PROSODIC PROFILES: SUSPECTS' SPEECH DURING POLICE INTERVIEWS

by

Lorna Fadden
M.A., Simon Fraser University, 2000
B.A., Simon Fraser University, 1996

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

In the
Department
of
Linguistics

© Lorna Fadden 2008

SIMON FRASER UNIVERSITY

Spring 2008

APPROVAL

Name: Lorna Fadden

Degree: Doctor of Philosophy

Title of Dissertation: Prosodic Profiles: Suspects' Speech during Police Interviews

Examining Committee:

Chair: Dr. Chunghye Han .
Associate Professor, Department of Linguistics

Dr. Nancy Hedberg
Senior Supervisor
Associate Professor, Department of Linguistics

Dr. Murray Munro
Supervisor
Professor, Department of Linguistics

Dr. Juan M. Sosa
Supervisor
Associate Professor, Department of Linguistics

Dr. Deborah Connelly
Internal Examiner
Associate Professor, Department of Psychology
Simon Fraser University

Dr. Alison Johnson
External Examiner
Lecturer, School of English
University of Leeds

Date Defended/Approved: Feb 14th, 2008

SFU SIMON FRASER UNIVERSITY
THINKING OF THE WORLD

STATEMENT OF
ETHICS APPROVAL

The author, whose name appears on the title page of this work, has obtained,
for the research described in this work, either:

(a) Human research ethics approval from the Simon Fraser University Office of
Research Ethics,

or

(b) Advance approval of the animal care protocol from the University Animal
Care Committee of Simon Fraser University;

or has conducted the research

(c) as a co-investigator, in a research project approved in advance,

or

(d) as a member of a course approved in advance for minimal risk human
research, by the Office of Research Ethics.

A copy of the approval letter has been filed at the Theses Office of the
University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the
relevant offices. Inquiries may be directed to those authorities.

Bennett Library
Simon Fraser University
Burnaby, BC, Canada

ABSTRACT

This dissertation presents a descriptive study of the prosodic characteristics of suspects' speech during investigative interviews with police. During police interviews, investigators direct conversation by asking suspects questions and making assertions thereby placing the suspect in the position of responding. Based on whether the suspect is a first-time or repeat offender, and the type of information suspects produce, responses are categorized and examined for their properties of pause, tempo, and pitch.

Response types explored in this study are *affirmative*, in which suspects confirm information in the investigators' questions or assertions; *negative*, in which information is rejected; *relevant* and *irrelevant*, in which suspects offer information pertaining or not pertaining to the investigators' questions; and *confessions*. Pausal features — response latency, and pause-to-speech ratio — are found to differ across certain response types in both groups. In general, for example, first-time suspects pause more than repeat offenders, both before and during turns, particularly when offering *relevant* responses. Among the temporal features, first-time suspects' speech and articulation rates are lower when producing *relevant* information than repeat offenders' rates. Furthermore, first-timers' *irrelevant* temporal rates are higher than *relevant* temporal rates. Pitch

characteristics show less distinction across response types than pause and tempo, although first time suspects' pitch values cluster somewhat more consistently within response types than repeat offenders', whose pitch values vary more widely.

The findings noted above are discussed in relation to previous studies that address the prosodic characteristics of discourse and emotion. Furthermore, on the basis of these findings, prosodic profiles of response categories for first-time offenders and repeat offenders are created. The potential for forensic application of these profiles is discussed, particularly with respect to deception.

In loving memory of James

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 PROSODY AND POLICE INTERVIEWS

> *...the Orientals used a deception test based on sound principles. The accused was given some rice to chew and if he were unable to spit it out he was deemed guilty, and if he could spit it out, innocent. (Larson, 1932: 65).*

This introductory chapter has four aims. First I discuss in very general terms what a police interview is, and why it is worthy of academic interest. I then introduce key concepts relevant to a discussion of the prosodic features that accompany the responses that criminal suspects produce during police questioning during an investigative interview. Second, the chapter continues with an overview of the literature on two distinct areas. I survey the studies addressing prosody as it is affected by discourse, and prosody as it is affected by emotion, after which comes a survey of the studies performed on police interviews and deception, both real and simulated. Third, I discuss the contribution of this dissertation, to situate it among the surveyed areas of study. Fourth, is the introduction of the research questions to be answered in subsequent chapters. The chapter closes with an outline of the remainder of the dissertation.

1.1 Introduction

As long as there have been laws, there have been members of society charged with the responsibility of upholding the laws. Regardless of geographic location or cultural background, those who deal with citizens who are accused of deviating from acceptable behaviour, are engaged in some form of policing. And an important part of policing is questioning those who are believed to have been involved in a criminal act.

Police interviews in North America are popularly known to be one-on-one, intense, rapid-fire questioning periods, that, thanks to television and movies, may even involve physical brutality to persuade the undoubtedly guilty party to confess. In North America, however, 'third degree'[1] tactics such as physical abuse are no longer tolerated in the interview room, and have not been for several decades. While some may argue that physical coercion has been replaced by psychological coercion, that debate is not addressed here. Rather, the purpose of this dissertation is to explore how suspects respond when questioned by police in a contemporary, Western Canadian context, and more specifically, what kind of prosodic features are present in suspects' voices when responding. This purpose will be reiterated again below, with elaboration.

---

[1] The term 'third degree' is believed to be borrowed from Freemasonry, making reference to the rigorous adjudication process through which one must pass before attaining the rank of master (Deakin, 1988).

During criminal investigations, it is standard practice for investigating

officers to interview those who may have committed a criminal act.[2] The type of

dialogue that takes place between an investigator and a suspect is commonly

referred to as a police interview. During an interview, the investigator, or

interviewer, primarily performs two conversational tasks: she or he asks the

suspect questions about the suspect's knowledge of the events of interest, and

she or he makes assertions about the suspect's actions. The primary task for the

suspect, or interviewee, is to respond to questions with relevant information and

confirm or reject the investigator's assertions. Of course, this is a highly

simplified model of the interactive roles played by each. However, it captures

fairly succinctly what takes place in a police interview.[3]

The following two extracts provide examples of exchanges that occurred

in the corpus I constructed (see Chapter Two), to provide a picture of what is

found in the interviews in this study. The first extract, (1), comes from an

interview with a 31-year-old male with no prior convictions or interaction with

police, suspected of physically abusing his 5-month-old child who suffered a

fractured skull. In his account, he asserts that the child fell from a carseat placed

---

[2] Police also conduct interviews with people who witness actions and events. However, this type of interview will not be of concern here. While witness interviews are worthy of study in their own right, a corpus of suspect and witness interviews would be akin to comparing apples and oranges. For suspects who face potential legal consequences, linguistic behavior is likely to be quite different than for witnesses who face lesser or no legal consequences.

[3] In the early stages of most interviews, the suspect is invited to discuss on their own terms their knowledge of events or circumstances, however in none of the interviews in this corpus did any suspect choose to do so. Therefore these interviews consist largely of this question/response, assertion/response pattern.

on a table, sharply contradicting the medical examiner's report which concluded that the child had suffered a blunt force trauma to the head. "I" indicates the investigator's speech, and "S" indicates the suspect's.

Extract (1)

(1)   I: Where was your baby when you found him?

(2)   S: Ok, this is the kitchen counter, this is the kitchen counter, [*gesturing with hands*] this is where I make the milk, the microwave's here, I go that way to the washroom, when I come back, he's on his back this way. And--

(3)   I: Where's the chair?

(4)   S: I know you say, the chair is still here, the carseat is still facing this way.

(5)   I: And where is he?

(6)   S: Here, back. On his back.

(7)   I: Where's the bar under the chair?

(8)   S: Pardon me?

(9)   I: Where's the bar? You said there's a bar under the chair.

(10)   S: Yes, it's here. It it runs — [*gesturing with his hand how the bar runs along the floor and up to support the seat of the chair*]

(11)   I: Now where is his head in relation to that bar?

(12)   S: In front of it.

(13)   I: How far?

(14)   S: Right there. Now I know you say he should have fallen on his face,

(15)   I: Yeah.

(16)   S: I gotta tell you how I found him. I'm not gonna say, ok, now he was on his face. That's how I found — I should have said, if I wanted it to be

consistent, I found him on his right side with his temple on that bar —

(17)   I: No, not on the bar you said--

(18)   S: No, now now why wouldn't I say that to be consistent with where it happened. Why why wouldn't I say I put him on the table to make the fall look more great? You know, I'm not a dummy, why why

(19)   I: I don't think you're a dummy.

Later in the interview:

(20)   I: Did you hurt your son?

(21)   S: No sir, I did not.

Extract (1) illustrates the question-answer pairs led by the investigator. On three separate occasions, lines 3, 10, and 17, the investigator demonstrates control over the conversation when he interrupts the suspect in order to continue his line of questioning. The suspect complies, and offers information relevant to the investigator's questions. Extract (2) comes from an interview with a 26-year-old suspect with a number of prior convictions, alleged to have stolen multiple vehicles in a crime spree involving several sexual assaults.

Extract (2)

(1)   I: Where was that? *(location of vehicle theft)*.

(2)   S: Downtown. I went and got a bunch of dope, cus there's a guy down there, drug dealer right? Thinks he's big right? And I don't like the guy to begin with so I said fuck it, I'm gonna rob him.

(3)   I: Ok.

(4)   S: So I went and punched his face in in the alley.

(5)   I : That night?

(6)   S: Yeah, that night.

(7)   I: What's his name?

(8)   S: Aah, it doesn't matter what his name is.

(9)   I: It is. It is kind of important what his name is
        because he's probably, he has probably made a
        complaint against you that, that uh, I don't know
        about yet.

(10)   S: He's a drug dealer! Come on man! I'm well sure,
        they're fuckin rats man. Look at the heads on them!
        I did you guys a fuckin favour.

Extract (2) illustrates an exchange in which the suspect is much more

confrontational than the suspect in (1). However, the investigator-led framework

still applies, and we see the investigator asking the questions, challenging the

suspect, and making assertions, and the suspect responds, even if

antagonistically.

The two extracts above introduce a number of aspects that will appear in

this study. First, as previously mentioned, police interviews exhibit a fairly rigid

framework in which the investigators do the questioning and asserting, which

leaves the suspect in the position of responding. Second, we're given a glimpse

into the kinds of responses that suspects produce. For example, they can offer a

confirmation of the information contained in the investigator's question, as

shown in Extract (1) turn 10, and Extract (2), turn 6; a rejection as in Extract (1),

turn 21; or they can volunteer some amount of relevant information, usually in

response to a WH-question as shown in Extract (1), turns 2, 14 and 16, and

Extract (2), turn 2. Other response types will be defined at later stages of the dissertation. Third, we are also shown two different types of suspects. Extract (1) comes from an interview with a first time suspect, while extract (2) comes from an interview with a suspect who has had more than one prior conviction, and hence plenty of opportunity to find himself in an interview with police. It will be shown that first time offenders and repeat offenders show different linguistic behaviour when interviewed by police. Fourth, are the vocal characteristics that accompany the responses that the suspects give, and in fact, the goal of this dissertation is to examine the prosody, that is, pause, tempo, and pitch for the different types of responses exemplified above, comparing first time and repeat offenders. The prosodic characteristics will be identified below, and defined in Chapter Two.

Continuing on with discussion of some of the features of police interviews as a conversational genre, I identify the goals of the participants engaged, i.e., the investigators and suspects. While this dissertation does not employ frameworks of Conversation Analysis (CA), and Discourse Analysis (DA), it is convenient to use CA and DA terms to describe the type of dialogue that a police interview is. The police interview therefore, is a type of institutional discourse, an exchange between a professional and a layperson for the purpose of conducting an institutional task. What sets institutional discourse apart from casual discourse are the roles played by the speaking participants, and how they orient

themselves to the context and manage their turns to accomplish the institutional

task at hand (Drew & Heritage, 1992).

Drew and Heritage (1992: 22) demonstrate how in institutional settings,

'both lay and professional participants generally show an orientation to

institutional tasks or functions in the design of their conduct. By listing a small

set of institutional discourse examples, one gets a clear picture of the goals

shared by participants. These include service counter interaction in which the

customer and clerk negotiate a commercial exchange (Peräkylä & Vehvilänen,

2003; Ylanne-McEwen, 2004), classroom interaction, in which put simply,

teachers prompt students, students respond, and teachers then provide feedback

(Hellerman, 2005; Wells, 1993), doctor-patient interaction (Frenkle, 1990;

Maynard, 1991; ten Have, 1991), where the doctor asks questions and the patient

provides information so that the doctor may make a diagnosis. The common

thread throughout these studies, within the CA framework, is the participants'

orientation toward the collaborative accomplishment of the institutional task,

and the maintenance of their own institutional roles in which one participant is

the knowledgeable professional and the other is the layperson in need.

The difference between police interviews and other types of institutional

discourse is that the participants' goals are often quite divergent, as shown in the

extracts above. The investigator aims to uncover information regarding the

suspect, and in many cases, the suspect aims to conceal that information or

provide information that is not helpful in achieving the institutional goal.

(Fadden, 2006a; Levy, 1999; Walton, 2003).

Taking a more philosophical approach, Walton (2003) labels the interrogation as "a subspecies of information-seeking dialogue"(p.1775) with a set of normative rules for each participant, setting up a non-mutual, adversarial, type of exchange. The rules, listed below, "express the goals and practical strategies of the participants" (p. 1780) and present a concise demarcation of the opposing orientations that CA describes. They are as follows:

> Rule 1: The respondent needs to take care not to inadvertently say something that might give out the information that he wants to conceal, or allow the proponent to infer it.
>
> Rule 2: The proponent may coerce the respondent to reveal information through threats or sanctions but only by the means allowed.[4]
>
> Rule 3: The proponent needs to pose questions to the respondent and these questions can and often should be leading, loaded and deceptive.
>
> Rule 4: The respondent should answer in formulations that are vague, ambiguous, misleading or confusing, if that will serve his ends.
>
> Rule 5: The proponent should probe critically into the respondent's prior replies, and try to use them to extract information.
>
> Rule 6: The respondent should take care to try to be consistent in his replies and in the commitments that can be inferred from them.

---

[4] I am not aware of any Canadian police agency that allows overt coercion or threat, but 'interrogation' is also conducted on suspects of, for example, terrorism or other political activity, by non-police agencies, and they are governed by other policies which may condone the use of such means.

Rule 7: If the proponent finds inconsistencies in the respondent's commitments, or implausible statements, or statements that are inconsistent with information from other sources, she should ask questions that critically examine them.

Rule 8: If the proponent extracts the information that she wants from the respondent, then she has achieved her goal and the dialogue concludes in her favour.

Rule 9: If the proponent terminates the interrogation without getting the information she wants, and the respondent preserves his interests, the dialogue concludes in the respondent's favour.

Rule 10: The two parties can use any arguments, even ones considered irrelevant or fallacious from the viewpoint of a critical discussion, to achieve their ends. (p. 1780)

The rules illustrate quite well the more immediate concerns of the participants of police interviews, in particular suggesting how participants might advance their interests as an interview proceeds. Walton concluded that the goal of an interrogation, of which police interviews are one type, is to obtain correct and complete information that will be useful in the pursuance of an investigation to create a case for trial. Others claim that the goal is to secure a confession, a point I return to below.

With police interviews adequately defined as a type of institutional discourse, I identify key terms related to police interviews, and I offer a definition of each as they will be used in this dissertation.

Police interview: a type of institutional discourse in which a police investigator questions a suspect in relation to the suspect's involvement in a

criminal event or events. All interviews that are the focus of this work take place at police headquarters in designated interview rooms.

Investigator: the police official who conducts the investigative interview with a suspect. All interviews in my corpus are conducted by officers at the rank of detective or inspector.

Suspect: the citizen who has been arrested and brought to police headquarters for questioning, or who is not under arrest but who has agreed to come to police headquarters for questioning regarding their involvement in a crime or crimes.

Prompt: a question or assertion made by the investigator addressed to the suspect.

Response: the suspect's response to a prompt. In Chapter 2, response categories will be defined.

1.1.1 Why study police interviews?

Besides its being inherently intriguing as one of many institutional discourse genres, the police interview is worth examining closely for legal and forensic purposes. Many studies in the literature conclude with remarks on how findings may be applied to improve practice in the field and shed light on an integral part of police investigative work that can have far reaching consequences on the lives of individuals. This dissertation is no exception, and

will address the application of my findings to the disciplines of forensic linguistics and forensic psychology.

1.2 Research on police interviews

Police interviews have not been very widely studied from a linguistic perspective, which hopefully positions this work well as a complement to existing and future linguistic studies. That police interviews have not been in the linguistic spotlight much, does not mean that researchers from other disciplines have not undertaken rigorous study. Below, I survey some of the linguistic work on police interviews, as well as those studies where the boundary blurs between linguistics and other social sciences such as psychology and sociology.

There are a few common themes in the existing research on police interviews in the areas of linguistics, and to some extent, where linguistics intersects with psychology and philosophy. In this section I survey three areas of research and afterwards, discuss how a dissertation on prosody and discourse complements existing work. The first area, often grounded in the theoretical frameworks of Conversation Analysis and Critical Discourse Analysis, is the establishment and maintenance of a power hierarchy between the investigator and the suspect. The second area explores the nature of confessions from psychological and linguistic perspectives. The third area examines laboratory speakers' and some criminal suspects' behaviour in order to better understand

how deception can be detected. These three areas will be dealt with in this order, followed by a discussion of where this dissertation on the prosody and discourse of suspects is situated in the literature on police interviews.

