown Press.

- Egger, S. (2015). *Interprofessional communication and self-efficacy of non-native English speaking and native English speaking baccalaureate nursing students*. Retrieved from ProQuest Dissertations Publishing. (AAT 3724887)
- Egolf, D. B. (2012). Chapter 4: Human communication. In *Human communication and the brain: Building the foundation for the field of neurocommunication* (pp. 41-60). Plymouth: Lexington Books.
- Ekman, P. (2006). Cross-cultural studies of facial expressions. In P. Ekman (Ed.), *Darwin and facial expressions: A century of research in review* (pp. 169-256). New York: Academic Press.
- Ekman, P., & Friesen, W. V. (1972). Hand movements. *Journal of Communication*, 22, 353–374.
- Epstein, R. M., Franks, P., Fiscella, K., Shields, C. G., Meldrum, S. C., Kravitz, R. L., & Duberstein, P. R. (2005). Measuring patient-centred communication in patient-physician consultations: Theoretical and practical issues. *Social Science & Medicine*, 61(7), 1516-1528.
- Fassaert, T., van Dulmen, S., Schellevis, F., & Bensing, J. (2007). Active listening in medical consultations: Development of the Active Listening Observation Scale (ALOS-global). *Patient Education and Counseling*, 68, 258-264.
- Fellegly, A. M. (1995). Patterns and functions of minimal response. *American Speech*, 70(2), 186-99.
- Fortanet-Gómez, I., & Camiciottoli, B. C. (2015). Introduction. In B. C. Camiciottoli & I. Fortanet-Gómez (Eds.), *Multimodal analysis in academic settings: From research to teaching* (pp. 1-16). New York; London: Routledge.

- Friesen, W. V. (1972). *Cultural differences in facial expressions in a social situation: An experimental test of the concept of display rules* [Unpublished doctoral dissertation]. University of California, San Francisco.
- Fuse, A., Navichkova, Y., & Alloggio, K. (2018). Perception of intelligibility and qualities of non-native accented speakers. *Journal of Communication Disorders, 71*, 37-51.
- Gao, Z., & Naugler, C. (2013). *Clinical skills review: Scenarios based on standardized patients* (3rd ed.). Edmonton: Brush Education Inc.
- Goodwin, C. (1980). Restarts, pauses, and the achievement of a state of mutual gaze at turn-beginning. *Sociological Inquiry, 50*, 272-302.
- Guenther, F. H. (2016). Chapter 5: Auditory feedback control. In *Neural control of speech* (pp. 153-176). Cambridge, MA: The MIT Press.
- Halliday, M. A. K. (1985). *An Introduction to Functional Grammar*. London; Baltimore: Edward Arnold Press.
- Hammer, M. R. (2001). Conflict negotiation under crisis conditions. In W. Eadie & P. Nelson (Eds.), *The language of conflict resolution* (pp. 57–80). Thousand Oaks, CA: Sage.
- Harper, N. L. (1979). *Human Communication Theory: The History of a Paradigm*. Rochelle Park, NJ: Hayden.
- Henzel, V. M. (1989). Linguistic means of social distancing in physician-patient communication. In W. von Raffler-Engel (Ed.), *Doctor-patient interaction* (pp. 77-92). Amsterdam; Philadelphia: John Benjamins Publishing Company.

- Hopkins, P., & Balint Society. (1972). *Patient-centred medicine: Based on the first international conference of the Balint society in Great Britain on "The doctor, his patient, and the illness", held March 23-25 at the Royal College of Physicians, London*. London: Regional Doctor Publications Ltd.
- Hurley, K. (2011). *OSCE and clinical skills handbook* (2nd ed.). Toronto: Elsevier.
- Hymes, D. (1974). *Foundations in Sociolinguistics: An Ethnographic Approach*. Philadelphia: The University of Pennsylvania Press.
- Iacoboni, M. (2008). Mesial frontal cortex and super mirror neurons. *Behavioural and Brain Sciences*, 31(1), 30.
- Iacoboni, M., & Dapretto, M. (2006) The mirror neuron system and the consequences of its dysfunction. *Nature Reviews Neuroscience*, 7(12), 942–951.
- Improta, M. W. (2011). *The role of effective communication during the medical interview between native and non-native English speaking dyads*. Retrieved from ProQuest Dissertations Publishing. (AAT 3454474)
- Jaspars, J., King, J., & Pendleton, D. (1983). The consultation: A social psychological analysis. In D. Pendleton & J. Hasler (Eds.), *Doctor-patient communication* (pp. 139-160). Orlando: Academic Press Inc.
- Jones, S. M., Bodie, G. D., & Hughes, S. D. (2016). The impact of mindfulness on empathy, active listening, and perceived provisions of emotional support. *Communication Research*. <https://doi.org/10.1177/0093650215626983>
- Kagan, P. (2008a). Feeling listened to: A lived experience of human-becoming. *Nursing Science Quarterly*, 21(2), 59-67.

- Kagan, P. (2008b). Listening: Selected perspectives in theory and research. *Nursing Science Quarterly*, 21(2), 105-110.
- Karen, R. (1998). *Becoming attached: First relationships and how they shape our capacity to love*. New York: Oxford University Press.
- Kartchava, E., & Ammar, A. (2014). The noticeability and effectiveness of corrective feedback in relation to target type. *Language Teaching Research*, 18(4), 428-452.
- Kelly, S. D., Barr, D. J., Church, R. B., & Lynch, K. (1999). Offering a hand to pragmatic understanding: The role of speech and gesture in comprehension and memory. *Journal of Memory and Language*, 40, 577-592.
- Kendon, A. (1967). Some functions of gaze direction in social interaction. *Acta Psychologica*, 26(1), 22-63.
- Kendon, A. (1996). An agenda for gesture studies. *Semiotic Review of Books*, 7, 8-12.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Kenny, D. A., Veldhuijzen, W. van der Weijden, T., LeBlanc, A., Lockyer, J., Légaré, F., & Campbell, C. (2010). Interpersonal perception in the context of doctor-patient relationships: A dyadic analysis of doctor-patient communication. *Social Science & Medicine*, 70(5), 763-768.
- King, H. (2005). Women's health and recovery in the Hippocratic corpus. In H. King (Ed.), *Health in antiquity* (pp. 150-161). London; New York: Routledge.
- Kita, S. (2009). Cross-cultural variation of speech-accompanying gesture: A review. *Language and Cognitive Processes*, 24(2), 145-167.