1.2.1 Power hierarchies and institutional discourse

Language used in any official capacity comprises a large conversational genre called institutional dialog (Drew & Heritage, 1992). The police interview is one type of institutional dialog that falls within the legal domain. Others of this subcategory include witness interviews and courtroom discourse. In this section, I introduce a number of studies, which from the perspective of Conversation Analysis and Critical Discourse analysis, demonstrate the imbalance of power that maintains the professional role of the investigator and the layperson role of the suspect.

Institutional discourse is marked by interaction involving at least one participant who is engaged in a professional role. In the police interview, the professional/layperson pairing gives rise to a power imbalance in which the investigator holds the institutional and conversational power, and the suspect does not.

Heydon (2005) discusses at length the extent to which authority is encoded into police turns in interviews with suspects. Whereas under normal conversational practice, participants move from topic to topic in a planned manner, linking one idea to the next, (Jefferson, 1984, citing Sacks, Lecture 5

Spring, 1972) interviewing officers do not employ such strategies. Instead, police

select the topics, smoothly or abruptly, and limit the suspect's ability to initiate

topics. Of this, she notes "suspects do not have access to a role that will allow

them to place information on the floor without invitation — any information

provided by the suspect can only be provided in response to police request"

(2005: p115).

The control over topics and the assignment and taking conversational

floor-time is well described in both Haworth (2006) and Newbury & Johnson

(2006) on an interview with Dr. Shipman, a high profile case of a UK physician

who was alleged to have murdered several of his patients. Both illustrated —

Haworth through CDA, and Newbury & Johnson through Hallidayan systemic

functionalism — the power hierarchy typically associated with institutional

discourse by highlighting the suspect's resistance strategies. Such resistance

strategies included the avoidance of confirmation of investigators' questions and

assertions, topic avoidance, and direct challenges to institutional roles. Both

conclude that while the power hierarchy is well intact throughout the interview,

certain sequences arise leaving open the possibility for the suspect to challenge

the investigator's authority.

In a conversation analysis of police interviews, Fadden (2006a) illustrated

strategies suspects use in order to evade collaboration, and hence attempt to

resist the power hierarchy while still conforming to the typical

question/assertion and response pattern. Suspects can, for example, attempt to change topics within their turns, they can produce related but irrelevant, non-consequential information, they can extol their own virtues as a means of damage control when a line of questioning attacks their character. Suspects can also blame their victims, either for the incident in question, or evaluate them in a derogatory manner, to remove themselves from the focus of questioning. All of these strategies see the suspect adhering to his or her lower status in the interview room, without submitting entirely to the investigator's authority.

While the purpose of her article was not to address the power imbalance between suspect and interviewer, Benneworth (2006) illustrated the suspects' and investigators' differences in perspective toward paedophilia. Her findings demonstrated police preference for using explicit physical reference to sexual vocabulary (e.g. penis, vagina, anal, intercourse, penetration, and so on) while suspects tend to use more informal, or emotionally laden terms (e.g. corresponding to those above, thingy, fanny, between the cheeks, closeness, cuddle, and so on, in UK English). Benneworth asked whether the euphemistic manner in which paedophiles describe contact with their victims should be employed by investigators in interviews to even out the discourse clash between the two parties, which raises an even bigger question: to what extent should investigators linguistically accommodate to suspects speech?

Ainsworth (in press) has made strides toward answering this very question in her work on police interaction with citizens. She has found that when police use coarse language, much the same way suspects often do (as will be seen in later portions of the present work), there comes a sense of inappropriateness in the best case, and intimidation and fear in the worst. Linguistic accommodation then, would seem to backfire in instances where the power hierarchy is quite pronounced.

A resolution to the debate on discourse imbalance then, may be far off. If, when police adopt informal, i.e., non-institutional, manners of discourse, then the power hierarchy may not be maintainable, preventing police from maintaining authority. If they do not, though, the power imbalance may be significant enough to hinder communication.

1.2.2   Confessions

My own conversations with investigating officers make clear the notion that obtaining a confession is the ideal outcome when interviewing a suspect, particularly a confession that is spontaneously produced. The same is found in Baldwin (1993, 1994), who worded it more strongly: "officers are in many cases hell-bent on securing a confession" (1994, p. 67). The importance of a confession to an investigation is clear: confessions provide singularly compelling evidence in finding a defendant guilty (Kassin & Neumann, 1997; Wrightsman & Kassin, 1993). Much of the existing work on confessions focuses on the differences

between confessions spontaneously produced by suspects and confessions which arise more out of investigators' assertions which are met with passive confirmation of details. Other studies focus on understanding the difference between true confessions and false confessions, and what might lead an innocent person to falsely confess to having committed a crime. In this section I survey some of the widely cited work in this area.

False confessions are a troubling issue at all levels of justice system, because not only is the wrong person convicted and sentenced, but the person who committed the crime is not held responsible and may remain free. It would seem difficult to argue against someone's guilt if they admit to having committed a crime even if they did not. However, Hill (2003) devised a means to demonstrate a suspect's guilty knowledge. He examined the degree to which the investigator and the suspect each contribute to the information revealed in an interview. A case study of a false confession produced by a young suspect maintaining loyalty to his peers involved examining question and answer pairs to determine that the investigators were producing details, which the suspect mostly confirmed, rather than the suspect spontaneously describing events and actions. It was shown that the suspect produced none of the critical details that would demonstrate whether he did indeed have genuine guilty knowledge. Hill's response categories inspired three that I describe in the next chapter.

While some suspects readily confess, others do not, and certainly investigators are not guaranteed a confession from the suspect by the time the interview is over. A number of studies show (see below), in fact, that some suspects are more likely to confess than others, and factors contributing to suspects' inclination or disinclination to confess include age, sex, legal representation, previous convictions, among others.

In a study of 160 police interviews conducted in London, UK, Pearse, Gudjonsson, Clare, & Rutter (1998) found that younger suspects were more likely to confess than older suspects. Results showed almost 60% of suspects under the age of 25 confessed, while more than 60% of suspects over 25 denied allegations. They also found that suspects who had consumed illicit drugs in the previous 24 hours were more likely to confess than those who had not. However, alcohol was not found to be a factor. Those who had spoken with a lawyer or had a lawyer present during the interview were also much less likely to confess than those who did not. Self-reported knowledge of legal rights was not found to be a factor.

Other studies have focused less on the suspect himself or herself, and instead looked to police practice to account for those who confess and those who do not. In many studies that approach the police interview from observations of police conduct, this conduct is criticised for being coercive, confrontational, or persuasive. In response to this, a number of studies have shown that when

interviewing officers appear respectful of a suspect, a confession is more likely to be forthcoming than if the investigator is domineering or aggressive. See especially Holmberg & Christianson (2002), and Ressler & Burgess (1988).

The observation that "being nice" in the interview room results in more cooperation and less resistance, from the suspect, i.e., willingness to confess, has given rise to interview techniques and philosophies that emphasize a less condemning, less confrontational manner, particularly when the suspect appears cooperative. Among these are Ethical Interviewing (Shepherd, 1991); Cognitive Interviewing (Fisher & Geiselman, 1992; Fisher, Geiselman, & Amador, 1989; Geiselman, Fisher, MacKinnon, & Holland, 1985), Investigative Interviewing (Milne & Bull, 1999; Williamson, 1993) and Therapeutic Jurisprudence (Finkelman & Grisso, 1996; Slobogin, 1996). In each of these, common approaches see the investigator demonstrating empathy and understanding of the suspect's situation, and the investigator aims to reduce the suspect's anxiety. It might be argued naturally, that sympathy toward the suspect in the interview room is yet another coercive tactic at the interviewer's disposal in moving toward the goal of obtaining a confession, but that debate is well beyond the scope here. I direct the reader to the work of Gudjonsson and colleagues (Gudjonsson, Sigurdsson, Asgeirsdottir, & Sigfusdottir, 2006, 2007; Gudjonsson, Sigurdsson, Bragason, Einarsson, & Valdimarsdottir, 2004; Gudjonsson, Sigurdsson, & Einarsson, 2004; Gudjonsson & Williamson, 2006; Gudjonsson, Young, & Bramham, 2007) for

criticism of police practice. A substantial one-stop review of this work is found in Kassin & Gudjonsson (2004).

I refer to one final study in this section on confessions, which highlights the effect of bias on those who interview suspects. In an experimental setting, Kassin, Goldstein, & Savitsky (2003) demonstrated interrogator bias against guilty parties. In their study, professional interrogators conducted interviews with people who were alleged perpetrators of a mock theft, and were informed of the 'suspects" guilt or innocence. The interrogators were shown to put more pressure on the 'guilty suspects' to obtain a confession than they did when they believed the suspect to be innocent. The lesson to be learned, they cautioned, is that in standard police practice, especially in North America, interrogation takes place only once the suspect is deemed of significant interest (Inbau, Reid, & Buckley, 2001) and hence, the presumption of guilt is already there.

1.2.3 Deception detection

In this section, I provide a brief overview of the relevant work on deception in speech[5]. While there are countless studies that have sought to describe what differentiates lies from truths, I restrict this survey to studies that are linguistically focussed, i.e., they address at least in part, phonetic or other linguistic variables.

---

[5] There is a vast literature on the identification of deception in text as well, and for an introduction to the subject, I refer the reader to the dissertation of Adams (2002).

It is worth mentioning that the initial idea for this dissertation started with the goal to explore the prosody of deception. To explore such a topic, it was apparent that I would need to identify a type of discourse in which lying is likely to be found, without having to construct elaborate laboratory based means. The police interview seemed a natural starting point. However, no work had been done on the prosody of this conversational genre, and hence I was without a starting point. I decided then that I would undertake a study on suspects' prosody in order to establish an understanding of the prosodic features that suspects exhibit under the condition of police questioning.

The relevance of deception to police interviews is obvious. These are situations in which an individual is being questioned about his or her involvement in a criminal act, and the outcome of an investigation may result in punitive consequences including incarceration. It is therefore reasonable to assume that such an individual may be highly motivated to lie, assuming, of course, that they are indeed guilty. While my goal is not to address deception specifically in my study, those who have blazed a trail in this field offer a wealth of theories and methodologies in discourse and speech science as they seek to uncover the cues to what constitutes lying, and how it might be distinguished from telling the truth.

An abundance of work by psychologists has addressed the behaviour of liars. The common approach in all published work on deception of which I am

aware aims at identifying clusters of verbal (i.e., vocal) and non-verbal (i.e., gestural) phenomena, and as yet, I am unaware of any study claiming that a single behaviour might hold the key to distinguishing the difference between truthful and deceptive speech.

Vrij and his colleagues (Vrij, 1995, 2000; Vrij, Edward, & Bull, 2001; Vrij & Mann, 2001) have contributed significantly to this field, figuring prominently in the literature reviews in the subsequent data chapters of this dissertation. Collecting data both in the lab and in the real world, he has contributed to our understanding of how people behave when they lie. His approach has been to examine several variables at once while speakers are engaged in the act of lying, and when they are telling the truth.

Table 1.1 below, focussing on more recent studies for the sake of brevity, gives an overview of the variables explored in a range of studies on deception. I have omitted physical and gestural variables, since none of those are addressed herein. For a much more comprehensive summary of more studies dating further back, see Vrij (2000, p.59-60).

Table 1.1   Summary of findings: features of deceptive speech

Empty fields denote prosodic elements not addressed in a particular study.

Authors	Pitch	Pause	Tempo	Type of data
Mann, Vrij, & Bull (1998)		more pauses		police interview
Anolli & Ciceri (1997)	higher mean pitch and range when suspected of lying	more pauses when suspected of lying	minimal difference	lab
Vrij, Edward, & Bull (2001)		greater latency before speaking more hesitations of all types	faster	lab
Vrij & Mann (2001)		more frequent and longer	slower	police interview
Granhag & Strömwall (2002)		fewer pauses	minimal difference	lab
Rockwell, Buller, & Burgoon (1997)	higher pitch found by coders' impressions	greater latency before speaking	slower	lab

The table above shows that sometimes there is concurrence in findings, as is the case with pitch quite often where lies seem to be told at a higher pitch. Other times, no concurrence is found, as speech rate for example shows opposite results from one study to another, or non-differentiating results. The purpose of surveying studies of deception is to identify the types of prosodic features and modes of measurement as they have been used in forensic contexts. In later chapters, I refer to the variables and methods of deception studies to support the methods of my own work.

1.3 Prosody in discourse and emotion

As this study focuses specifically on the prosodic features of suspects' speech during police interviews, it is necessary to identify key studies that have brought to light the effects of emotional state, and discourse on prosody, and the effect of prosody on what the hearer takes from the message.

First, in the interest of defining the terms as they will be used in later sections, prosody is often taken to be the variation in duration, pitch and loudness of utterances (Lehiste, 1970). While pitch is readily explained as changes in the highs and lows of the voice, duration is a much more widely encompassing dimension because durational measurements of a variety of units can be taken. In this dissertation, duration of pauses is measured, both in terms of between-turn and within-turn lapses in speech. Duration is also used to calculate the frequency of units to analyse speaking tempo. Pauses per unit of time, syllables per unit of time, speech segments per unit are all examined herein, and each is further explained in subsequent and appropriate chapters. The only one of Lehiste's prosodic features I do not examine is loudness, sometimes referred to as intensity. Because I acquired the data from police agencies as they gather it following standard procedure (see Chapter 2), I had no control over the suspects' proximity to recording equipment, and therefore absolute intensity

could not be reliably measured in my data.[6] I do not, however, discount the important contribution that a section on intensity would have made.

I do not intende to attribute the prosodic phenomena studied herein with either discourse behaviour or emotional states of suspects specifically, and in fact that would not be possible without consulting the suspects' themselves. It is, nonetheless, worthwhile to tie together my own findings with those of previous studies, if only as a starting point for discussion on the contribution that prosody makes to speakers' meanings and intentions. Below, I survey the studies that have made broad generalizations on the interface of prosody and emotion, and prosody and discourse. More specific findings and methods are discussed in Chapters Three, Four, and Five, where the prosodic phenomena themselves are reported.

A compilation of common findings across the literature is summarized by Banse and Scherer (2003). In it, they attribute higher mean F0, greater F0 variability, wider F0 range, and a faster rate of articulation with anger. Fear is associated with similar prosodic correlates. Sadness and boredom, on the other hand, see a decrease in all F0 dimensions as well as a decrease in articulation rate. Joy and disgust are more difficult to characterize, with F0 values and articulation rates being more variable, and differing across studies. More specific findings are discussed in later chapters as they become relevant.

---

[6] It may have been possible to measure relative intensity within some speakers, but intensity would be lacking in relation to the prosodic variables that were investigated.

Why vocal, i.e., prosodic characteristics pattern consistently across emotional states stems from the speaker's physiological conditions at the time that emotion is felt. Studies have shown (Ekman, O'Sullivan, Friesen, & Scherer, 1991; Scherer, 1989, in particular) that emotional response triggers changes in breathing, phonation, and articulation due to muscular constriction and that changes in these parameters have clear effects on the speech signal.[7]

Emotion is only one factor contributing to the prosodic content of an utterance. Speakers' prosody also varies due to discourse structure. The type of information a person conveys, for example, is accompanied by a set of typical prosodic features. It has long been held that new topics are introduced with higher pitch, and wider range which narrows as the speaker comes to a close (Crystal, 1969; Lehiste, 1970, 1975, 1976, 1979, 1982; Swerts & Geluykens, 1994; van Donzel, 1999). Rate of speech too, signals different discourse elements. Peripheral, or parenthetical information is produced at a significantly slower speech rate than pertinent information (Grosz & Hirschberg, 1992; Hirschberg & Grosz, 1992). And pauses have long been shown to act as acoustic punctuation marks, acting as boundaries between one coherent block of information and the next (Chafe, 1989; Goldman-Eisler, 1968; Lounsbury, 1954). These features are further discussed as necessary in later chapters. The point thus far has been to

---

[7] That emotion affects the speech signal is not unique to humans. Primates vocalize differently as well according to various affective states (Scherer & Kappas, 1988).

illustrate that prosody contributes to the spoken message, and that these features are encoded into all utterances that we make.

1.4   Research goals and the contribution of this dissertation

The previous subsections highlighted studies from three divergent areas of research. In this section, I identify the aims of this dissertation, followed by a discussion of where this work fits within the larger body of research on police interviews. This dissertation is a compilation of studies focussing on the prosody and discourse of 16 criminal suspects during investigative interviews with police. Eight suspects are first time offenders (FT), and eight are repeat offenders (RO).

The primary aim of this dissertation is to complement existing research on police interviews by bringing together prosodic characteristics of suspects' speech with commonly occurring discourse patterns, specifically, a limited set of responses that are frequently produced by suspects during questioning by police. I will show that analysis of vocal properties is best achieved by taking into consideration the content of the message.

The secondary aim of this dissertation is to determine whether suspects who have been through the process of being arrested and interviewed by police show measurably different linguistic behaviour than suspects who are being interviewed for the first time and who have otherwise had no significant dealing

with police. Applications for findings presented in subsequent chapters will be considered.

In reaching these goals, a number of research questions are posed, organized by prosodic feature.

1) Do suspect group and response type have an effect on

a) the latency (i.e., delay) between the investigators' prompts and suspects' responses?

b) the pause-to-speech ratio of responses?

c) the frequency of pauses in responses?

d) the duration of pauses in responses?

2) Do suspect group and response type have an effect on

a) the speech rate of responses?

b) the articulation rate of responses?

3) Do suspect group and response type have an effect on

a) the average pitch of responses?

b) the pitch variability of responses?

c) the pitch range of responses?

The research questions give a general outline of structure of the dissertation. More importantly though, the research questions show how this

dissertation will approach police interviews in terms of the types of responses suspects give to complement discourse and conversation analytic approaches, as well as make use of a subset of the numerous variables supplied by researchers of deception.

As a unique type of institutional discourse, police interviews have not been the focus of many linguistic studies, and of those, none address a suite of prosodic characteristics. The contribution of this dissertation, therefore, is a descriptive, empirical treatment of the authentic speech of suspects under the condition of genuine police interviews.

1.5 Outline of the dissertation

This introductory chapter closes with an overview of the dissertation. Chapter 2 introduces the corpus upon which the research is based. I offer a few words on how police interviews are conducted by the agencies who have graciously provided the data. I introduce the types of responses that suspects routinely produce in response to police questioning, as well as the suspects themselves whose speech forms the basis for this study. In Chapter 2, I also introduce the prosodic features that comprise the dependent variables for the study.

The following three chapters address the individual studies carried out on suspects' prosodic behaviour. Chapter 3 deals with four pausal phenomena. First, I examine the degree of latency between the investigator's question or

assertion and the suspect's response. Second, the ratio of pause to speech is explored for each response type. Third, pause frequency, and fourth, pause duration are both studied for each response type. Chapter 4 examines the temporal aspects of suspects' responses using two measures. Speech rate, as a measure of syllables per second for the duration of whole utterances, and articulation rate, as a measure of segments per second, minus pauses, are examined. Chapter 5 addresses the pitch values of suspects' responses with respect to average, variance, and range. Chapters 3, 4, and 5 are each constructed in such a way as to provide an overview of previous prosodic studies followed by discussions of the relevant methodological considerations and rationales for each study. Results are presented, followed by a summary of variables in the form of profiles for each response type, comparing the two groups of suspects. Chapter 6 concludes the dissertation by correlating the results of the three previous chapters to produce a set of prosodic profiles for each response type. A note on the forensic application of the findings is made, and proposed research directions are given.

2 THE CORPUS AND THE METHODOLOGY

It is a capital mistake to theorize before one has data.
Insensibly, one begins to twist facts to suit theories,
instead of theories to suit facts.

--Sir Arthur Conan Doyle

This chapter serves a number of purposes. First, the corpus used to

examine the prosody of suspects' speech is introduced in terms of its origins and

suitability for this study. Second, police interviews are shown to consist of three

basic parts, an opening, a question phase and a closing. Third, I introduce the

suspects whose responses are analysed in this study. Fourth, the categories of

responses that suspects produce during police interviews are defined, and

examples are given. Finally, I discuss prosodic features that will be the focus of

Chapters Three through Five of the dissertation, and the measurements

appropriate to their description. The overall design of the study concludes the

chapter.

2.1   The corpus

In Canada, in recent years, police interviews are at least audio recorded,

and most often video recorded as a matter of routine. Interviews are conducted

and kept at police headquarters for use during the investigation and for use as

evidence in court if a case goes to trial. They are archived at police headquarters for future reference, should a case be re-examined at a later date.

For this study, I relied on the cooperation of the Vancouver Police Department and the New Westminster Police Service to amass a corpus of 31 interviews with suspects who were held for questioning for their involvement in one or more crimes. The interviews were loaned to me in DVD, VHS, and audiocassette formats. Each of these was digitized to an audio CD and then each saved at 44,100Hz, in .wav format, in 30 minute segments. Dealing with interviews in 30 minute blocks made data management and extraction easier than dealing with several hours at once.

In order to conduct the present study, 16 interviews out of the 31 in my larger corpus interviews were selected on the basis of four criteria:

1. All suspects are native speakers of Western Canadian English to control for any effect on prosody that non-native accent might have.

2. All suspects are male. Female suspects were excluded at this stage of the research for two reasons. First, female suspects tend to cry a great deal more during interviews which makes phonetic analysis awkward. Some male suspects occasionally cry, in which case, I exclude that portion of the interview. Second, any differences that typically occur between male and female prosody is beyond the scope of the present study. There are four interviews with women in the larger corpus.

3. While the suspects in the larger corpus were questioned for a variety of crimes, only those who were questioned for violent or disturbing crimes were selected. This decision was motivated by the need to maintain some level of control over the degree of penalty faced by the suspects. Those who break into cars for example, face less punitive consequence generally than those who commit sexual assault or attempted murder. It is assumed that an effect of stress could potentially influence the amount of excitability in suspects' voices.

4. Interviews with the clearest speech were chosen for phonetic analysis. While all 31 interviews in the larger corpus were intelligible, not all recordings were of adequate quality because of either noise interference — two interviews had a noisy hiss throughout — or in five interviews, the suspect spoke too quietly for the software (identified below) to perform phonetic analyses.

Once the interviews were selected on the basis of these criteria, each interview was transcribed, and the responses produced by each of the 16 suspects were extracted from the 30-minute segments as smaller digital sound files, and sorted according to the type of suspect (see 2.3) and the type of response (2.4). Analyses on prosodic features (see 2.7.1) were performed using the phonetic software package Praat, version 4.4.20 (Boersma & Weeknink, 2006).

2.2   Structure of the police interview

The purpose of this section is to describe how interviews with suspects typically unfold in this corpus. Both cooperating agencies conduct their

interviews at police headquarters. Interviews take place in small rooms,

approximately three metres by three metres, furnished with one small table and

two chairs. There are no windows in an interview room, and it is lit from above.

For interviews that are video recorded, a video camera is mounted on the wall,

beside or above the door, which is opposite the table. Figure 2.1 shows the

ordinary configuration of an interview room.

Figure 2.1   Typical configuration of the interview room.
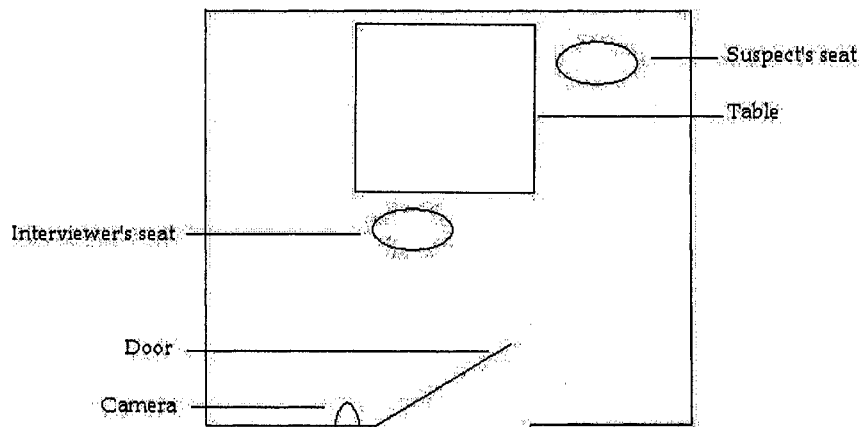
Interviews take place in their entirety in the interview room. While they

can vary greatly in length, they ordinarily consist of three phases, each of which

are described below.

2.2.1   Phase one: the opening

Before the suspect is brought into the interview room, the investigating

officer usually enters the room alone to record the date and time of the interview,

the name of the suspect, the charges for which he or she is being questioned, and

the file number associated with the case. This step takes less than one minute to complete, after which time the investigating officer or a colleague leads the suspect into the interview room. If the suspect is handcuffed, the cuffs are removed before the interview begins. If the suspect is in prison clothes, he is often offered something more comfortable to change into. The suspect is then invited to take a seat in the chair which most faces the camera (see Figure 2.1).

The dialog in the opening stage does not normally address the charges right away. Rather the investigator inquires of the suspect's physical wellbeing, whether he is under the influence of any alcohol or narcotics at the time, and whether he has eaten since his detainment. The suspect is offered food and drink. The investigator often then shifts the conversation to the suspect's occupation, family life, level of education and a number of other demographic topics. Occasionally, small talk is made regarding current events, sports news, recent movies, local politics, and various other matters. The suspect is asked as well about his prior dealings with police, e.g., whether he or she has ever been arrested before, and whether charges have ever been brought against him or her. Often when the suspect is already known to police, background information is not gathered and small talk is kept to a minimum. This information-gathering, rapport-building stage ranges in my corpus from a very few minutes to approximately one hour.

2.2.2 Phase two: substantive questioning

The transition from the first stage to the second is often marked by the investigator asking the suspect if he knows why he is speaking with police. The suspect may state why he believes he has been detained, or the suspect may claim to have no knowledge as to investigators' interest in him. At this time, the investigator informs the suspect of the charge or charges and then he or she informs the suspect of the right to legal representation and that he or she is not required by law to provide any information to police. As a matter of routine procedure, the Charter Cautions are recited to the suspect as follows with occasional slight variation in wording:

> "It is my duty to inform you that you have the right to retain and instruct counsel without delay. You may call any lawyer you want. A legal aid duty lawyer is available to provide legal advice to you without charge and explain the legal aid plan to you. Do you understand? You are not obliged to say anything unless you wish to do so, but anything you do say may be given in evidence."

The suspect is asked whether he understands the caution. While most suspects report understanding their rights, investigating officers frequently reiterate the caution in plainer language.

After the caution is recited, the investigator moves to the substantive part of the interview during which the questioning is directly related to the investigation. In my corpus, this phase ranges from 18 minutes to 3.5 hours. The pattern of turn taking is driven by the investigator, placing the suspect almost solely in a position of responding. During this phase, a simple pattern quickly

emerges in which the investigator asks questions or makes assertions with the expectation that the suspect will respond with a confirmation or further relevant information. To both questions and assertions, the suspect ordinarily provides some type of response. The types of responses will be described in section 2.4.

This stage of the interview may be interrupted at any time should the investigator leave the room to answer a telephone call or a knock on the door, or should the suspect request a washroom break. Often during lengthy interviews, the suspect is offered another drink and something to eat. Phase Two nears its conclusion when the investigator asks a very open ended question such as, "Is there anything you feel we have not discussed here?" or, "Is there anything else you think I need to know?"

Phase Two is the portion of the interview with which we are most concerned in this study. Questions asked during Phase One are not examined because they do not address the charge for which the suspect is of interest. Furthermore, the questions asked during phase two evoke more emotional response than seemingly benign questions about a suspect's level of education or employment.

2.2.3 Phase three: the closing

The third phase of an interview is its closing. The investigator informs the suspect of the next steps in the process which ordinarily means that he will be taken back to his cell and that he can expect to be seen by a Justice of the Peace

within a day. The interview is then concluded when the investigator explicitly indicates so, and he or she states the current time for the record. In one interview in my corpus, the suspect is shown being finger-printed, and blood samples are taken for DNA evidence. His body is photographed for evidence of injury. Under most circumstances, the closing phase takes approximately a minute or two.

2.3   The suspects

In this section, the suspects are introduced. First, two groups are identified: first-timers and repeat offenders. Second, basic demographic information is provided for each suspect, as well as the charge or charges for which the suspect is questioned.

2.3.1   First-time suspects: type FT

The suspects in this group report no significant dealing with police prior to the charge for which they are being questioned. Some may have had minor dealings with police as youths, but none report prior convictions. It is observed, although not without exception, that first time suspects appear anxious and apprehensive. Other outward characteristics include nervousness, fear, remorsefulness, apprehension, hesitance, non-confrontational demeanour, or some combination of these traits. Type FT suspects can sometimes be overly cooperative with police, and offer much more information than is asked for.

2.3.2 Repeat offenders: type RO

The suspects in this group have had previous dealings with police, as well as prior convictions for one or more crimes. This type of suspect shows rather bold, confident behaviour. He appears self-assured, impatient with the process, often uncooperative, confrontational, challenging of authority, and sometimes bored or annoyed with the process. Some suspects of this type occasionally seem proud of the crime of which they are accused. As with Type FT suspects, there are exceptions.

2.3.3 Homogeneity of groups?

By no means would it be reasonable to claim that all first time offenders behave in the "Type FT" manner, nor would all repeat offenders consistently behave in the "Type RO" manner. However, this pattern emerged quickly in the early stages of the study, and therefore made it convenient in terms of identifying these two types as independent variables.

It would be unwise to discount the probability of other suspect types on the basis of psychological profiles employed by law enforcement or forensic psychology, even within the small set of suspects herein. However that avenue remains for future research.

2.3.4 Suspects' details

In the tables below, the 16 suspects in this study are divided into the appropriate groups, and information about their ages and occupations is provided, along with the event for which they are of interest to police.

Table 2.1   First-time (FT) suspects' ages, occupations, and charges

Suspect	Age[8]	Occupation	Questioned for
FT-1	50s	unemployed	sexual assault
FT-2	65	retired	sexual assault of a minor
FT-3	20s	labourer	assault causing bodily harm (attempted murder)
FT-4	26	fast-food restaurant manager	stalking, sexual assault
FT-5	24	unemployed	child physical abuse
FT-6	29	unemployed	sexual assault of a minor
FT-7	34	courier	child physical abuse
FT-8	31	youth group coordinator	sexual assault of a minor

[8] Where possible the suspect's age at the time of the interview is recorded. When unspecified in the interview, I have estimated. Unknown ages are estimated to the decade. It is necessary to make more estimates among repeat offenders because repeat offenders' demographics are likely already known and therefore investigators do not inquire during the interview.

Table 2.2   Repeat offender (RO) suspects' ages, occupations, and charges

Suspect	Age	Occupation	Questioned for
RO-1	40s	scrap metal collector	armed robberies
RO-2	60s	unemployed	armed robberies
RO-3	40s	moving company operator	extortion, unlawful confinement, assault causing bodily harm
RO-4	26	labourer	serial sexual assault
RO-5	57	delivery truck driver	sexual interference with minors
RO-6	40s	unemployed	assault causing bodily harm (attempted murder)
RO-7	20s	unemployed	forcible confinement
RO-8	50s	unemployed	sexual assault

The demographics of suspects in the larger corpus are similar to those listed with respect to a wide range of ages, and their holding unskilled jobs or being unemployed. Only one suspect in the larger corpus held a highly-skilled, well-paid job as a family doctor.

2.4   Response types

In this section I describe the types of responses that suspects give when prompted by police. Because the conversation is led and topics chosen largely by the investigating officer, the suspects' turns are generally limited to responses to

investigators' questions and assertions. Many responses fall into easily identified types, depending on the type of information they contain. Those response types to be analysed for prosodic features in this study are *Affirmative, Negative, Relevant, Irrelevant,* and *Confession.* Each is described in Sections 2.4.1 through 2.4.5. Other response types that occur in the corpus but which will not be analysed prosodically are defined in section 2.4.6.

2.4.1 Affirmative

Similar to 'accept details' as defined by Hill (2003), affirmative is a response in which a suspect answers affirmatively to the information in the investigator's prompt. Often, the prompt is either a yes/no question, or an assertion, wherein the details are provided by the investigator. It does not include the offer of any new information by the suspect, so typically, this type of response is quite brief.

> I: Did you touch her?
> S: Yes I did.

> I: So after that, you were in New Westminster.
> S: Yeah.

2.4.2 Negative

Much like Affirmative, Negative responses tend to be fairly short and do not offer any new information. Again, they are most likely prompted by a

yes/no question, but an assertion can also be followed by a Negative. This response type, is similar to 'decline details' defined by Hill (2003).

> I: Were you trying to kill her?
> S: No.

> I: And she drew blood where she bit you.
> S: No, she didn't.

2.4.3 Relevant

Relevant responses provide information that directly addresses the information contained in the investigator's prompt. This response type ordinarily follows an assertion or an open-ended, WH question and it shows the most variation in terms of the amount of information offered. *Relevant* responses can vary greatly in length, anywhere from a short proposition containing one or few details, to several details. In this study, I chose to limit the category Relevant responses to those not longer than a few clauses, provided they do not consist of elements typically attributed to narratives (see section 2.4.6), which tend to uniquely have a range of internal prosodic structure.

> I: Other drugs, whatever, what other drugs you were on?
> S: I was on valiums and rivitols.

In the example above, the investigator asks a WH question, to which the suspect provides a concise and relevant response. In the example below, the investigator makes an assertion that the suspect has been identified as the individual who injured someone.

> I: That's sort of this information we have received at this time in regards to to that, you know, they put you there. They put you being the person that has injured this individual and has placed this individual into the hospital.
>
> S: That's all true, but, he tried to kill me three fucking times. And the third time he had an object in his hand as I was leaving the house. He confronted me in the front yard and he went for my throat.

Responses that begin with a 'yes' or 'no', and that also contain some relevant details are categorized as Relevant rather than Affirmative or Negative because of the additional offer of relevant information. An example follows.

> I: So is he coming from the house? *[present progressive to guide past narrative]*
>
> S: No, he's outside waiting for me. He's on the sidewalk.

2.4.4 Irrelevant

Suspects may offer information in response to a question or assertion that does not directly address the details contained in the prompt, in which case that response is categorized as Irrelevant.

I: If you could, ok, how many paces from the couch to
the wall?