- Klaus, M., & Kennell, J. (1972). Maternal attachment: Importance of the first postpartum days. *New England Journal of Medicine*, 286(9), 460-463.
- Knapp, M. L. (1972). The field of nonverbal communication: An overview. In C. J. Steward & B. Kendall (Eds.), *On Speech Communication: An Anthology of Contemporary Writings and Messages* (pp. 57-71). New York: Holt, Rinehart & Winston.
- Kourmoussi, N., Amanaki, E., Tzavara, C., & Koutras, V. (2017). Active Listening Attitude Scale (ALAS): Reliability and validity in a nationwide sample of Greek educators. *Social Sciences*, 6(1), 28.
- Kress, G., & van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. London: Arnold.
- Lamiani, G. Meyer, E. C., Rider, E. A., Browning, D. M., Vegni, E., Mauri, E., Moja, E. A., & Truog, R. D. (2008). Assumptions and blind spots in patient-centredness: Action research between American and Italian health care professionals. *Medical Education*, 42(7), 712-720.
- Larina, T. (2015). Culture-specific communicative styles as a framework for interpreting linguistic and cultural idiosyncrasies. *International Review of Pragmatics*, 7(2), 195-215.
- Leech, G. N. (1983). *Principles of pragmatics*. London; New York: Longman.
- Lensgraf, A. F. (1989). Ten keys to better doctor-patient relations. In W. von Raffler-Engel (Ed.), *Doctor-patient interaction* (pp. 285-288). Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Levine, A., & Heller, R. (2010). *Attached: The new science of adult attachment and how it can help you find- and keep- love*. New York: J. P. Tarcher of Penguin Group.

- Levitt, D. H. (2001). Active listening and counselor self-efficacy: Emphasis on one microskill in beginning counselor training. *The Clinical Supervisor, 20*(2), 101-115.
- Little, P., Everitt, H., Williamson, I., Warner, G., Moore, M., Gould, C., Ferrier, K., & Payne, S. (2001a). Observational study of effect of patient centredness and positive approach on outcomes of general practice consultations. *British Medical Journal, 323*(7318), 908-911.
- Little, P., Everitt, H., Williamson, I., Warner, G., Moore, M., Gould, C., Ferrier, K., & Payne, S. (2001b). Preferences of patients for patient-centred approach to consultation in primary care: Observational study. *British Medical Journal, 322*(7284), 468-472.
- Loehr, D. P. (2004). *Gesture and Intonation* [Unpublished doctoral dissertation]. Washington, DC: Georgetown University.
- May, S. E. (2007). "Does your throat hurt more in the morning or throughout the day?" "Yes.": *Intercultural medical discourse* [Doctoral dissertation]. Retrieved from ProQuest Dissertations Publishing. (AAT 3288917)
- McCarthy, A., Lee, K., Itakura, S., & Muir, D. W. (2006). Cultural display rules drive eye gaze during thinking. *Journal of Cross-Cultural Psychology, 37*(6), 717-722.
- McNeill, D. (2005). *Gesture and Thought*. Chicago: University of Chicago Press.
- McWhinney, I. R. (2014). The evolution of clinical method. In M. Stewart, J. B. Brown, W. W. Weston, I. R. McWhinney, C. L. McWilliam & T. R. Freeman, *Patient-centered medicine transforming the clinical method* (pp. 18-32). London; New York: Radcliffe Publishing.
- Medical Council of Canada (2017a). 2017 MCCQE Part II annual technical report [PDF]. Retrieved from <https://mcc.ca/media/MCCQE-Part-II-Annual-Technical-Report-2017.pdf>

- Medical Council of Canada (2017b). Assessment Evolution for international students and graduates. Retrieved from <https://www.youtube.com/watch?v=Ib8EQG5wQDY>
- Medical Council of Canada (2019a). Our story. Retrieved from <https://mcc.ca/about/story/>
- Medical Council of Canada (2019b). Governance. Retrieved from <https://mcc.ca/about/governance/>
- Medical Council of Canada (2019c). Vision, mission and strategic goals. Retrieved from <https://mcc.ca/about/vision-mission-goals/>
- Medical Council of Canada (2019d). Medical Council of Canada Qualifying Examination Part I. Retrieved from <https://mcc.ca/examinations/mccqe-part-i/>
- Medical Council of Canada (2019e). Medical Council of Canada Qualifying Examination Part II. Retrieved from <https://mcc.ca/examinations/mccqe-part-ii/>
- Medical Council of Canada (2019f). History of the OSCE. Retrieved from <https://mcc.ca/examinations/osce-orientation/history/>
- Medical Council of Canada (2019g). Common mistakes made on Objective Structured Clinical Examination stations. Retrieved from <https://mcc.ca/examinations/common-mistakes-osce/>
- Mehrabian, A. (1971). *Silent Messages*. Belmont, CA: Wadsworth Publishing Company, Inc.
- Mineyama, S., Tsutsumi, A., Takao, S., Nishiuchi, K., & Kawakami, N. (2007). Supervisors' attitudes and skills for active listening with regard to working conditions and psychological stress reactions among subordinate workers. *Journal of Occupational Health, 49*, 81–87.
- Mishima, N., Kubota, S., & Nagata, S. (2000). Development of a questionnaire to assess the attitude of active listening. *Journal of Occupational Health, 42*(3), 111-118.