S: Well, like, cause we've moved around our, our
living room since then like, two or three times
because we're, just she just moved there right? So
we've been changing things around.

Often times, the suspect may start his or her response on the appropriate

topic (Relevant), and move off topic, offering Irrelevant information.

I: So what were you doing just before you sat down?

S: I went to grab him to feed him.

I: From who? Who'd you take him from?

S: Relevant → From the, I think I took him from his
car seat. Like we usually take him from the crib to his
car seat because we don't have a swing, right?
Irrelevant → Because we just moved like we do have
a swing but it's at the mother's place. Or at the
grandmother's.

Producing Irrelevant responses appears to be a strategy employed largely

by FT suspects. RO suspects produced Irrelevant responses only rarely.

Therefore there were too few to be analysed in any meaningful manner.

2.4.5 Confession

When confronted with the accusation for which the suspect is of interest,

he or she may confess to having committed the crime. A confession may come

quite early in an interview, or after a considerable amount of questioning. Not all

interviews conclude with a confession.

I: Well then tell me what went on there. Tell me about the straight goods. Ok? What happened?

S: I did it.

I: You did that one?

S: I did it.

I: Did you do it? Because if you did, let's talk about it. Did you touch VICTIM sexually? Did you touch her vagina?

S: [long exhale] Yeah, I must have.

I: Tell me about it.

S: She comes over with my brother and she stays with us. I had a couple of beers. And I just went to go check on her, and then uh, I went and uh and laid on the bed and she was saying uh, 'don't do this'.

Confessions were not as numerous as other response types, and in fact only 6 of the 16 suspects in this study produced them. Confessions however, constitute pertinent evidence in an interview and therefore are included in this study. It should also be noted that where some confessions may look like affirmative responses (I: "Did you break his leg?" S:"Yes, I did"), they are categorized as confessions when they directly respond to a criminal charge, as opposed to affirming a less significant detail, as shown in section 2.4.1.

2.4.6   Other response types

The response types above are the most frequently observed types in the corpus and are produced by the largest number of suspects with the noted exception of confessions. They are not, however, the only response types to be

found. Others occur with less frequency, and are reported in this section, or they occur often, but for other reasons are unsuitable for study at this time. Beyond defining them below, they will not be addressed further in the present work.

Concession

When confronted with an accusation, the suspect may concede to having committed some questionable act, or a violation of the law, but not the one to which the investigator is referring. Occasionally, it is a crime that may be perceived as less severe than the one of which the suspect is accused. In the case of sexual assaults, the suspect may initially deny sexual contact with a victim and later reluctantly admit to having had sex with the victim, but insisting on it having been consensual. This type of response is not frequent in my corpus, and not all suspects produce it.

> I: So we both know what happened. Really, about how you obtained the car.
>
> S: The only thing that gets me, is is that, is that it says in there, in there, *[pointing at report]* sexual assault. I beat her up, I didn't rape her.

The suspect is willing to concede that he physically assaulted the victim, but at this stage in the interview, he does not own up to a sexual assault. In another example, the suspect who had steadfastly denied all sexual contact with the victim is confronted with the information that his DNA was recovered from the victim's genitals, at which point he concedes to their having had sex.

I: Can you tell me why your DNA would be on her private parts?

S: Like, we did have sex *[after earlier insisting there had been no sexual contact.]*

Plead innocent

In this type of response, the suspect directly declares his or her innocence, or he or she pleads with the investigator to believe his or her version of events.

I: So now, how do I find out the truth? Because that's not where it happened. That's terrible and that's very upsetting, but that's not where it happened. So where did it happen?

S: Ok, you've been through a thousand cases. Or thereabouts. Can you not look at me, can you not look at me and tell, seriously I think you can tell.

I: But let me explain something else SUSPECT , that, this this incident, let's say, let's just say hypothetically now, let's say you did this incident to your son, you did something to your son, that doesn't make you a bad person.

S: Yeah but I don't I didn't do it. I'm innocent.

Refusal to answer

Some suspects refuse to respond to investigators' prompts. They may take advantage of their right to remain silent, or they may become uncomfortable with the direction the questioning takes. Sometimes, a suspect who had started out cooperative and talkative decides to discontinue the interview.

I: What have you got to say?

S: I exercise my rights under the charter. I got nothing to say.

I: Ok. Now, you own that one for sure [referring to a recent confession of sexual assault]. Let's deal with them all, let's make sure everything's out on the table here.

S: That's all I'm dealing with.

I: I think you want to take this opportunity to tell your side of the story.

S: I want to go back to my cell.

Remaining silent is also a type of refusal to answer. Considering that remaining silent is an option, it is interesting how few suspects seem to exercise it. In the larger corpus of 31 interviews, only one suspect completely declined to answer questions, two suspects declined to answer certain questions, and one who had been answering questions stopped, laid down on the floor and passed wind repeatedly instead of answering questions.

Lack of knowledge or recollection

In response to any kind of question or assertion, a suspect may claim to have no knowledge of a detail or event, or that he or she cannot recall.

I: You touched her and then you had sex with her. It just happened. Didn't it?

S: I don't know.

> I: Now please tell me that you went down to check on
> her, and hey, you just got aroused, and it just
> happened. You didn't plan it.
>
> S: I can't remember anything.

It may be that the suspect genuinely does not know the answer to a

prompt, or has genuinely forgotten. It may also be the case that the suspect does

in fact recall, or does in fact know the answer, but states otherwise. It would be

difficult or impossible to distinguish genuine lack of knowledge from feigned

lack of knowledge given the scope of this study, although at a later date,

determining the prosodic or contextual cues would be an interesting course of

study.

Self-aggrandizement

When faced with an accusation, the accused may extol his or her own

virtues or otherwise point out positive attributes. An example follows:

> S: Alright, if I did it, I fucking did it. I'll tell you I did
> it.
>
> I: And not a lot of people would do that.
>
> S: Alright? And and and if you look at my file and
> and you look at my court cases, I've never taken a
> fucking charge to trial. You know my last sexual
> assault conviction in 94, I could have walked on that!
> I could have beat that fucking beef if I took it to trial. I
> had it beat. They had no DNA they had no fucking
> nothing. Sure I robbed the store, I was there to rob the
> store, but they couldn't prove the sexual assault, they
> had no fucking proof of it

> I: But there was a sexual assault.
>
> S: But, because I'm a man, I own that. I said, yeah, I did that.
>
> I: Fair enough.

In the example above, the suspect highlights his morality by disclosing his readiness to accept responsibility for crimes he has committed. Not all are of such an extreme nature, however, as the following example shows.

> I: When uh we met, I told you I was gonna be talking to, what the hell's his name. Witness. Witness. Ok. I told you he's in jail.
>
> S: Alright.
>
> I: And he is.
>
> S: It has nothing to do with me. Hey, I don't fuckin'- My problem is I my wife says it all the time, you're always helping these dead beats. You know, taking people into the fuckin' house, giving money away. I'm just, I'm bad at that.

The extract above is an example of a suspect, anecdotally and through the reported speech of his wife, attempting to explain how his generosity puts him in contact with criminals, presumably implying that he would not find himself in legal trouble otherwise.

Rhetorical questions

Some suspects, not all, ask a rhetorical question in response to a question or assertion.

> I: Now SUSPECT, we need to know what happened to
> VICTIM.
>
> S: Why would I hurt my own child?

Evaluation

Some suspects, usually Type RO suspects, produce evaluative comments

regarding the investigation or the interview process itself. Frequently and

unsurprisingly, their evaluative comments are often negative of others or are

intended to register the suspect's disbelief or disappointment with his situation.[9]

> I: Any other robbery of drug dealers down in CITY?
>
> S: They're fucking goofs anyway man. I'm doing you
> guys a fucking favour by robbing those pieces of shit.

> I: We're here to talk about your involvement, not his.
>
> S: This is so fucked! You're fucked!

> I: I'd like to ask you a bit about that.
>
> S: I can't believe I'm sitting here. He's there and I'm
> here.

Narratives

In all 16 of the interviews in the study, suspects of both types were apt to

embark on a narrative recounting their version of events, which were often,

although not always, in relation to the crime or crimes for which they were being

---

[9] See Fadden (2007) for discussion of suspects' evaluative comments in police interviews.

interviewed. Like narratives occurring in casual or other non-confrontational speech, narratives in police interviews range from several seconds to several minutes, displaying all the features that casual narratives display, meaning a sequence of events containing some or all of the following Labovian narrative elements: abstract, orientation, complicating actions, evaluation, resolution, and coda (Labov, 1972) An example follows with the narrative in italics, and narrative features marked within the text:

I: So let's let's go right back to that to the beginning of that case. You're at WITNESS 's place, right? You walked out of WITNESS's place, there was a woman there, taking stuff out of her car. Right?

S: No

I: Ok, then tell me what happened.

S: She wasn't taking anything out of her car, she was standing BY her car. *(orientation →)* I thought Fuck I need a car, I gotta go. *(orientation →)* We're all fuckin stoned and shit, WITNESS runnin around doin dope all night and stuff, I don't know if he told you that part.

I: yep, oh yeah.

S: *(orientation →)* And even while we were doin dope I was sittin there thinking, fuck I gotta go, I gotta go, I got, I got things to do right? *(complicating action →)* And uh, so then I left right? *(evaluation →)* Opportunity right? There's opportunity right?*(complicating action →)* So I jumped on her. *(complicating action →)* And she fuckin fought. Fuckin punchin me, kickin me, everything and, and for like two minutes,*(evaluation →)* I don't know how this guy *[an eyewitness]* only said a fuckin minute, *(resolution →)* cause I fuck, we ended up outta the car.

While fascinating and worthy of study themselves, narratives were not examined as a response type in this study owing to the fact that narratives are shown to exhibit an internal prosodic structure from beginning to end (Oliveira, 2000; Schleef, 2003).

2.5 Limitations of the corpus study

The previous subsections present a relatively thorough sampling of the types of responses that suspects produce. However, it cannot be claimed that these response types are exhaustive of all that can occur during a police interview because of suspects' individual differences, investigators' individual methods of questioning, and unique situations call for unique exchanges of information. As an example of institutional discourse, though, in which the institutional goal is to elicit relevant information pertaining to a crime, the examples above would capture the most common response types, produced by the majority of suspects.

It is important to acknowledge the one-sidedness of this study in that only suspects' responses are analysed. That investigators' questions and assertions are not categorized and analysed is the result of time and space constraints. A fuller treatment of the prosody of prompt-response exchanges would be well informed by the previous work of Stenström (1984) who provides a robust account of questioning and responding in English, as well as those works in which such exchanges have been examined specifically within the context of police

interviews (Heydon, 2004, 2005; Hill, 2003; Johnson, 2006; Shuy, 1997). I do not discount the strong likelihood that in some cases the prosodic features of the suspects' responses would mirror the immediately preceding prompt from the investigator, however an analysis of that aspect of police interviews remains for future study.

2.6 Category reliability

In order to test the validity of the proposed types listed in Section 2.4, a reliability analysis was carried out. Two professors of discourse analysis in the Department of Linguistics at Simon Fraser University were asked to categorize 100 randomly ordered responses. Response categorization training consisted of a labelled set of exchanges including pairs of prompts and responses. Multiple examples of each response type were given and both participants reported the labels to be clearly defined an exemplified.

The reliability test itself was carried out on 100 prompt-response pairs, and raters were presented with a forced-choice task in which they were required to categorise the responses. Raters' categorizations were taken together with my own and a Kappa coefficient was obtained on the basis of three raters' judgements. A kappa value of .81 was achieved and was deemed acceptable, given the number of categories.[10]

---

[10] In the reliability study, raters labelled examples from each of the categories in Section 2.4.6. However, only those response types that would be examined in the study were calculated for the Kappa value of .81.

Cases of disagreement most often involved categories of responses not included in the study at this time. In particular, there had initially been a category called "*Excuse*" which was often given in response to a "why" type of question posed by police. These tended to be labelled as Relevant by raters, prompting the blending of Relevant and *Excuse* under the category label Relevant.

As the kappa value was sufficiently high to signal strong category reliability, I proceeded with categorizing the responses according to the initial definitions, with the one exception noted for *Excuse*.

2.7   Design of the study

In this section, I introduce the prosodic features in question and how they are to be measured and then provide an overview of the study as a whole. The following three chapters, each devoted to an individual prosodic feature will address more specific information and rationale about the appropriate methods for pause, tempo and pitch.

2.7.1   Prosodic features

It is uncontroversial to define prosody as the collection of suprasegmental phenomena which includes pitch, duration, and intensity (Lehiste, 1970). In this study, pitch, pause, and tempo are analysed and each is discussed below with respect to different types of measurements. Unfortunately, it is not possible to

examine intensity in this study because of the lack of control over suspects' proximity to recording equipment. In each of the subsections below, a few words on the types of measures to be taken are provided.

Pause

A set of four pausal features will be examined in this portion of the study. First, turn latency, or between-turn pause, is examined for many of the response types, reported the nearest 0.01s. Also reported is the pause-to-speech ratio, expressed as the sum of all pauses, divided by the total duration of the utterance, excluding latency. Within-turn pause frequency, measured as the number of pauses per second, and within-turn pause duration are also reported.

Tempo

The rate at which suspects speak is measured in two ways. First, speech rate is analysed, as a measure of syllables per second, taken over the duration of the response in its entirety, including pauses. Second, articulation rate is also explored, as a measure of segments per second, excluding pauses for the entire response.

Pitch

Pitch variables to be examined for all response types for both groups of suspects are average pitch, pitch variability, and pitch range. These phenomena are determined by measuring the minimum, maximum, and average pitch values

of each intonation phrase in each response. Intonation phrases are judged according to the criteria in Cruttenden (1997). In order to better capture the differences within and between speakers across the response types in a manner that closely matches the human capacity to hear differences in pitch, the semitone scale will be used (Henton, 1989), rather than the linear Hz scale.

2.7.2 Response distribution and analysis matrix

Previous sections of this chapter have introduced the nature of police interviews and the suspects whose speech will be examined, the categories of responses that suspects produce, and the prosodic features to be explored. This section provides schematic information about the study, with respect to a) the number of responses produced by each suspect, for each type (Table 2.3 and Table 2.4); and b) the organization of response categories and the use of the appropriate prosodic measurement to summarize the design of the study in its entirety (Table 2.6)

Table 2.3   Response distribution of FT suspects

Suspect	Affirmative	Negative	Relevant	Irrelevant	Confession	Total
FT-1	27	21	24	19		91
FT-2	9	16	20	12	5	62
FT-3	22	6	34			62
FT-4	22	23	22	20	4	91
FT-5	31	19	26	16		92
FT-6	11	9	12	20	7	59
FT-7	10	19	18	16		63
FT-8	18	20	18	15		71
Total	**150**	**133**	**174**	**118**	**16**	**591**
%	**25.4**	**22.5**	**29.4**	**20.0**	**2.7**	**100.0**

The header row above spans under "Response Type".

Table 2.4   Response distribution of RO suspects

Suspect	Affirmative	Negative	Relevant	Irrelevant	Confession	Total
RO-1	12	21	28			61
RO-2	18	14	14		2	48
RO-3	10	9	23			42
RO-4	41	18	18		5	82
RO-5	23	12	21		2	58
RO-6	11	11	7			29
RO-7	12	11	26			49
RO-8	19	10	14			43
Total	**146**	**106**	**151**	**0**	**9**	**412**
%	**35.4**	**25.7**	**36.7**	**0.0**	**2.2**	**100.0**

The header row above spans under "Response Type".

Table 2.5   Total responses by response type including both groups

	Affirmative	Negative	Relevant	Irrelevant	Confession	Total
Total FT+RO	296	239	325	118	25	1003
%	**29.5**	**23.8**	**32.4**	**11.8**	**2.5**	**100.0**

Responses vary in length. The shortest response is 0.19s, a brief "yeah",

and the longest is 23.61s, where a suspect offers relevant information to an

investigator's question. The average response length is 5.72s. Figure 2.2 illustrates

the distribution of responses by duration, including within-turn pauses.

Figure 2.2   Distribution of responses by duration
              Values are given in seconds.

□ 0 to 2.99
⊠ 3 to 5.99
■ 6 to 10.99
■ 11 to 24
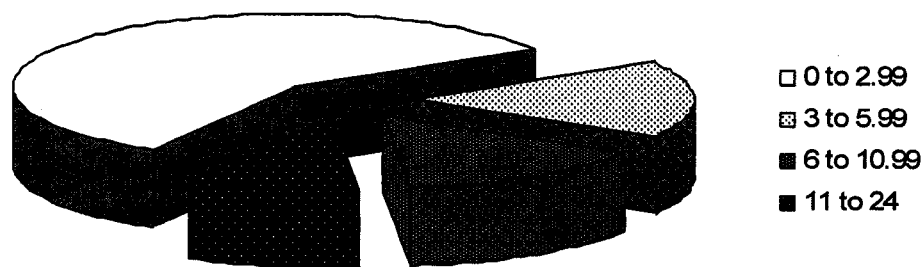
More than half of the responses are quite short, owing to the fact that

approximately half of all responses included in the study belong to the

affirmative or negative categories, where the suspect responds with a brief 'yes'

or 'no' style of answer. The large proportion of affirmative and negative

responses likely indicates that police ask more yes/no questions than any other

type. However, police discourse and question type is beyond the scope of this

study.[11]

---

[11] See Hill (2003) for discussion of police interviewer question types.