- Montserrat, D. (2005). 'Carrying on the work of the earlier firm': Doctors, medicine and Christianity in the *Thaumata* of Sophronius of Jerusalem. In H. King (Ed.), *Health in antiquity* (pp. 230-242). London; New York: Routledge.
- Nassaji, H., & Kartchava, E. (Eds.) (2017). *Corrective feedback in second language teaching and learning: Research, theory, applications, implications*. New York: Routledge.
- Nation, I. S. P., & Newton, J. (2009). *Teaching ESL/EFL Listening and Speaking*. New York: Routledge.
- Neustein, A. (1989). Medical history-taking as an interactive event. In W. von Raffler-Engel (Ed.), *Doctor-patient interaction* (pp. 61-76). Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Nobe, S. (2000). Where do *most* spontaneous representational gestures actually occur with respect to speech? In D. McNeill (Ed.), *Language and gesture* (pp. 186-198). Cambridge: Cambridge University Press.
- Norman, G. R., Arfai, B., Gupta, A., Brooks, L. R., & Eva, K. W. (2003). The privileged status of prestigious terminology: Impact of "medicalese" on clinical judgements. *Academics Medicine*, 78(10), 82-84.
- Norris, S. (2004). *Analyzing multimodal interaction: A methodological framework*. London: Routledge.
- O'Grady, W., & Archibald, J. (2015). *Contemporary linguistic analysis: An introduction* (8th edition). Toronto: Pearson Education Canada.
- O'Halloran, K. L. (2008). Systemic functional-multimodal discourse analysis (SF-MDA): Constructing ideational meaning using language and visual imagery. *Visual Communication*, 7(4), 443-475.

- O'Halloran, K. L. (2011). Multimodal discourse analysis. In K. Hyland & B. Paltridge (Eds.), *The continuum companion to discourse analysis* (pp. 120–137). London; New York: Continuum.
- Patterson, M. L. (1983). *Nonverbal behaviour: A functional perspective*. New York: Springer Verlag.
- Patterson, M. L. (1988). Functions of nonverbal behaviour in close relationships. In S. Duck, D. F. Hay, S. E. Hobfoll, W. Ickes & B. M. Montgomery (Eds.), *Handbook of personal relationships: Theory, research and interventions* (pp. 41-56). Oxford: John Wiley & Sons.
- Philippot, P., Feldman, R. S., & McGee, G. (1992). Nonverbal behavioural skills in an educational context: Typical and atypical populations. In R. S. Feldman (Ed.), *Applications of nonverbal behavioural theories and research* (pp. 191-213). New York: Psychology Press.
- Putnam, S. M., Stiles, W. B., Jacob, M. C., & James, S. A. (1988). Teaching the medical interview. *Journal of General Internal Medicine*, 3(1), 38-47.
- Raider, E., Coleman, S., & Gerson, J. (2006). Teaching conflict resolution skills in a workshop. In M. Deutsch, P. Coleman & E. Marcus (Eds.), *The handbook of conflict resolution: Theory and practice* (2nd ed.) (pp. 695–725). San Francisco: Jossey-Bass/Wiley.
- Robertson, K. (2005). Active listening: More than just paying attention. *Australian Family Physician*, 34(12), 1053–1055.
- Roever, C. (2011). Testing of second language pragmatics: Past and future. *Language Testing*, 28(4), 463–481.

- Rogan, R. G., & Hammer, M. R. (2002). Crisis/hostage negotiations: A communication-based approach. In H. Giles (Ed.), *Law enforcement, communication, and community* (pp. 229–254). Amsterdam: John Benjamins Publishing.
- Rogers, C. R., & Farson, R. E. (1988). Active listening. In P. B. Dubose (Ed.), *Readings in management* (pp. 184-195). Englewood Cliffs, NJ: Prentice-Hall.
- Rost, M. (2011). *Teaching and researching listening* (2nd ed.). Abingdon, Oxon: Routledge.
- Royce, T. (2007). Intersemiotic complementarity: A framework for multimodal discourse analysis. In T. D. Royce & W. L. Bowcher, *New directions in the analysis of multimodal discourse* (pp. 63-109). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Schegloff, E. (1982). Discourse as an interactional achievement: Some uses of “uh huh” and other things that come between sentences. In D. Tannen (Ed.), *Analyzing discourse text and talk* (pp. 71-93). Washington, DC: Georgetown University Press.
- Schegloff, E. A. (1984). On some gestures' relation to talk. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 266-295). Cambridge: Cambridge University Press.
- Scollon, R., & Scollon, S. W. (2003). *Discourses in place: Language in the material world*. London: Routledge.
- Scollon, C. N., Diener, E., Oishi, S., & Biswas-Diener, R. (2004). Emotions across cultures and methods. *Journal of Cross-Cultural Psychology*, 35, 304-326.
- Scollon, C. N., Diener, E., Oishi, S., & Biswas-Diener, R. (2005). An experience sampling and cross-cultural investigation of the relation between pleasant and unpleasant affect. *Cognition and Emotion*, 19, 27-52.

- Scott, J. G., Cohen, D., DiCicco-Bloom, B., Orzano, A. J., Jaen, C. R., & Crabtree, B. F. (2001). Antibiotic use in acute respiratory infections and the ways patient pressure physicians for a prescription. *Journal of Family Practice*, 50(12), 853-858.
- Shipley, S. D. (2010). Listening: A concept analysis. *Nursing Forum*, 45(2). 125-34.
- Silverman, J., Kurtz, S., & Draper, J. (2005). *Skills for communicating with patients* (2nd ed.). Abingdon, UK: Radcliffe Medical Press.
- Stafford, E. (2005). 'Without you no one is happy': The cult of health in ancient Greece. In H. King (Ed.), *Health in antiquity* (pp. 120-135). London; New York: Routledge.
- Stewart, M. A. (1995). Effective physician-patient communication and health outcomes: A review. *Canadian Medical Association Journal*, 152(9), 1423-1433.
- Stewart, M., Brown, J. B., & Freeman, T. R. (2014). The fourth component: Enhancing the patient-doctor relationship. In M. Stewart, J. B. Brown, W. W. Weston, I. R. McWhinney, C. L. McWilliam & T. R. Freeman (Eds.), *Patient-centred medicine transforming the clinical method* (pp. 142-159). London; New York, NY: Radcliffe Publishing.
- Stewart, M., Brown, J. B., McWilliam, C. L., Freeman, T. R., & Weston, W. W. (2014). The first component: Exploring health, disease, and the illness experience. In M. Stewart, J. B. Brown, W. W. Weston, I. R. McWhinney, C. L. McWilliam & T. R. Freeman (Eds.), *Patient-centred medicine transforming the clinical method* (pp. 35-60). London; New York, NY: Radcliffe Publishing.

- Stewart, M., Brown, J. B., Weston, W. W., Freeman, T. R., & McWilliam, C. L. (2014). Introduction. In M. Stewart, J. B. Brown, W. W. Weston, I. R. McWhinney, C. L. McWilliam & T. R. Freeman (Eds.), *Patient-centred medicine transforming the clinical method* (pp. 3-17). London; New York, NY: Radcliffe Publishing.
- Stickley, T. & Freshwater, D. (2006). The art of listening in the therapeutic, relationship: the role of the modern mental health nurse is becoming more technical when, argue Theodore Stickley and Dawn Freshwater, what would really benefit patient the often misunderstanding art of listening. *Mental Health Practice*, 9(5), 12-18.
- Stivers, T. (2008). Stance, alignment, and affiliation during storytelling: When nodding is a token of affiliation. *Research on Language & Social Interaction*, 41(1), 31–57.
- Streeck, J. (2009). *Gesturecraft: The manu-facture of meaning*. Amsterdam; Philadelphia: John Benjamins
- Sue, D. W., & Sue, D. (1977). Barriers to effective cross-cultural counseling. *Journal of Counseling Psychology*, 24(5), 420-429.
- Tang, K. S. (2013). Instantiation of multimodal semiotic systems in science classroom discourse. *Language Sciences*, 37, 22-35.
- Tate, P. (1983). Doctor's style. In Pendleton, D. & Hasler, J. (Eds.), *Doctor-patient communication* (pp. 75-88). Orlando: Academic Press Inc.
- Thomas, J. (1983). Cross-cultural pragmatic failure. *Applied Linguistics*, 4(2), 91-112.
- Ting-Toomey, S. (1988). Intercultural conflicts: A face-negotiation theory. In Y. Y. Kim & W. Gudykunst (Eds.), *Theories in intercultural communication* (pp. 213–235). Newbury Park, CA: Sage Publications.

- Ting-Toomey, S. (2007). Intercultural conflict training: Theory-practice approaches and research challenges. *Journal of Intercultural Communication Research*, 36(3), 255-271.
- Ting-Toomey, S., & Kurogi, A. (1998). Facework competence in intercultural conflict: An updated face-negotiation theory. *International Journal of Intercultural Relations*, 22, 187–225.
- Tsuchiya, K. (2013). *Pragmatics & beyond new series: Listenership behaviours in intercultural encounters: A time-aligned multimodal corpus analysis*. Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Valbonesi, L., Ansari, R., McNeill, D., Quek, F., Duncan, S., McCullough, K. E., & Bryll, R. (2002). *Multimodal signal analysis of prosody and hand motion: Temporal correlation of speech and gestures*. Toulouse: 11th European Signal Processing Conference. Retrieved from <https://ieeexplore.ieee.org/abstract/document/7072137/authors#authors>
- Vandergrift, L. (1997). The Cinderella of communication strategies: Reception strategies in interactive listening. *The Modern Languages Journal*, 81(4), 494-505.
- Vogel, D., Meyer, M., & Harendza, S. (2018). Verbal and non-verbal communication skills including empathy during history taking of undergraduate medical students. *BMC Medical Education*, 18(157). <https://doi.org/10.1186/s12909-018-1260-9>
- Weger, H. J., Castle, G. R., & Emmett, M. C. (2010). Active listening in peer interviews: The influence of message paraphrasing on perceptions of listening skill. *The International Journal of Listening*, 24(1), 34-49.

- Weir, C., & Vidaković, I. (2013). The measurement of listening ability 1913-2012. In C. J. Weir, I. Vidaković & E. D. Galaczi (Eds.), *Measured constructs: A history of Cambridge English language examinations 1913-2012* (Vol. 37) (pp. 347-444). Cambridge: Cambridge University Press.
- White, S. (1989). Backchannels across cultures: A study of Americans and Japanese. *Language in Society*, 18(1), 59-76.
- White, G. M. & Robillard, A. B. (1989). Doctor talk and Hawaiian 'talk story': The conversational organisation of a clinical encounter. In W. von Raffler-Engel (Ed.), *Doctor-patient interaction* (pp. 197-212). Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Wood, T. J. (2014). Exploring the role of first impressions in rater-based assessments. *Advances in Health Sciences Education: Theory and Practice*, 19(3), 409-427.
- Wood, T. J., Chan, J., Humphrey-Murto, S., Pugh, D. & Touchie, C. (2017). The influence of first impressions on subsequent ratings within an OSCE station. *Advances in Health Sciences Education: Theory and Practice*, 22(4), 969-983.
- Wood, T. J., Pugh, D., Touchie, C., Chan, J. & Humphrey-Murto, S. (2018). Can physician examiners overcome their first impression when examinee performance changes? *Advanced Health Science Education: Theory and Practice*, 23(4), 721-732.
- Wright, D. & Mullally, S. (2016). 'Not everyone can be a Gandhi': South Asian-trained doctors immigrating to Canada, c. 1961-1971. *Ethnicity & Health*, 21(4), 340-354.
- Yedidia, M. J. (2007). Transforming doctor-patient relationships to promote patient-centred care: Lessons from palliative care. *Journal of Pain and Symptom Management*, 33(1), 40-57.

Yngve, V., (1970). On getting a word in edgewise. In *Papers from the Sixth Regional Meeting of the Chicago Linguistic Society* (pp. 567–578). Chicago, IL: Department of Linguistics, University of Chicago.