Table 2.6 Analysis matrix

The prosodic features in the leftmost column are measured for each of the response types shown across the top, where an x occurs in the inner cells. Gaps in the table are the result of linguistic and methodological consideration, and are explained in subsequent data chapters.

Prosodic Feature	Suspect Type→	Affirmative		Negative		Response Type Relevant		Irrelevant		Confession	
		FT	RO	FT	RO	FT	RO	FT	RO	FT	RO
Turn Latency		x	x	x	x	x	x		x	x	x
Pause Ratio						x	x	x		x	x
Pause Frequency						x	x	x		x	x
Pause Duration						x	x	x		x	x
Speech Rate						x	x	x		x	x
Articulation Rate		x	x	x	x	x	x	x		x	x
Average Pitch		x	x	x	x	x	x	x		x	x
Pitch Variability		x	x	x	x	x	x	x		x	x
Pitch Range		x	x	x	x	x	x	x		x	x

In this chapter, the corpus has been introduced, along with an overview of how police interviews are conducted, the types of responses suspects produce, and demographic information about the suspects in the corpus. With the prosodic variables defined and matched with the response types for which they are appropriate measures, the next three chapters treat each pause, tempo, and pitch as individual studies with respective literature reviews, methods, results and discussions.

3 PAUSE

Never miss a good chance to shut up.
–Will Rogers

In this chapter, suspects' pause behaviour during police interviews is examined. Four pause phenomena are explored. First, I look at the latency between the investigators' prompts and the suspects' responses. Results show that overall, suspects' latency values vary as a function of both response type and suspect group. Second, the ratio of pause-to-speech is examined in relevant responses to find differences between the groups. Third, the frequency of pauses within responses is also explored and I conclude that pause frequency varies as a function of group. Fourth, I examine the duration of pauses to find that suspects' pause duration varies according to group.

3.1 Introduction

In her widely cited work on temporal phenomena in spontaneous speech, Goldman-Eisler (1968) gave the following characterization of the flow of speech which is naturally disrupted by the occurrence of pauses.

> "Speech is a serial phenomenon, an activity spread
> out over time. It does not however, fill time
> continuously, particularly when it is spontaneous, but
> comes in fits and starts with intermittent periods of
> non-speech. A passage of speech extending into time

consists of two sorts of time: time of vocal action and
time of silence." (p.11)

Disruptions in speech production are attributed to an array of

physiological factors including anxiety, pain (Goldman-Eisler, 1968; O'Connell &

Kowal, 1983) and breathing (Hird & Kirsner, 2002); cognitive factors such as

lexical availability (Goldman-Eisler, 1968); emotional state (Scherer, 2003;

Scherer, Banse, & Wallbot, 2001; Sobin & Alpert, 1999); linguistic factors such as

intonation and discourse boundaries (Ford & Thompson, 1996; Oliveira, 2000;

Schleef, 2003; Swerts, 1997); mental health (Nilsonne, 1987, 1988); and individual

and social factors such as age, sex, social context (Kowal, O'Connell, & Sabin,

1975), and discourse genre involving the degree of spontaneity (Barik, 1977;

Goldman-Eisler, 1961; Osada, 2003).

A number of authors have concluded, (Kowal, O'Connell, & Sabin, 1975,

most strongly; O'Connell & Kowal, 1983), that a study which neglects to account

for "multiple psychological determinants of speech behaviour", i.e., a study

which isolates any member of the set of factors that contribute to pausal

phenomena ought to be "rejected as inadequate" (1975: p.195). While it is

assumed that multiple determinants are behind pause and other prosodic

phenomena, one need not be deterred from studying pausal phenomena without

knowing what all the causal factors are because the alignment of prosodic

features with various discourse acts in a variety of contexts is in itself worthy of examination. I return to this point below.

In this chapter, the pause behaviour of suspects as they respond during police questioning is examined. Before identifying the research questions to be addressed in this study, I offer a word on what this portion of this chapter is and is not intended to describe. The goal is to address quantitatively speakers', i.e., suspects', pause behaviour. The goal is not to address the underlying factors that cause suspects to pause during interviews. In fact, to do so would be extremely difficult without interviewing the suspects to collect self-assessment post-police interview with respect to their emotional states, their physical wellbeing, or any other factors that may contribute to varying degrees of pause frequency and duration. In the absence of this information, however, it is worthwhile to examine the differences in pause behaviour between the two groups of suspects, and the differences between their response types, to produce a robust picture of this element of prosody in the context of police interviews. To that end, the following research questions are posed:

Do response types and suspects' Type FT (first timer) or RO (repeat offender) status affect:

1) the latency between investigators' prompts and suspects' responses?

2) suspects' within-turn pause-to-speech ratio?

3) suspects' within-turn pause frequency?

4) pause duration?

Examples of the types of pausing analysed in the data are shown in the tables below using two prompt and response pairs (3.1 contains a relevant response, and 3.2, an irrelevant response). Latency, silent pause, and filled pause duration are indicated in parentheses, measured in seconds. Within-turn pauses greater than 250ms are counted.[12]

Table 3.1   Suspect FT-2, sample pauses

Suspect FT-2	
I: So she fit into your family quite well at that time. S: **(0.6)** I sure think so. **(1.94)** We all **uh (0.36)** really liked her. **(3.12)** I probably **(1.76)** fucked that up.	
Total duration of response	11.94s
Total pause time within turn	7.18s
Pause-to-speech ratio	0.60

Table 3.2   Suspect FT-4, sample pauses

Suspect FT-4	
I: What date was that that she found out? S: **Um (0.96)** What I know from **(0.41)** WITNESS is that **(0.63)** she did approach one of my friends in a club.	
Total duration of response	4.71s
Total pause time	1.04s
Pause-to-speech ratio	0.22

---

[12] The decision to set the filled and unfilled pause threshold at 250 milliseconds is discussed in section 3.3.

In Suspect FT-2's example, the latency period between the investigator's prompt and the suspect's response is 600 milliseconds. Within his turn, he pauses four times, with pause lengths varying from a relatively short 360 milliseconds, to a relatively long 3.12s. The response taking 11.94s, including 7.18s pausing, results in a pause ratio of 0.60. Suspect FT-4's turn begins with a latency period, during which he utters an 'um' of 0.96s[13]. He pauses twice, totalling 1.04s pausing, within his 4.71-second turn, producing a pause ratio of 0.22. These two examples illustrate the pausal dimensions to be discussed in the remainder of this chapter. Further examples are given as necessary in the results subsections in Section 3.4.

Following this introduction, Section 3.2 is divided into two subsections. The first provides an overview of the previous prosodic work where pause intersects with discourse and emotion, and the second surveys a handful of studies on police interviews and deceptive speech where pause behaviour has been examined. Section 3.3 outlines the methodology by which this study was conducted, guided by methodological issues in the literature. Section 3.4 presents the results of the study, addressing each of the hypotheses in separate subsections. Results are compared to findings in previous studies on pause. Section 3.5 offers a look at pause feature correlates and a discussion of the results.

---

[13] A pause is considered filled if it contains an *um* or *ah*-style vocalization. Any silent period before or after the filler is considered to be part of the same pause. Rationale for collapsing the two pause types is given in section 3.3.

3.2 Pause in the literature

At least a century's worth of work has gone into understanding where in the stream of speech speakers pause, how often and for how long, and why, among other phenomena. The literature review that follows divides into two subsections. In the first, I survey the pause literature that forays into discourse and emotion, giving due attention to the historically more monumental studies which gave rise to the contemporary work on pause. In the second subsection, an overview of the relatively recent field of research on police interviews and deceptive speech is given. While my intention is not to deal with deception in the present work, considerable headway has been made on this topic particularly with respect to how pause ought to be examined in spontaneous and laboratory speech.

3.2.1 Pause as a prosodic feature in discourse and emotion

Pike (1945) was among the earlier investigators to have listed pause among the suite of prosodic phenomena that include pitch, contour, and rhythm. He coupled final pauses with the ends of intonation contours that fall to a low pitch, and tentative pauses with level contours intended to indicate non-finality (p.31). He, and later others (Essen, 1949; Hegedüs, 1953; Lounsbury, 1954, for example) all concluded that pausing is a means of punctuating the stream of speech in such a way as to group logically cohesive units of thought.

On a more quantitative level, Goldman-Eisler (1956a) showed that significant portions of speakers' utterance time is taken up by pauses. In her study, speakers discussing selected subjects paused anywhere from 13% to 63% of the total utterance time, while impromptu discussions resulted in pause proportions of 35% to 67%.

Much of the recent work on temporal phenomena, particularly pause, builds on the work of Goldman-Eisler (1968) whose multiple experiments demonstrate that pauses are for the benefit of speakers and hearers. She makes clear the difference between pauses that occur at natural linguistic boundaries and those which occur at non-linguistic boundaries. Pauses that occur at linguistic boundaries such as ends of sentences, before conjunctions, before relative and interrogative pronouns, indirect or implied questions, before adverbial clauses, and parenthetical information are instrumental in the grouping of semantically cohesive phrases thereby assisting the hearer. Pauses that do not occur at linguistic boundaries but rather elsewhere in the utterance, result from speakers' real-time processing of upcoming speech. Among the reasons for speakers' pausing, Goldman-Eisler lists lexical choice, syntactic construction, and overall conception of the message. The effect of complexity on the degree of pausing was determined in two ways: the comparison of pauses during spontaneous versus repeated speech tasks, and the comparison of pause time in description versus interpretation tasks.

Goldman-Eisler (1972) developed the notion of semantically coherent units (see above) to demonstrate "thought units" as hierarchically organized pieces of information, set apart by pauses which tend to be longer than those found elsewhere in spontaneous speech. Looking into the locations of pauses in speech through the examination of narratives, Chafe (1989) refers to similarly conceived utterances as "idea units" and shows their regular demarcation by pauses.

It has been noted that a pause need not be a period of silence in order to be a pause. The 'filled pause', that is an *um*, *ah*, or *well* type of filler, occurring turn-initially or within a turn, has also been studied. Fox Tree (2002) examines silent and filled pauses as indicators of speakers' level of commitment to their message, giving hearers a window into speakers' comfort with a topic, degree of honesty, or difficulty with speech production. From the point of view of the speaker, "the presence or absence of an *um* has implications for how far in advance speakers are aware of their upcoming delay." (Fox Tree, 2002, p.29)

The work on pause as an emotional cue is scant compared to that which takes into account pitch.[14] However, a number of studies demonstrating the effect of emotion on prosody, including pause, have been carried out in the past three decades. In order to examine the prosodic characteristics that accompany various emotions, actors are often employed to produce utterances while

---

[14] See Chapter 3 for an overview of the literature on pitch with respect to emotion.

conveying a particular emotion, or elaborate means are constructed in effort to induce emotional states in subjects.[15]

Actor-produced or artificially induced speech can be criticised for not reflecting what would happen in genuinely spontaneous speech; however, the findings on pause and emotion from this branch of study show some consistent and interesting results. Some of these are summarized in Table 3.3.

Table 3.3   Broad summary of some findings on emotion and pause[16]

	Fear	Anger	Sadness	Joy
Duration of pauses	short[b]	short[b,d]	long[b,d]	long[b]
Number of pauses	few[b]	few[b]	many[b]	some[b]
Duration of articulation	short[b,c]; moderate[a]	short[a,b]; moderate[c]	long[a,b,c]	long[b]; short[a]; moderate[c]

[a]Thompson and Balkwill (2006); [b]Sobin and Alpert (2002); [c]Banse and Scherer (1996); [d] Luakka, Juslin, Bresin (2007)

Overall, sadness appears to have the most consistent effect among the emotional states, frequently leading to long and frequent pauses (Thompson and Balkwill, 2006; Sobin and Alpert, 2002; Banse and Scherer, 1996). Speech produced under the conditions of fear and anger would appear to be indistinguishable on the basis of pause alone, both being characterized by short

---

[15] Of this approach, Scherer (2003) notes that vocal simulations of emotions yield "much more intense, prototypical expressions" than are found under natural conditions. Methodologically, however, obtaining genuine emotion is difficult.
[16] Many other emotional states are examined, only those emotions that appear consistently across studies are included in Table 3.3. Other emotional states include boredom, anxiety, pride, disgust, interest, despair and distinctions are made between hot anger and cold anger.

and infrequent pauses (Sobin and Alpert, 2002). For all three of those emotions, the duration of articulation, that is whether segments are articulated slowly or rapidly, ranges from short to moderate. Joy, it seems, has the least consistent effect on pause behaviour (Banse & Scherer, 1996; Sobin & Alpert, 1999; Thompson & Balkwill, 2006).

In a study of natural speech that examined anger and fear in call logs of a European Stock exchange call centre, and perhaps contrary to the findings in Table 3.3, fear was found to be accompanied by more and longer pauses than anger (Devillers, Vasilescu, & Vidrascu, 2004). Stronger relative comparisons could be made if their study had included sadness and joy as well.

The findings across the studies of emotion and prosody, while rather inconsistent perhaps due to methodological variability, reveal that pause behaviour varies from one emotional state to another and that hearers are sensitive to prosodic cues, including pause, allowing them to identify the emotion that underlies the message. Many authors note, however, that pause should not be studied in isolation, but rather should be considered in conjunction with other prosodic cues to adequately characterize the emotional content of speakers' messages. This concern will be addressed in Chapter 6, where all prosodic features will be examined in correlation and reference will be made to some of the findings above as they relate to suspects' responses.

3.2.2 Pause in police interviews and deception studies

A number of recent studies have aimed to describe the discourse behaviour of suspects and investigators in police interviews, both in authentic police interviews and in lab-based studies which simulate police interviews and other situations in which subjects are motivated to deceive by various incentives. Among the variables analysed in these studies are pauses within and between turns.

In casual conversation, long pauses are dispreferred, and speakers cooperatively move to fill gaps (Sacks, Schegloff, & Jefferson, 1974). When suspects produce hesitation-filled turns during police interviews, they may do so either out of lack of knowledge, or out of fear that they may reveal incriminating information, and hence take more time to plan their turn. In a conversation analysis of police interviews, Heydon (2005) identifies investigators' tolerance for long pauses as a strategy which often succeeds in urging the suspect to fill the conversational void. Specifically, Heydon uses the participation frameworks designed by Goffman (1981) to demonstrate the roles that participants might play. These are the *principle* role ("the party who is held responsible for taking up the position to which the meaning of the utterance attests", p. 517), the *animator* role (the party who actually vocalizes the utterance), and the *author* role (the party who creates the utterance). Investigators' tolerance of long pauses prompts suspects to take up *principle, animator,* and *actor* roles, thereby ending the silence and producing information (p. 67).

Psychologists and forensic psychologists have made progress both in the lab and in the field by looking at a variety of vocal and physical gestures with the aim of uncovering constellations of behaviours that signal deceptive speech acts. The studies in which pause behaviour is addressed are surveyed below.

Anolli and Ciceri (1997) prompted 31 subjects to tell truths and then lies to an experimenter who, in one case, appeared to be unsuspecting, and in another feigned suspicion. Among such vocal features as fundamental frequency range and speaking rates, they examined the frequency and duration of pauses over 100ms. In short, it was found that the pause duration was similar in all three conditions, but that pauses were more frequent in the "unsuspecting" condition, and significantly less so when the suspect told the truth, or lied in the "suspicious" condition.

Vrij and Mann (2001) performed an intensive study of a single video-recorded police interview with a murder suspect in which two observers coded the vocal and gestural behaviours of the suspect with the aim of comparing behaviours displayed during times of confirmed truth telling and confirmed lying, both before and after confession. In their study, pauses were longer and more frequent during periods of lying both before and after confessing. In addition, *ah*-like disturbances were more frequent during truth telling before the confession, and much less frequent during lies, suggesting that the suspect was actively attempting to manage commonly-assumed indicators of uneasiness.

Extending the work of Vrij and Mann (2001), a follow up study examined the vocal and gestural behaviour of 16 suspects during police interviews (Mann, Vrij and Bull, 2002). In this study, a pause was defined as an interruption in speech of approximately 0.5s or more (p. 370). Significant differences were found between deceptive and truthful speech acts in terms of pausal behaviour. Specifically, suspects delayed longer before lying than before truth telling, and displayed more hesitations during speech while lying. They also tended to display more speech errors, but results were not significant.

In a lab based study, Vrij, Edward, and Bull (2001) measured response latency in deceptive speech and found that subjects delayed significantly longer before telling a lie than before telling a truth. In another lab-based study comparing the results of first-time and repeat interrogations, Granhag & Strömwall (2002) found a significant difference between truths and lies for frequency of silent pauses of greater than two seconds, but no significant difference in terms of filled pause frequency or latency in truths and lies.

In all studies above focussing on deception except Granhag and Strömwall (2002), the authors cite Goldman-Eisler (1968) who attributes pausing and other disfluencies to a greater amount of cognitive processing experienced by the subject or suspect when performing a complex speaking task. They conclude that increased cognitive demand is placed on speakers as a result of having to be more creative in the production of deceptive information. More and longer

pauses and filled pauses afford the speaker time to think before and during speech (Anolli & Ciceri, 1997; Mann, Vrij, & Bull, 2002; Vrij, Edward, & Bull, 2001; Vrij & Mann, 2001).

While the topic of deception is not addressed in the present work, it is reasonable to suggest that the speakers, having found themselves in a high-stakes situation comparable to that described in Vrij & Mann (2001) and Mann, Vrij, & Bull (2002), would be under considerable duress. How their pause behaviour manifests will be compared to the results found in the studies mentioned in this subsection and the one preceding.