Appendices

Appendix A

Illustration of common MCCQE Part II errors

Error	Description of error
Error 1. Asking too many questions.	Asking as many questions as you can, especially if done in a rapid-fire, disorganized fashion, in hopes of asking enough of the right questions, is not a good strategy. The relatively short amount of time allowed for each task requires that you organize your approach, preferably around a differential diagnosis or around generating a differential diagnosis.
Error 2. Using too many directed questions.	Using open questions is helpful in this exam, just as it is in clinical practice. While the broad question “Tell me why you are here” may not lead patients to share all the information you need, it will certainly provide a starting point. Questions such as “Can you describe the pain for me?” or “Have you noticed any other changes?” are helpful. You will likely have to follow up with more directed questions to get the full picture from the patient but it is easier to start with their initial comments than to run through endless “yes/no” questions that limit the patients to saying yes or no. Remember that, as in real life, patients cannot report what they do not have and may not report changes that do not relate to the presenting problem. However, your score is based on eliciting critical information, not only on eliciting specific pertinent positives. If a patient says he or she has no other symptoms, but you need to know whether there has been any weight loss or specific neurological changes to determine the diagnosis, then you should ask specifically about these points.
Error 3. Not listening to patients.	Patients often report that candidates do not listen to them and therefore miss crucial information. If a patient’s answer is not clear AND the issue is important, then explore further. For instance, you ask the

	<p>patient if he or she has ever had this problem before, and the patient responds with “not like this”. You may want to know what was meant by this statement. If you appear indifferent to their concerns, the patients are expected to react accordingly – you will get less information from them. They are not required to give everyone their whole story. You are expected to elicit their story. And as in clinical practice, patients will be reluctant to confide their underlying fears or embarrassing information if they feel you are indifferent, judgmental or hostile to them.</p>
<p>Error 4. Talking too much.</p>	<p>Sometimes candidates talk so much while with the patient that they lose valuable time and miss information. Efficiency does not exclude the professional courtesies of introducing yourself or explaining briefly why you need to ask a difficult question or conduct a painful maneuver; however, it is not necessary to talk incessantly. You are expected to speak to the patient in language that he or she will understand, not to quote textbooks to the examiner. Your manner tells both the patient and the examiner a lot about your attitude to the patient; you do not need to do it all with words. Avoid lecturing and showing off.</p>

Appendix B

MCCQE part II scoring scheme

Approach to Physical Examination [One bubble only]	Inadequate performance	Marginal performance	Adequate performance	Superior performance
<ul style="list-style-type: none"> ● Selects appropriate manoeuvres for the patient problem ● Performs examination in a logical order ● Demonstrates technical skill for eliciting findings ● Executes the examination respectfully ● Attentive to the patient's physical comfort and dignity 	○	○	○	○
Attentiveness to Ethical Issues (as relevant to this case) [One bubble only]	Inadequate performance	Marginal performance	Adequate performance	Superior performance
<ul style="list-style-type: none"> ● Exhibits professional behaviours including compassion, respect, and maintenance of confidentiality ● Addresses ethical issues (e.g., disclosure and consent) with clarity and respect ● Demonstrates awareness of and adherence to regulatory and legal requirements 	○	○	○	○
Collaboration (non-patient / colleague) [One bubble only]	Inadequate performance	Marginal performance	Adequate performance	Superior performance
<ul style="list-style-type: none"> ● Engages respectfully (verbal and non-verbal) ● Discusses pertinent information ● Shares responsibility for safe and effective patient care ● Manages differences and resolves conflicts to enhance collaboration 	○	○	○	○
Interpersonal Behaviour [One bubble only]	Inadequate performance	Marginal performance	Adequate performance	Superior performance
<ul style="list-style-type: none"> ● Demonstrates respectful management of the interaction (e.g., non-judgmental, culturally sensitive, avoids interrupting) ● Listens appropriately to facilitate conversation ● Uses appropriate body language; remains composed ● Avoids offensive or aggressive behavior 	○	○	○	○
Interviewing Skills [One bubble only]	Inadequate performance	Marginal performance	Adequate performance	Superior performance
<ul style="list-style-type: none"> ● Gathers relevant biomedical and psychosocial information ● Establishes a timeline or sequence of events ● Elicits pertinent positives and negatives as they relate to the differential diagnosis ● Uses open and closed-ended questions ● Avoids jargon and leading questions ● Clarifies and summarizes what the patient has said 	○	○	○	○
Organization of Encounter [One bubble only]	Inadequate performance	Marginal performance	Adequate performance	Superior performance
<ul style="list-style-type: none"> ● Pursues a purposeful encounter with a logical flow ● Explores the most pertinent data; does not lose time on less relevant information ● Intervenes with the patient as appropriate 	○	○	○	○
Patient-Centred Approach to History-Taking / Physical Examination [One bubble only]	Inadequate performance	Marginal performance	Adequate performance	Superior performance
<ul style="list-style-type: none"> ● Explores patient's experience with health problem (e.g., feelings, ideas, impact, and expectations) ● Attends to patient's verbal and/or physical responses ● Attempts to understand patient's psychosocial context 	○	○	○	○

Patient-Centred Approach to Treatment Planning [One bubble only]	Inadequate performance	Marginal performance	Adequate performance	Superior performance
<ul style="list-style-type: none"> Engages patient in healthcare planning (e.g., encourages questions, discussion and feedback) Integrates treatment approach with patient's psychosocial context and priorities Encourages adherence to treatment by finding common ground 	○	○	○	○

Patient Education [One bubble only]	Inadequate performance	Marginal performance	Adequate performance	Superior performance
<ul style="list-style-type: none"> Provides clear, concise, and accurate information Provides a summary of expected disease course, progression, or resolution Verifies that the information is understood Avoids jargon Facilitates informed decision-making by exploring risks and benefits of each option Discusses timeline for follow-up 	○	○	○	○

Patient Summary Report: Assessment and Management Plan [One bubble only]	Inadequate performance	Marginal performance	Adequate performance	Superior performance
<ul style="list-style-type: none"> Reports a well-reasoned plan that considers impact on patient Provides rationale for treatment 	○	○	○	○

Patient Summary Report: Key Issues [One bubble only]	Inadequate performance	Marginal performance	Adequate performance	Superior performance
<ul style="list-style-type: none"> Report reflects comprehensive understanding of key issues Report includes pertinent negative/normal findings, in addition to the positive/abnormal issues Communication is concise, with no irrelevant information 	○	○	○	○

Patient Summary Report: Organization [One bubble only]	Inadequate performance	Marginal performance	Adequate performance	Superior performance
<ul style="list-style-type: none"> Report is well organized, coherent, and logical Reflects a complete information transfer 	○	○	○	○

Appendix C

Ethics clearance



Office of Research Ethics
5110 Human Computer Interaction Bldg | 1125 Colonel By Drive
Ottawa, Ontario K1S 5B6
613-520-2600 Ext: 2517
ethics@carleton.ca

CERTIFICATION OF INSTITUTIONAL ETHICS CLEARANCE

The Carleton University Research Ethics Board-A (CUREB-A) has granted ethics clearance for the research project described below and research may now proceed. CUREB-A is constituted and operates in compliance with the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans* (TCPS2).

Ethics Protocol Clearance ID: Project # 109402

Project Team Members: Katharina Klassen (Primary Investigator)
Dr. Michael Rodgers (Research Supervisor)

Project Title: Assessing the Active Listening Behaviours of Non-Native English Speaking Medical Candidates in Canada

Funding Source (If applicable):

Effective: **September 10, 2018**
2019.

Expires: **September 30,**

Please ensure the study clearance number is prominently placed in all recruitment and consent materials: CUREB-A Clearance # 109402.

Restrictions:

This certification is subject to the following conditions:

1. Clearance is granted only for the research and purposes described in the application.
2. Any modification to the approved research must be submitted to CUREB-A via a Change to Protocol Form. All changes must be cleared prior to the continuance of the research.

3. An Annual Status Report for the renewal of ethics clearance must be submitted and cleared by the renewal date listed above. Failure to submit the Annual Status Report will result in the closure of the file. If funding is associated, funds will be frozen.
4. A closure request must be sent to CUREB-A when the research is complete or terminated.
5. Should any participant suffer adversely from their participation in the project you are required to report the matter to CUREB-A.

Failure to conduct the research in accordance with the principles of the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans 2nd edition* and the *Carleton University Policies and Procedures for the Ethical Conduct of Research* may result in the suspension or termination of the research project.

Upon reasonable request, it is the policy of CUREB, for cleared protocols, to release the name of the PI, the title of the project, and the date of clearance and any renewal(s).

Please contact the Research Compliance Coordinators, at ethics@carleton.ca, if you have any questions.

CLEARED BY:

Date: September 10, 2018



Bernadette Campbell, PhD, Chair, CUREB-A



Natasha Artemeva, PhD, Vice-Chair, CUREB-A

Appendix D

List of illustrative control group video codes and corresponding video length and content

Video code	Video length (minutes)	Demographics	Context
PA ³ 1 +	4:30	Doctor-M Patient-F	Consultation
PA 2 +	2:58	Doctor-F Patient-F	Consultation
PA 3 +	3:05	Doctor-M Patient-F	Consultation
PA 4 +	4:52	Doctor-M Patient-M	Consultation
PA 5 +	4:00	Doctor-M Patient-F	Consultation
PA 6 +	3:45	Doctor-M Patient-F	Consultation
PA 7 +	2:35	Doctor-M Patient-F	Consultation
PA 8 +	2:26	Doctor-F Patient-F	Consultation
PA 9 +	1:42	Doctor-M Patient-F	Consultation and Physical Examination
PA 10 +	1:00	Doctor-M Patient-F	Consultation and Physical Examination

³ Refers to Physicians Apply, an organization affiliated with the Medical Council of Canada and the source of the control group videos.

Appendix F

List of non-illustrative control group video codes and corresponding video length and content

Video code	Video length (minutes)	Demographics	Context
PA 1 -	2:51	Doctor-M Patient-M	Consultation
PA 2 -	2:45	Doctor-F Patient-F	Consultation
PA 3 -	2:18	Doctor-M Patient-F	Consultation
PA 4 -	3:12	Doctor-M Patient-M	Consultation
PA 5 -	3:20	Doctor-M Patient-F	Consultation
PA 6 -	2:41	Doctor-M Patient-F	Consultation
PA 7 -	2:50	Doctor-M Patient-F	Consultation
PA 8 -	1:35	Doctor-F Patient-F	Consultation
PA 9 -	1:26	Doctor-F Patient-F	Consultation and Physical Examination
PA 10 -	00:30	Doctor-F Patient-F	Consultation
PA 11 -	00:47	Doctor-F Patient-F	Consultation

Appendix F

List of test group video codes and corresponding video length and content

Video code	Video length (minutes)	Demographics	Context
Lake ⁴ -Doc1-1of5	10	Doctor-F Patient-M	History and physical condition consultation
Lake-Doc1-2of5	10	Doctor-F Patient-F	History and physical examination
Lake-Doc1-3of5	10	Doctor-F Patient-M	Breaking bad news
Lake-Doc1-4of5	10	Doctor-F Patient-M	History and physical examination
Lake-Doc1-5of5	10	Doctor-F Patient-F	History and emotional/psychological consultation
Lake-Doc2-1of5	10	Doctor-F Patient-M	History and physical condition consultation
Lake-Doc2-2of5	10	Doctor-F Patient-F	History and physical examination
Lake-Doc2-3of5	10	Doctor-F Patient-M	Breaking bad news
Lake-Doc2-4of5	10	Doctor-F Patient-M	History and physical examination
Lake-Doc2-5of5	10	Doctor-F Patient-F	History and emotional/psychological consultation
Lake-Doc3-1of5	10	Doctor-M Patient-M	History and physical condition consultation
Lake-Doc3-2of5	10	Doctor-M Patient-F	History and physical examination

⁴ Refers to the name of the project's partner at AIMGA and C2I who provided the test group videos.