3.3 Methodology

In this section, I discuss the methods by which pause is analysed in this study. The challenge in arriving at an appropriate method for studying pause, and other linguistic behaviours, stems from striking a balance between producing results that are easily comparable with findings from previous work, yet identifying the most suitable way to describe the data at hand. The methods and measures used in related work are discussed below.

Although many studies differentiate between silent pauses and filled pauses (Clark & Fox Tree, 2002; Fox Tree, 2002; Lalljee & Cook, 1969), this approach has not been taken in this study. Instead filled and unfilled pauses are grouped together and simply reported as either the break between participants' turns or within turns pauses. The rationale behind this decision is to keep the

focus on content-laden vocalizations rather than semantically void fillers, in order to get a clear picture of how much time suspects spend or do not spend producing information.

More specifically, in this study a pause is defined as any pause, filled or unfilled, or a combination of silent and filled pause amounting to 250ms or greater during a response. In other words, any phase during a turn (>0.25s) spent not producing content-laden speech is counted as a pause. Speech errors such as false starts are not counted as pauses; rather they are figured in as part of the speaking duration (see Chapter 5). Latency is considered to be any period of silence, filled pause, or combination of silence of *um* or *ah* that occurs between the end of the investigator's prompt and the suspect's response. In a few cases, there is no period of latency, but rather there is turn overlap, to be discussed below.

The decision to set the pause threshold at 250ms was not an arbitrary one, and throughout the literature, what constitutes a pause in speech has been set at vastly differing thresholds. Some count a pause as being any hesitation as low as 80ms (Levin, Silverman, & Ford, 1967). Pauses less than 130ms are believed to be due to articulatory necessity, and beyond 130ms, pauses are deemed psychologically relevant by Hieke, Kowal, and O'Connell (1983). Some pause studies take into account only pauses that are judged impressionistically by raters (Maclay & Osgood, 1959; Vrij, Edward, & Bull, 2001; Vrij & Mann, 2001).

Still others disregard any pauses less than two seconds (Granhag & Strömwall, 2002), and some as high as three seconds (Siegman & Pope, 1966).

The majority of studies in which pause is examined set the pause threshold in the range of 200 to 270ms. Boomer (1965) and Sobin and Alpert (1999) consider pauses in their studies at 200ms and beyond; pauses of 230ms and more have been considered as well (Kowal & O'Connell, 1987); a threshold of 250ms has been set by Goldman-Eisler (1968); and still others count pauses of 270ms (Kowal, O'Connell, & Sabin, 1975; Sabin, 1976; Sabin, Clemmer, O'Connell, & Kowal, 1979).

The pause threshold in this study is set at 250ms in part because it puts the pause threshold reasonably close to the majority of other investigations, thereby making the results of this study comparable to results in other studies. Pauses in the very low ranges, besides being only necessary for articulatory purposes, are also not often perceived reliably as pauses by hearers (Spinos, O'Connell, & Kowal, 2002). It should be noted too that the majority of pauses in spontaneous speech fall within the range of 250ms to 1s (Goldman-Eisler, 1968) which makes it clear that setting the threshold at more than one second would exclude a significant amount of pausing behaviour.

Four dimensions of pause are examined for each response. First, turn latency was measured to the nearest 10ms as the time lapsed between investigators' prompts and suspects' responses. The latency duration includes

any turn-initial *um* or *ah* vocalizations. In the case of overlap, latency was recorded as a negative value (e.g., -0.6 for a turn overlap of 0.6s.) At the outset of the study, turns with greater than 750ms overlap were excluded because it is not always possible to reliably examine other prosodic elements while two speakers are talking at once. Furthermore it can also be difficult to reliably understand word for word what is spoken when there is considerable overlap.

Second, to establish the ratio of pause-to-speech, all periods of silence and all filled pauses such as *uhm* and *ah* were added together to determine how many seconds, to the nearest 10ms, were spent not producing content-laden speech. The ratio of pause-to-speech is computed by dividing the pause time by the duration of the whole response, beginning after any latency period.

Third, the number of pauses of all lengths beyond the threshold is totalled for the purpose of obtaining a measure of pause frequency. Because the responses in this study average approximately 6s, it was decided that pause frequency would be given as a "pauses per second" measure. While the measure of "pauses per minute" is a more standard figure for pause frequency throughout the literature, many of these studies focus on longer spans of speech such as free narratives or other retellings of events (e.g., Vrij & Mann, 2001). Pauses occurring in longer spans of speech serve to separate discourse segments (Swerts, 1997). This type of global-level discourse segmentation is not a concern in the present

data; rather, examining the frequency of pauses local to individual categories of responses is, and as such, a pauses per second measure is the most appropriate.

Fourth, the duration of all pauses beyond the 250 millisecond threshold occurring during responses is examined with the goal of making a comparison between the groups, the response types, and assessing within-speaker differences.

The response types examined for latency are affirmative, negative, relevant and confession. Irrelevant information is often embedded in other response types, and as such is not necessarily at the onset of a suspect's turn. Response types examined for the pause-to-speech ratio, frequency, and duration are relevant and *confessions*. Affirmative and negative are too short, and therefore do not incorporate within-turn pauses. The table below summarizes the design of the study on pause.

Table 3.4   Design of the study on pausal phenomena

Pausal Phenomena	Measure	Response Types
Latency	milliseconds	affirmative negative relevant confession
Pause-to-speech ratio	total pause time divided by total speaking time	relevant irrelevant (FT suspects only) confession
Pause frequency	pauses per minute	relevant irrelevant (FT suspects only) confession
Pause duration	milliseconds	relevant irrelevant (FT suspects only) confession

All response types except for confession will be analysed and treated statistically because there is a substantial number of them for each suspect in both groups. Confessions occur in much smaller numbers and not for all suspects, but they should not be overlooked owing to their importance in police interviews. Therefore, confessions will be treated on a more qualitative basis. Table 3.5 shows the number of responses for which pause is studied by suspect type.

Table 3.5   Distribution of response types analysed for pausal phenomena by suspect

	FT-1	FT-2	FT-3	FT-4	FT-5	FT-6	FT-7	FT-8	Total
Affirmative	27	9	22	22	31	11	10	18	**150**
Negative	21	16	6	23	19	9	19	20	**133**
Relevant	24	20	34	22	26	12	18	18	**174**
Irrelevant	19	12		20	16	20	16	15	**118**
Confession		5		4		7			**16**
Total	**91**	**62**	**62**	**91**	**92**	**59**	**63**	**71**	**591**

	RO-1	RO-2	RO-3	RO-4	RO-5	RO-6	RO-7	RO-8	Total
Affirmative	12	18	10	41	23	11	12	19	**146**
Negative	21	14	9	18	12	11	11	10	**106**
Relevant	28	14	23	18	21	7	26	14	**151**
Irrelevant									**0**
Confession		2		5	2				**9**
Total	**61**	**48**	**42**	**82**	**58**	**29**	**49**	**43**	**412**

A total of 1003 responses are examined for pausal features. Results follow in Section 3.4.

3.4 Results

In the following four subsections, the results of response latency, pause-to-speech ratio, pause frequency, and pause duration are reported. Discussion of the results follows in section 3.5.

3.4.1 Response latency

The time that elapses between an investigator's prompt and a suspect's response varies as a function of the type of response given, as well as the

suspect's first-time or repeat offender status. In this section I present the results

of the study on response latency in the corpus. An example of each response type

is shown in the table below. FT suspects are used here because the difference is

more apparent.

Table 3.6   Latency samples from four FT suspects

Suspect FT-6	Response Type
I: She uh she calls you Dad? S: (.41) Yeah.	Affirmative

Suspect FT-3	
I: Had you guys been arguing or anything like that? S: (-.75) No, no.	Negative

Suspect FT-8	
I: Ok how did you contact her? S: (2.63) Uh, through the office phone, my cell phone um, um even the phone in the um the um hallway.	Relevant

Suspect FT-6	
I: Ok tell me more. S: (1.5) I was touching her leg and then um you know and I was starting to get closer and she moved over so I moved over a little more uh when it's up top there's not much room right? So she moved back over so I started to put my hand down on her vagina and then I put my hand down her pyjamas.	Confession

In the examples above, the two short responses, affirmative and negative

are preceded by a considerably shorter latency period — both less than 1s, than

the two longer responses of the relevant and confession types, which are preceded by more than one second's delay. Despite some within-speaker variation, this difference holds well throughout the data. Table 3.7 shows the mean length of latency for each suspect in the two groups by response type.[17]

Table 3.7   Response latency by response type for both suspect groups

FT	FT-1	FT-2	FT-3	FT-4	FT-5	FT-6	FT-7	FT-8	mean	s.d.
Affirm	.64	.38	.31	.17	.23	.32	.09	.02	.27	.19
Negative	.26	.19	.94	-.01	.12	1.64	.14	.87	.52	.57
Relevant	1.00	1.81	.44	2.48	.75	2.57	2.04	4.06	1.89	1.18

RO	RO-1	RO-2	RO-3	RO-4	RO-5	RO-6	RO-7	RO-8	mean	s.d.
Affirm	.46	.45	.27	.89	.34	.43	.20	.15	.40	.22
Negative	.65	.11	.55	.40	1.73	.30	.36	.04	.52	.50
Relevant	.93	.62	.84	.92	2.11	.61	.45	.23	.84	.53

In general for both groups of suspects, affirmative and negative are preceded by a delay of approximately 0.5s or less, whereas relevant is preceded by a mean delay of more than a second or two in most FT cases and just below one second for most RO cases. Figure 3.1 shows the average latency values for each group of suspects across response types.

---

[17] Confessions will be dealt with separately.

Figure 3.1   Latency values of both groups by response type
Error bars indicate one standard deviation.
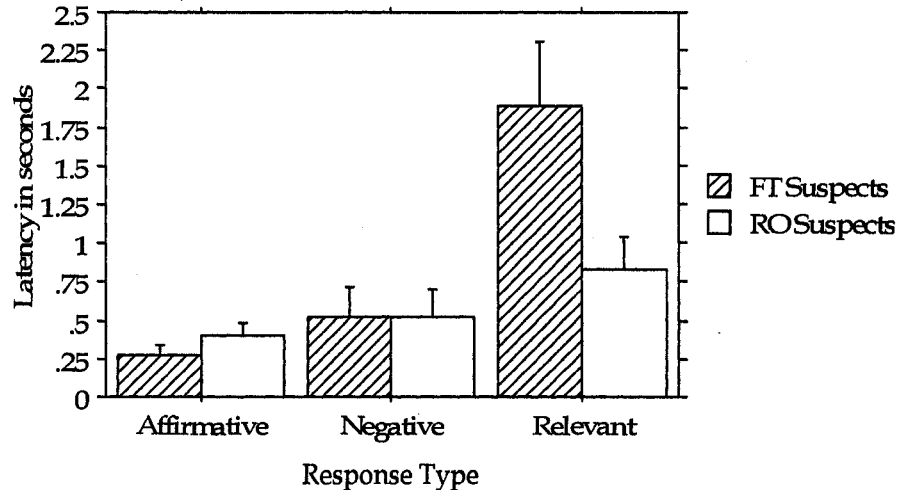
The response latency data were submitted to a mixed design ANOVA

with group (FT and RO) and response type as factors. There is an interaction

between group and response type $F(2,28)=4.919$, $p=.0148$. Post hoc Tukey tests

with Bonferroni adjustment (p<.0167) indicate significant difference between

affirmative and relevant, and negative and relevant. There is no between-group

difference in latency for either affirmative or negative. However, the difference in

time taken to commence a relevant response is considerably greater for FT

suspects than it is for RO suspects.

For those who produced them, confessions were preceded by a

considerable range of latency values. Examples follow from one of each type of

suspect.

Table 3.8   Confession latency value samples

Suspect FT-6
I: Now you said it happened more than once. How many times? S: (4.10) It happened the last couple of times. I: Tell me about this time in November. S: (3.20) I started you know, playing with her hair and stuff and just uh you know started tryin to touch *there* and she started moving around so.

Suspect RO-2
I: Can you tell me about the STORE A? S: (3.21) I just went in, tried to buy a couple of computers, told the clerk I wanted to buy a couple of computers, four computers, and when they brought them out I told the clerk to help carry'em out, clerk said no, I asked the manager to carry'em out, I told the manager I had a gun, which I didn't. I: Can you tell me about then, the STORE B? S: (0.62) Same thing. Exact same. Driving along. Saw a sign [points up] said they had computers, drove in and looked, same exact thing.

Suspect FT-6's confessions tend to come after considerable delay compared to Suspect RO-2's, which come quite readily. Figure 3.2 shows the latency values of all confessions in the data.

Figure 3.2   Latency values of confessions in seconds

Black dots represent the latency value of individual confessions, white dots indicate average latency values of suspects' relevant responses.
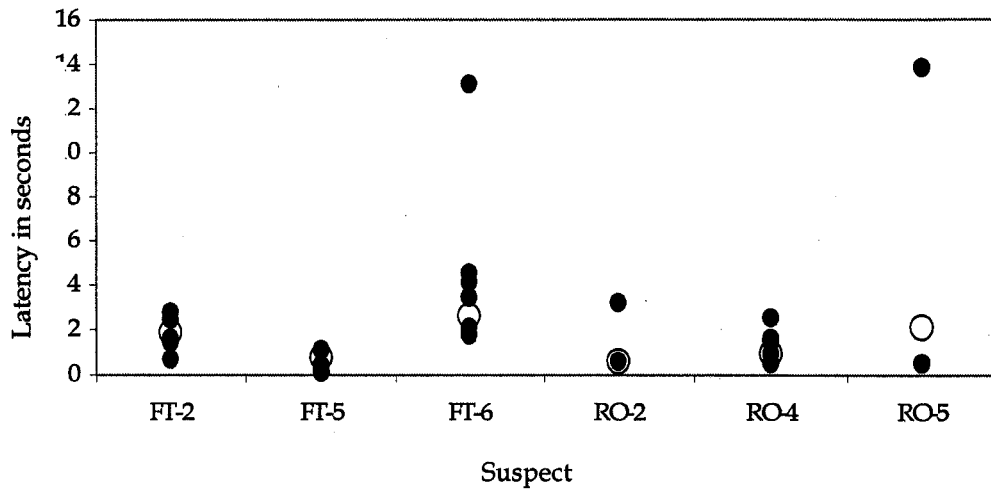
Figure 3.2 illustrates the wide range and lack of consistency among the latency values of confessions. Some confessions follow almost immediately after the prompt, while others may take several seconds. With these few figures, it is not possible to draw any firm conclusions from the data set. While relevant responses are preceded by significantly more delay than affirmative or negative, the correlation (see Figure 3.3) between the length of the confession and the degree of latency is not statistically significant (r=.204, p=.3421), as expected in situations where longer utterances require more planning. It should be noted that the longest latency values in the entire data set occurred during confessions.

Figure 3.3   Relationship between latency and response length of confessions
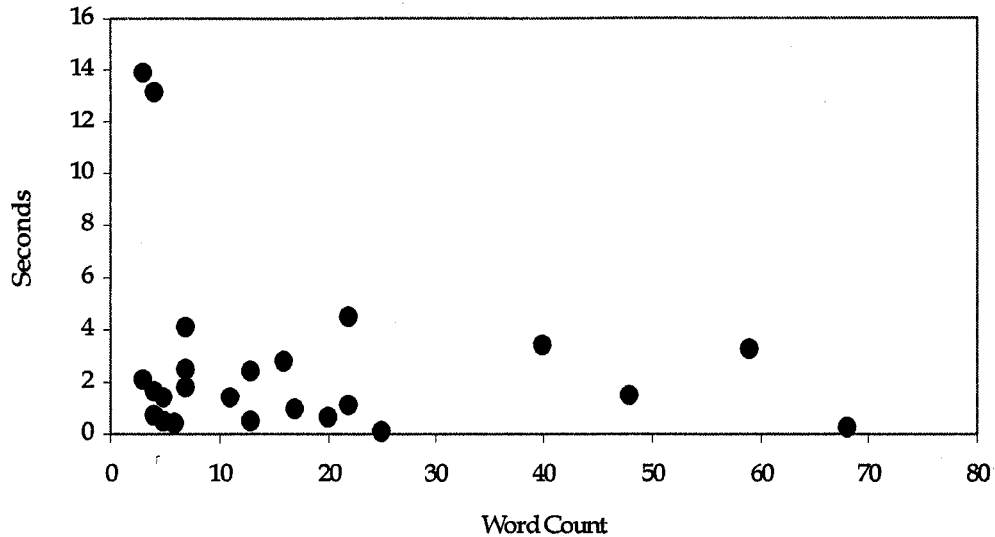
Inconsistencies aside, with respect to individuals' confession latency values, the only reasonable inference to be drawn from the data is that *confessions* appear to show latency periods closer to relevant *info* than to affirmative and negative.