Lake-Doc3-3of5	10	Doctor-M Patient-M	Breaking bad news
Lake-Doc3-4of5	10	Doctor-M Patient-M	History and physical examination
Lake-Doc3-5of5	10	Doctor-M Patient-F	History and emotional/psychological consultation
Lake-Doc4-1of5	10	Doctor-M Patient-M	History and physical condition consultation
Lake-Doc4-2of5	10	Doctor-M Patient-F	History and physical examination
Lake-Doc4-3of5	10	Doctor-M Patient-M	Breaking bad news
Lake-Doc4-4of5	10	Doctor-M Patient-M	History and physical examination
Lake-Doc4-5of5	10	Doctor-M Patient-F	History and emotional/psychological consultation
Lake-Doc5-1of5	10	Doctor-M Patient-M	History and physical condition consultation
Lake-Doc5-2of5	10	Doctor-M Patient-F	History and physical examination
Lake-Doc5-3of5	10	Doctor-M Patient-M	Breaking bad news
Lake-Doc5-4of5	10	Doctor-M Patient-M	History and physical examination
Lake-Doc5-5of5	10	Doctor-M Patient-F	History and emotional/psychological consultation

Appendix G

Results for individual active and anti-active verbal and non-verbal listening traits

Positive or negative listening trait type	Dependent video set	Fixed video set	<i>p</i> -value
Non-reflective commentary	Illustrative control group	Non- illustrative control group	0
	Illustrative control group	Test group	0
	Non-illustrative control group	Illustrative control group	0
	Non-illustrative control group	Test group	0
	Test group	Illustrative control group	0
	Test group	Non-illustrative control group	0
Exclusive language (e.g. medicalese)	Illustrative control group	Non-illustrative control group	0
	Illustrative control group	Test group	0
	Non-illustrative control group	Illustrative control group	0
	Non-illustrative control group	Test group	0
	Test group	Illustrative control group	0
	Test group	Non-illustrative control group	0
Non-reflective question	Illustrative control group	Non-illustrative control group	0
	Illustrative control group	Test group	0
	Non-illustrative control group	Illustrative control group	0
	Non-illustrative control group	Test group	0.771
	Test group	Illustrative control group	0
	Test group	Non-illustrative control group	0.348
Inaccurate summarization of patient's commentary	Illustrative control group	Non-illustrative control group	0
	Illustrative control group	Test group	0
	Non-illustrative control group	Illustrative control group	0
	Non-illustrative control group	Test group	0
	Test group	Illustrative control group	0
	Test group	Non-illustrative control group	0
Reflective commentary	Illustrative control group	Non-illustrative control group	0.219
	Illustrative control group	Test group	0.168
	Non-illustrative control group	Illustrative control group	0.371
	Non-illustrative control group	Test group	0.255
	Test group	Illustrative control group	0.473

	Test group	Non-illustrative control group	0.492
Vocalized affirmation	Illustrative control group	Non-illustrative control group	0
	Illustrative control group	Test group	0
	Non-illustrative control group	Illustrative control group	0
	Non-illustrative control group	Test group	0
	Test group	Illustrative control group	0
	Test group	Non-illustrative control group	0.498
Inclusive language (e.g. lack of medicalese or medicalese is explained)	Illustrative control group	Non-illustrative control group	0
	Illustrative control group	Test group	0
	Non-illustrative control group	Illustrative control group	0
	Non-illustrative control group	Test group	0
	Test group	Illustrative control group	0
	Test group	Non-illustrative control group	0
Reflective question	Illustrative control group	Non-illustrative control group	0.471
	Illustrative control group	Test group	0
	Non-illustrative control group	Illustrative control group	0.386
	Non-illustrative control group	Test group	0.501
	Test group	Illustrative control group	0.347
	Test group	Non-illustrative control group	0.286
Whole or partial repetition of patient's previous statement	Illustrative control group	Non-illustrative control group	0.239
	Illustrative control group	Test group	0.435
	Non-illustrative control group	Illustrative control group	0.192
	Non-illustrative control group	Test group	0.325
	Test group	Illustrative control group	0.462
	Test group	Non-illustrative control group	0.391
Accurate summarization of patient's commentary	Illustrative control group	Non-illustrative control group	0.473
	Illustrative control group	Test group	0
	Non-illustrative control group	Illustrative control group	0.618
	Non-illustrative control group	Test group	0
	Test group	Illustrative control group	0.531
	Test group	Non-illustrative control group	0.509
Demonstrates distraction with hands	Illustrative control group	Non-illustrative control group	0
	Illustrative control group	Test group	0.224
	Non-illustrative control group	Illustrative control group	0

(e.g. writing, playing with hair, etc.)	Non-illustrative control group	Test group	0.772
	Test group	Illustrative control group	0.178
	Test group	Non-illustrative control group	0
Eye contact (occurrence (“O”) and duration (“D”))	Illustrative control group	Non-illustrative control group	O:0 D:0
	Illustrative control group	Test group	O:0 D:0
	Non-illustrative control group	Illustrative control group	O:0 D:0
	Non-illustrative control group	Test group	O:0 D:0
	Test group	Illustrative control group	O:0 D:0
	Test group	Non-illustrative control group	O:0 D:0
General commentary unrelated to patient’s commentary which is accessible to patient	Illustrative control group	Non-illustrative control group	0.833
	Illustrative control group	Test group	0
	Non-illustrative control group	Illustrative control group	0.226
	Non-illustrative control group	Test group	0
	Test group	Illustrative control group	0.446
	Test group	Non-illustrative control group	0.710
General commentary unrelated to patient’s commentary which is inaccessible to patient	Illustrative control group	Non-illustrative control group	0
	Illustrative control group	Test group	0
	Non-illustrative control group	Illustrative control group	0
	Non-illustrative control group	Test group	0
	Test group	Illustrative control group	0
	Test group	Non-illustrative control group	0.062
Beat hand gesture	Illustrative control group	Non-illustrative control group	0
	Illustrative control group	Test group	0.385
	Non-illustrative control group	Illustrative control group	0
	Non-illustrative control group	Test group	0.450
	Test group	Illustrative control group	0
	Test group	Non-illustrative control group	0
Metaphoric hand gesture	Illustrative control group	Non-illustrative control group	0.221
	Illustrative control group	Test group	0.323
	Non-illustrative control group	Illustrative control group	0.465

	Non-illustrative control group	Test group	0.261
	Test group	Illustrative control group	0.717
	Test group	Non-illustrative control group	0.564
Deictic hand gesture	Illustrative control group	Non-illustrative control group	0.607
	Illustrative control group	Test group	0.361
	Non-illustrative control group	Illustrative control group	0
	Non-illustrative control group	Test group	0.354
	Test group	Illustrative control group	0
	Test group	Non-illustrative control group	0.296
Interactive hand gesture	Illustrative control group	Non-illustrative control group	0
	Illustrative control group	Test group	0
	Non-illustrative control group	Illustrative control group	0
	Non-illustrative control group	Test group	0
	Test group	Illustrative control group	0.430
	Test group	Non-illustrative control group	0.298
Iconic hand gesture	Illustrative control group	Non-illustrative control group	0
	Illustrative control group	Test group	0.423
	Non-illustrative control group	Illustrative control group	0
	Non-illustrative control group	Test group	0.464
	Test group	Illustrative control group	0.351
	Test group	Non-illustrative control group	0.311
Interruption	Illustrative control group	Non-illustrative control group	0.605
	Illustrative control group	Test group	0.154
	Non-illustrative control group	Illustrative control group	0.187
	Non-illustrative control group	Test group	0.157
	Test group	Illustrative control group	0.316
	Test group	Non-illustrative control group	0.304
Head nod	Illustrative control group	Non-illustrative control group	0.291
	Illustrative control group	Test group	0
	Non-illustrative control group	Illustrative control group	0
	Non-illustrative control group	Test group	0
	Test group	Illustrative control group	0
	Test group	Non-illustrative control group	0.325
	Illustrative control group	Non-illustrative control group	0:0

Lack of eye contact when sought by patient (occurrence “O” and duration “D”)			D:0
	Illustrative control group	Test group	O:0.392 D:0.527
	Non-illustrative control group	Illustrative control group	O:0.180 D:0.221
	Non-illustrative control group	Test group	O:0.368 D:1.000
	Test group	Illustrative control group	O:0.249 D:0.257
	Test group	Non-illustrative control group	O:0 D:1.000
Demonstrates distraction with head (e.g. visibly shifts head away from patient during patient’s commentary)	Illustrative control group	Non-illustrative control group	0
	Illustrative control group	Test group	0
	Non-illustrative control group	Illustrative control group	0
	Non-illustrative control group	Test group	0.134
	Test group	Illustrative control group	0
	Test group	Non-illustrative control group	0.403

Appendix H

Descriptive statistics for positive listening set

Group	Mean	SD	N
Illustrative control group	24.66	9.448	10
Non-illustrative control group	18.53	9.651	11
Test group	34.06	8.359	25

Appendix I

Descriptive statistics for negative listening set

Group	Mean	SD	N
Illustrative control group	1.920	2.080	10
Non-illustrative control group	9.139	7.822	11
Test group	3.607	2.913	25

Appendix J

Descriptive statistics for beats

Group	Mean	SD	N
Illustrative control group	4.739	3.604	10
Non-illustrative control group	5.547	3.659	11
Test group	8.477	2.824	25

Appendix K

Descriptive statistics for eye contact frequency

Group	Mean	SD	N
Illustrative control group	11.115	6.495	10
Non-illustrative control group	13.837	7.294	11
Test group	6.673	3.709	25

Appendix L

Descriptive statistics for eye contact duration

Group	Mean	SD	N
Illustrative control group	5.866	2.313	10
Non-illustrative control group	4.631	2.665	11
Test group	3.463	2.901	25

Appendix M

General reviewer commentary following mock OSCEs performed by test group

Video	Further examination or diagnosis required	# of items needed to be further assessed by the doctor	Able to identify general diagnosis	Additional, incorrect diagnoses made	Positive acknowledgment of doctor's communication skills by reviewer
Lake-Doc1-1of5	Yes	3	Yes	No	Yes
Lake-Doc1-2of5	No	-	Yes	Yes	No
Lake-Doc1-3of5	-	-	-	-	No
Lake-Doc1-4of5	No	-	Yes	No	No
Lake-Doc1-5of5	Yes	Unclear	Yes	No	No
Lake-Doc2-1of5	Yes	6	Yes	No	No
Lake-Doc2-2of5	No	-	Yes	No	No
Lake-Doc2-3of5	-	-	-	-	No
Lake-Doc2-4of5	No	-	Yes	No	No
Lake-Doc2-5of5	No	-	Yes	Yes	No
Lake-Doc3-1of5	Yes	1	Yes	No	No
Lake-Doc3-2of5	Yes	1	Yes	No	No
Lake-Doc3-3of5	-	-	-	-	No
Lake-Doc3-4of5	No	-	Yes	No	No
Lake-Doc3-5of5	Yes	Unclear	Yes	No	No
Lake-Doc4-1of5	Yes	4	Yes	No	No
Lake-Doc4-2of5	No	-	Yes	No	No
Lake-Doc4-3of5	-	-	-	-	No

Lake-Doc4-4of5	No	-	Yes	No	No
Lake-Doc4-5of5	No	-	Yes	Yes	No
Lake-Doc5-1of5	Yes	4	Yes	No	Yes
Lake-Doc5-2of5	No	-	Yes	No	No
Lake-Doc5-3of5	-	-	-	-	No
Lake-Doc5-4of5	No	-	Yes	No	No
Lake-Doc5-5of5	Yes	1	Yes	No	No
Overall Occurrences	9	20	20	3	2
# of Occurrences Overall (%)	36	-	100	12	8
<i>Notes.</i> The use of the dash (-) indicates that the number of items was not made clear by the reviewers' comments.					