3.4.2   Pause-to-speech ratio

In this section I report the results of the study on the ratio of pause-to-speech within turns across the two groups. Only responses in which a substantial turn is produced are analysed for within-turn pause. In other words, only relevant, irrelevant and *confessions* are examined because there is enough speech for within-turn pausing to occur.[18] While it is possible for only relevant and

---

[18] Irrelevant responses are very infrequently produced by RO suspects, and suspect A3 of the FT group did not produce enough irrelevant responses to measure prosody reliably. See Chapter 2, section 2.3 for discussion.

irrelevant to be analysed statistically, due to there being a large number of tokens, confessions are reported but not included in statistical analyses. Affirmative and negative are excluded for being too short for within-turn pausing to occur. The pause-to-speech ratio was calculated as the sum of all pauses during an utterance, excluding latency, divided by the total duration of the utterance. number of examples follow.

Table 3.9   Examples of pause values for relevant and irrelevant responses for FT-1

Suspect FT-1	Relevant
I: What time did you leave the empress? S: Oh I dunno probly at uh **(0.54)** I dunno last time I looked it was probably **(1.2)** after nine or something	
Total duration of response	7.30 sec
Total pause time within turn	1.74 sec
Pause-to-speech ratio	0.24
Suspect FT-1	Irrelevant
I: Who was there at that time? S: She's got friends left and right all the time. **(0.68)** They're even poppin in just takin off.	
Total duration of response	4.80 sec
Total pause time within turn	0.68 sec
Pause-to-speech ratio	0.14

Table 3.10   Examples of pause values for a relevant responses for RO-1

Suspect RO-1	Relevant
I: Ok tell me what they said to you. What did they tell you they were going to do? S: They didn't go into fine detail, **(0.41)** "drive the damn van drop this guy off." One guy stayed in the vehicle the other one went into the store and I was told where to go, left right blah blah blah **(0.48)** go pick this guy up. And drive. **(0.46)** That van never stopped.	
Total duration of response	13.67 sec
Total pause time within turn	1.35 sec
Pause-to-speech ratio	0.10

The examples above are closely representative of the averages found in the data. Suspect FT-1's pause ratio is greater in his relevant response than in his irrelevant response, although it will be shown that that pattern does not hold throughout the data. Suspect RO-1's relevant response pause ratio is lower than Suspect FT-1's relevant response. Table 3.11 summarizes the average values for each speaker.

Table 3.11   Pause-to-speech ratio by response type for both groups

FT	FT-1	FT-2	FT-3	FT-4	FT-5	FT-6	FT-7	FT-8	mean	s.d.
Relevant	.08	.35	.16	.19	.09	.13	.16	.31	.18	.10
Irrelevant	.05	.29		.20	.17	.14	.15	.22	.17	.19

RO	RO-1	RO-2	RO-3	RO-4	RO-5	RO-6	RO-7	RO-8	mean	s.d.
Relevant	.13	.17	.12	.06	.06	.10	.12	.02	.12	.05

The pause-to-speech ratio data of relevant responses were submitted to a

t-test for independent samples. It was found that the ratio is higher among FT

suspects than it is for RO suspects: *t*(14)=2.236, p=.0422. Within the FT suspect

group, for whom it is possible to compare relevant and irrelevant responses, no

significant difference in the pause-to-speech ratio is found in a paired

comparison *t*-test (*t*(6)=.619, *p*=.5589). Figure 3.4 illustrates the differences

between the two groups for relevant and, for FT, the similarity between relevant

and irrelevant info.

Figure 3.4   Pause-to-speech ratio

Error bars indicate one standard deviation.
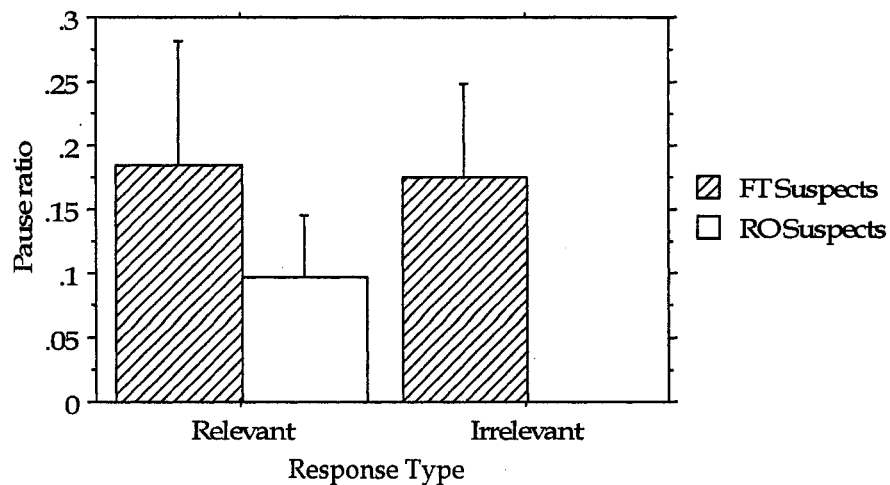
That FT suspects show a higher pause-to-speech ratio than RO suspects is

unsurprising, given that their response latency was higher as well (see Section

3.4.1). Overall, this result suggests that FT suspects are simply more hesitant

under the circumstances, with more interruption in their speech.

For those who produced them, confessions show a tremendously wide range of pause-to-speech ratio values. In Figure 3.5, the pause ratios are given for suspects' individual confessions.

Figure 3.5   Pause-to-speech ratio of confessions

Solid dots represent the pause ratio for each confession by suspect. Hollow dots represent the average pause ratio for relevant responses.
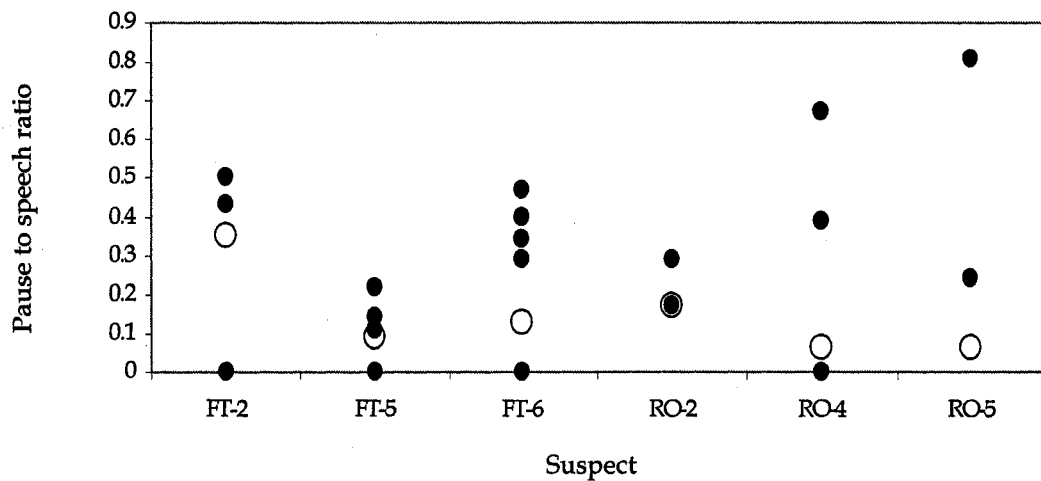
Except for suspects FT-5 and RO-2, whose pause ratios cluster reasonably close, others' pause ratios show a great deal of variation. Although suspects' relevant response pause ratios appear lower overall, they actually are not. Suspects FT-2 and FT-6 both produced three confessions having no pauses within them at all, and FT-5 and RO-4 also have at least one pause ratio between zero and 0.06. Therefore, the pause ratio values for relevant responses fall midway between high and low values for confession pause ratios.

3.4.3 Pause frequency

The pause-to-speech ratio is one means of assessing the amount of talk that takes place during a turn. Closely related to this ratio is the number of pauses in a given time span. Taking into account the same response types for the two groups, pause frequency is also examined. In this study, pauses greater than 250ms were counted on a per second basis. Examples follow:

Table 3.12   Within-turn pause samples from Suspect FT-2

Suspect FT-2	Relevant
I: She said there the nature of your hugs and kisses were different when you were in private with her than when you were in public with her family What do you think about that? S: I don't like to think they were **(1.17)** I don't think I ever **(0.89)** fondled her **(.59)** or grabbed her ass **(1.92)** or anything like that.	
Total duration of response	9.50 sec
Number of pauses	4
Pauses/sec	0.42

Suspect FT-2	Irrelevant
I: Is there any reason why if a swab of the internal cavity of her vagina, that there would be any of your DNA. S: Absolutely no. No. And if you want, you just go right ahead. [referring to taking DNA sample from suspect] I: What about the bra? S: The bra? I don't think I touched her bra. Irrel→ I mean I can't even get an erection anyway so what the fuck **(0.25)** what good would it do me?	
Total duration of response	3.41 sec
Number of pauses	1
Pauses/sec	0.29

Table 3.13 Within-turn pause sample from Suspect RO-7

Suspect RO-7	Relevant
I: And what's he doing? S: He's smoking crack the whole time. **(1.52)** He's smoking crack this whole time. **(0.80)** While we're waiting, the whole time he's smoking crack.	
Total duration of response	7.04 sec
Total pause time within turn	2
Pause-to-speech ratio	0.28

In the examples above, suspect FT-2's relevant response contains 4 pauses over 9.5 seconds yielding a pause frequency value of 0.42/second. The same suspects' irrelevant response contains a single pause within a 3.41 second turn resulting in a pause frequency value of 0.29/second. Suspect RO-7's relevant response contains 2 pauses over 7.04 seconds, at a rate of 0.28/second, considerably lower than suspect FT-2's. These examples are representative of the rest of the data, although differences between response types and groups will be shown not to be significant. The averages for each suspect are shown in Table 3.14

Table 3.14   Pause frequency for both groups, taken as the number of pauses per second

FT	FT-1	FT-2	FT-3	FT-4	FT-5	FT-6	FT-7	FT-8	mean	s.d.
Relevant	.14	.14	.15	.40	.15	.12	.27	.35	.22	.11
Irrelevant	.08	.36		.04	.37	.20	.27	.32	.23	.13

RO	RO-1	RO-2	RO-3	RO-4	RO-5	RO-6	RO-7	RO-8	mean	s.d.
Relevant	.26	.28	0.15	.04	.11	.17	.19	.01	.15	.10

The means for pause frequency in both groups is reported, and for all responses, parallels what is found for pause ratio in that FT suspects' relevant responses have higher pause frequency than ROs'. Unlike pause ratio, however, there is somewhat more variation resulting in non-significant differences. Figure 3.6 illustrates the findings.

Figure 3.6   Pause frequency taken as the number of pauses per second for both groups Error bars indicate one standard deviation.
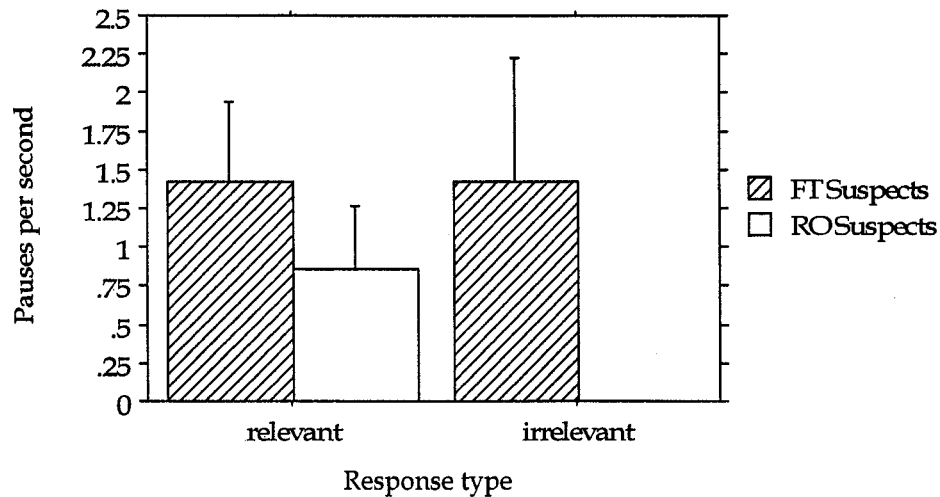
An unpaired *t*-test showed no significant difference between FT and RO suspects' pause frequency values for relevant responses ($t(14)=-1.238$, $p=.2361$). Similarly, a paired *t*-test showed no significant difference between FT suspects' pause frequency values in relevant and irrelevant responses: $t(6)=.133$, $p=.8984$. It is rather surprising that the pause frequency for irrelevant information would not be different from relevant information. In general, irrelevant information is also 'safe' information, in that it does not relate to the investigator's line of questioning. We might assume greater fluency or ease of speech when the suspect is addressing something less relevant to the investigation than more relevant, as would be the case with relevant informationresponses.

Pause frequency in confessions varied considerably, both between and within individuals' values, with some confessions having no pauses, and some having as many as 0.4 to 0.5 per second, which is higher than the means for relevant responses. This wide variation is shown in Figure 3.7 below and an example follows.

Figure 3.7   Pause frequency in confessions

Solid dots indicate the frequency of pauses per second within individual confessions. Hollow dots show suspects' relevant response pause frequency means.
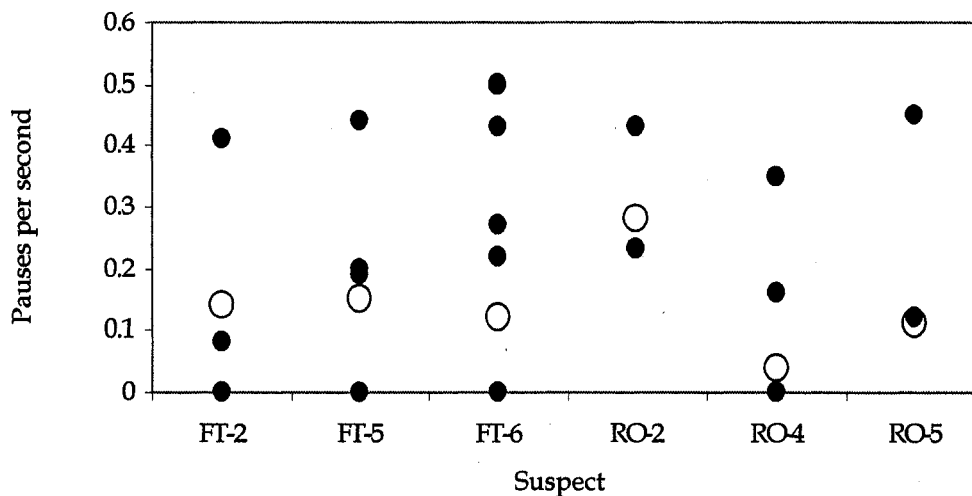
Table 3.15   Within-turn pause samples from Suspects FT-5 and FT-2 in confessions

Suspect FT-5	Confession
I: What happened? Did you hit him like this? S: I hit him like this [physically demonstrating] like it was just **(0.97)** like that and like that was it. **(0.44)** But it wasn't like full on boom boom it was just like, you know.	
Total duration of response	9.86 sec
Number of pauses	2
Pauses/sec	0.20

Suspect FT-2	Confession
S: (3.21) [19] I just went in, **(0.38)** tried to buy a couple of computers, told the clerk I wanted to buy a couple of computers, **(1.06)** four computers, **(1.19)** and when they brought them out I told the clerk to help carry'em out, clerk said no, I asked the manager to carry'em out **(0.35)**, I told the manager I had a gun, which I didn't.	
Total duration of response	9.24 sec
Number of pauses	4
Pauses/sec	.43

Despite the wide variation, overall, pauses were more frequent in confessions than in relevant responses and no consistent pattern turns up for any one suspect.

3.4.4   Pause duration

We have now seen that turn latency and pause-to-speech ratio differ on the basis of group and in some cases, the type of response produced, while no

---

[19] Latency values do not figure in to the analyses of within turn pause.

differences are found between or within groups for pause frequency. In this

section the results of pause duration are reported. As is the case with the pause-

to-speech ratio and frequency, only responses of substantial length are examined,

again because they contain sufficient speech for a combination of speaking and

pausing to occur. Table 3.16 shows the average values of suspects' pauses, across

groups for the examined response types.

Table 3.16   Suspects' average pause durations

FT	FT-1	FT-2	FT-3	FT-4	FT-5	FT-6	FT-7	FT-8	mean	s.d.
Relevant	.28	1.20	.57	.67	.46	.68	.67	1.21	.72	.33
Irrelevant	.19	.90		.64	.47	.63	.56	.78	.60	.22

RO	RO-1	RO-2	RO-3	RO-4	RO-5	RO-6	RO-7	RO-8	mean	s.d.
Relevant	.65	.80	.51	.52	.22	.39	.48	.10	.46	.22

FT suspects' relevant pause duration means generally longer than their

irrelevant pause durations, although the difference will not prove to be

significant. RO suspects' relevant pause durations average less than FT's

although, this difference too, will prove not to be significant. Figure 3.8 illustrates

the mean values for both groups.

Figure 3.8   Pause durations

Error bars indicate one standard deviation.
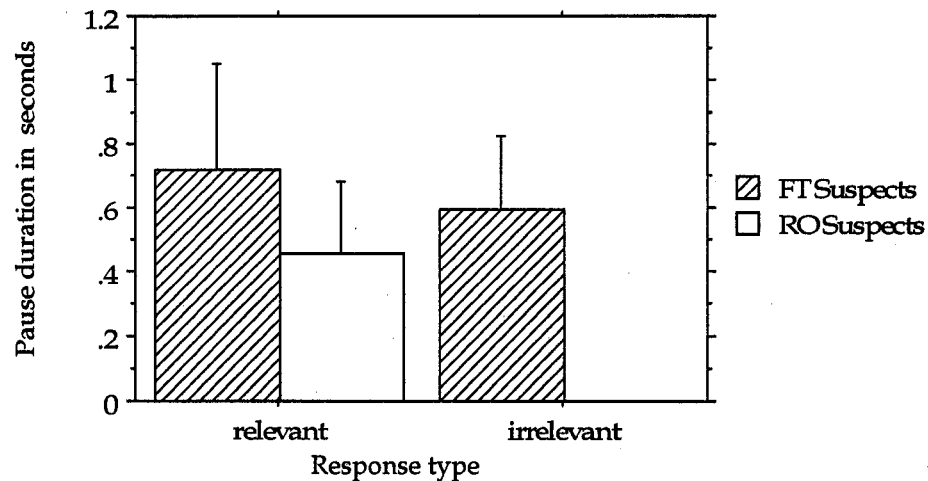
A *t*-test for independent samples shows no significant difference between FT and RO's average pause durations for relevant responses ($t(14)=1.837$, $p=.0875$). The difference between FT's relevant information and irrelevant information is also not significant ($t(6)=2.346$, $p=.0574$).

It is interesting to note that FT and RO suspects' differing significantly on pause ratio means while the differences for each pause frequency and duration do not differ significantly. It is therefore inferable that neither pause frequency nor duration alone makes a larger contribution to pause ratio.

Suspects who produced confessions show a great deal of variation, as has been the case with all other pause dimensions. The average values for pauses occurring in confessions is shown in Figure 3.9.

Figure 3.9   Average pause duration during confessions.

Solid bars illustrate average pause duration of confessions, dark diamonds illustrate the same suspects average pause duration for *relevant* info.
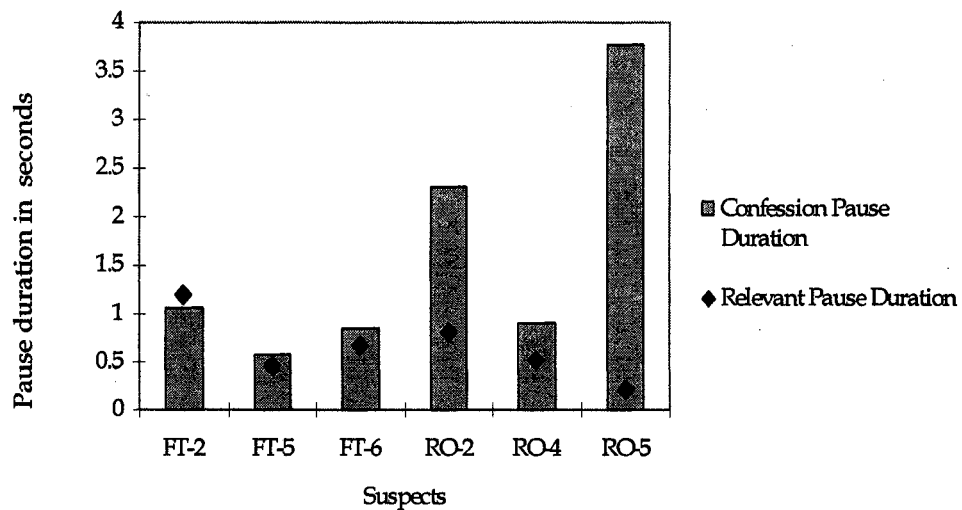
Pause durations, overall, tend to be longer during confessions than relevant information for all but Suspect FT-2. It also appears that RO suspects' relevant pauses are much shorter than their confession pauses.

3.5   Pause profiles and conclusions

The previous sections present the results of four pausological phenomena in the corpus: response latency, pause-to-speech ratio, pause frequency, and pause duration. In this section, profiles of the two groups' pause behaviour are given based on the preceding results. A summary of results relating to the research questions listed in Section 3.1 concludes the chapter.

3.5.1 Pause profiles and discussion

Taking into account all results from the previous sections paints a picture

of suspects' pause behaviour in this set of interviews. The figures below present a

subjective visual summary of the statistically significant differences as well as

strong tendencies within the pause results concurrently for each group, based on

the values illustrated in Figures 3.1, 3.4, 3.6, and 3.8. Confessions are added for

comparison. Greater pause feature values are represented with larger cones, and

lower pause feature values with smaller cones, and mid range values with

medium sized cones. Flat discs represent variables that vary too widely to be

considered a reliable profile.

Figure 3.10   FT pause features by response type, showing relative values
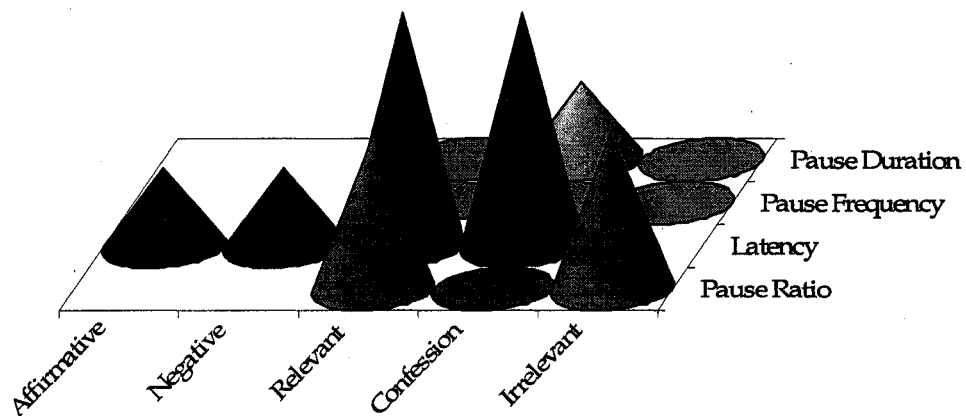
Figure 3.11   RO pause features by response type, showing relative values
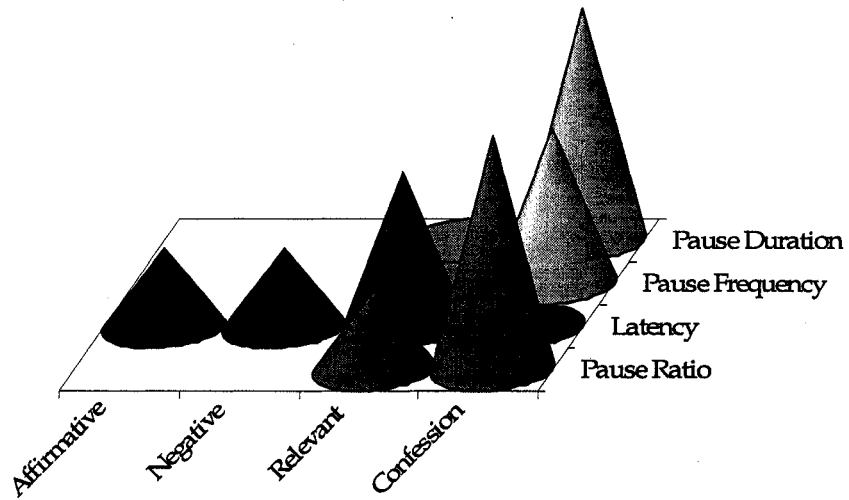
The figures above provide a simple visualization of the relative values of pause features found in the data. A summary of which differences reached levels of significance is given in the table below. Statistical *t* and *F* test values are reported in previous sections. Naturally, confessions are excluded from this summary since they were too few to treat statistically.

Table 3.17 Summary of significant and non-significant differences by group and response type

Feature	Comparison		Significance	Test
	Group	Response Type		
Latency	FT/RO	Aff/Neg	n.s.	F[a]
		Neg/Rel	<.05	F[a]
		Rel/Aff	<.05	F[a]
Pause to speech ratio	FT/RO	Rel	<.05	t[b]
	FT	Rel/Irrel	n.s.	t[c]
Pause frequency	FT/RO	Rel	n.s.	t[b]
	FT	Rel/Irrel	n.s.	t[c]
Pause duration	FT/RO	Rel	n.s.	t[b]
	FT	Rel/Irrel	n.s.	t[c]

[a]mixed design; [b]independent; [c]paired

That both groups' latency values are greater for relevant and confession than affirmative and negative is unsurprising given that a greater deal of turn planning (Goldman-Eisler 1968) ought to be expected ahead of long turns. When the suspect is prompted to provide more than a simple yes or no, it seems they pause longer before provide details that might bring about legal consequence.

The most striking differences between first-time (FT) and repeat offenders (RO) occur in the pause features for relevant responses and it is worth noting that FT suspects are significantly slower to respond, taking on average slightly more than twice the amount of time before offering relevant information than RO suspects. FT suspects also tended to pause more frequently, and for longer periods than RO suspects, although the results did not reach significance,

resulting in greater pause-to-speech ratios. Looking more closely at relevant

responses for both groups, and for FT suspects, the irrelevant responses,

correlations of pause values are given in Table 3.18. to give an idea of how the

pausal features co-occur in responses of substantial length.

Table 3.18   Correlation coefficients for relevant and irrelevant pause values

| | Relevant | | Irrelevant |
| | FT | RO | FT |
	r	r	r
Latency/Pause Ratio	.566	.135	na
Latency/Pause Frequency	.385	.193	na
Latency/Pause Duration	.686	.146	na
Pause Ratio/Pause Frequency	.758*	.881**	.112
Pause Ratio/Pause Duration	.965**	.889**	.289
Pause Duration/Pause Frequency	.683	.690	.721*

*p<.05, **p<.01

There are four statistically significant strong correlations found in

pairwise comparisons of pause ratio and pause frequency, and pause ratio and

pause duration for relevant responses in both suspect groups. This result is

unsurprising, given that relatively longer and relatively frequent pauses are

reasonably expected to yield a higher ratio of pausing, and the opposite would

be true as well: the lower the pause ratio, the shorter and less frequent the

pauses.

It is surprising that the differences between FT's relevant and irrelevant responses were all not significant. The nature of irrelevant information provided in an interview is generally 'safe' information, and for that reason, the same amount of hesitation as relevant information was not expected. Furthermore, to the listener, irrelevant information 'sounds' faster, perhaps giving the impression of less pausing. Chapter 5 will address the speech and articulation rates of the same, and it will be shown that although pauses do not differ in this respect, tempo does.

In a number of studies on police interviews and deceptive speech, the stress associated with being interviewed and with lying results in more and longer pauses (Mann, Vrij, & Bull, 2002; Vrij, Edward, & Bull, 2001; Vrij & Mann, 2001). The greater the amount of planning required to produce a turn, the greater the pausing (Goldman-Eisler, 1968). FT suspects' taking longer than RO to produce relevant responses and pausing longer and more frequently, veracity issues aside, may mean that the stress caused by being in an unfamiliar situation brings with it more pause-filled behaviours. If FT suspects are fearful of the investigation outcomes, then their longer pauses would be at odds with the findings of Sobin and Alpert (1999) who find short and infrequent pauses in fearful states.

The confession pause values for the suspect groups, while not particularly consistent, are on the whole greater than relevant pause values. A number of

studies conclude that sadness and remorse are accompanied by more and longer pauses, (Banse & Scherer, 1996; Johnstone & Scherer, 2000; Kehrein, 2003; Laukka, Juslin, & Bresin, 2007). It is tempting to suggest that feelings of sadness might accompany a confession, resulting in longer latency periods and longer more frequent within-turn pauses. However, suspects' emotional experiences are not confirmable at this stage. Nonetheless, it is curious that RO suspects' confessions have longer and more frequent pauses than their relevant responses and that the same was not found for FT suspects.

From a forensic point of view, the findings on pause behaviour, particularly with respect to relevant and irrelevant responses and confessions are important given that these are content-laden utterances, with details generated by the suspect.[20] Studies that address suspects' speech for purposes of detecting deception ought to take into account whether suspects are new to this discourse context, or they have prior experience because pause behaviour, which is examined in all deception studies, is demonstrably different when offering substantial information.

3.5.2 Conclusion

In this chapter, four aspects of suspects' pause behaviour have been examined. In this concluding section, I address the research questions put forth in section 3.1, repeated below:

---

[20] As opposed to details generated by the interviewer as is the case when yes/no questions are asked.

Do response types and suspects' FT (first timer) or RO (repeat offender) status affect:

1) the latency between investigators' prompts and suspects' responses?

2) suspects' within-turn pause-to-speech ratio?

3) suspects' within turn pause frequency?

4) pause duration?

The results of each portion of the study on pause showed some positive findings. First, latency was found to be affected by both the type of response, and the suspect group. Both groups of suspects took considerably less time before offering a "yes" or "no" than before producing a lengthier response such as relevant information or a confession. From the point of view of discourse planning (Goldman-Eisler, 1968), this is to be expected, since less planning is required to produce a one-word answer than a multi-word answer. It was also found that first-time suspects took longer before producing relevant responses than repeat offenders. It is tempting to suggest that first time interviewees could be less at ease than repeat offenders due to a lack of familiarity with legal process, causing them to delay more before responding.

Second, FT suspects' pause-to-speech ratios are higher than ROs' pause-to-speech ratios when producing relevant information. It is rather surprising that there were no differences between FTs' relevant and irrelevant responses in terms of the pause ratio. If irrelevant information is safer information, one might

expect such responses to be produced with the relative ease, and hence less pausing. Confessions showed considerable variation in terms of the pause-to-speech ratio, although in general, for each speaker, the ratio tended to be higher than the relevant responses.

Third, the frequency of pauses was significantly different for the two groups for relevant responses, with FT suspects pausing more frequently than RO suspects. Pause frequency was not different, however, for relevant and irrelevant among the FT suspects.

Fourth, analysis of pause duration showed no significant differences between the two groups for relevant responses, an unexpected result given the differences found in terms of all other pause dimension for the same comparison. For FTs' relevant and irrelevant responses, no significant difference was found for pause duration.

While differences were not found in all areas expected, first-time and repeat offenders did show significantly different behaviour in time taken to begin a turn and pause ratios.

4 TEMPO

The trouble with talking too fast is you may say something you haven't thought of yet.

--Ann Landers

Tempo, the second prosodic property in this study on suspects' speech, is examined in this chapter. The rate at which suspects speak is explored in two ways. First, I examine the speech rate of suspects' responses in syllables per second to conclude that in some cases, speech rate varies according to response type and suspect category. Second, the articulation rate of responses, taken as a measure of segments per second is examined, with similar conclusions drawn.

4.1 Introduction

On an individual speaker basis, a range of factors contribute to considerable variation in speaking rates (Crystal & House, 1982, 1988; Goldman-Eisler, 1968), many of which also account for the frequency and duration of pauses that occur during speech.[21] These factors include emotional state (Banse & Scherer, 1996; Fairbanks & Hoaglin, 1940; Johnstone & Scherer, 2000; Scherer, 2003; Sobin & Alpert, 1999) and mental health (Nilsonne, 1987, 1988); genre, (Barik, 1977; Hegedüs, 1953; Sabin, 1976; Sabin, Clemmer, O'Connell, & Kowal,

---

[21] See Chapter 4 for an overview of pause.

111

1979); linguistic factors such as sentence type (van Heuven & van Zanten, 2005);

biological effects such as neuromuscular constraints (Tsao & Weismer, 1997;

Tsao, Weismer, & Iqbal, 2006); complexity of speaking task (Essen, 1949;

Goldman-Eisler, 1968); and age (Kowal, O'Connell, & Sabin, 1975; Shipp &

Hollien, 1972).

The rate at which people speak has been claimed to vary from one

language to another, although troubling to any cross-linguistic comparisons are

the vastly different units of measurement. British English, for example, is

produced at 3.16 to 5.33 syllables per second (syll/s) (Tauroza & Allison, 1990),

and American English at 3.66 syll/s (Robb, Maclagan, & Chen, 2004). French and

Dutch tend to be spoken a little more rapidly, from 4.31 to 5.73 syll/s for the

former (Grosjean & Dechamps, 1973; Malecot, Johnston, & Kizziar, 1983), and 5.2

syll/s for the latter (Blaauw, 1995). Spanish, at 6.57 syll/s and Portuguese at 7.81

syll/s are spoken seemingly quicker still (Rebollo Couto, 1997). While it might

seem that these studies ought to be producing comparable results by using the

same metric, syllables per second, this is not necessarily the case. Languages with

simpler syllable structures, such as Spanish and Portuguese, and those with more

complicated syllable structures such as English, show higher and lower syllable

per second rates respectably.

Before discussing the rate at which speakers talk, it is important to note a

few key definitions of technical terms to be used in this chapter. Unless referring

112

to the work of others, in this study the terms *speech rate* or *rate of speech* will refer

to a measure of syllables per second, including pauses, in an utterance..

Articulation rate will refer to a measure of segments per second within an

utterance, excluding pauses, to describe only the actual rate of vocalization. The

term *speaking rate* will be used to discuss in general terms the *rate of speaking*

without referring to a specific measure.

A point to make clear in this chapter, as has been made in the previous

two chapters, is that the goal is not to attribute causes to suspects' variable

speaking rates along the lines of the factors mentioned in brief above, but rather

to describe and compare speaking rate phenomena as they occur within and

between the two suspect groups, for the various response types produced during

police interviews. It is assumed that many of the known causes of speaking rate

differences are at work among the suspects in this study, but at this stage, the

means to collect self-assessments from suspects is not possible. The aims,

therefore, of this portion of the study are descriptive rather than explanatory.

In this chapter, I offer an examination of two temporal aspects of suspects'

speech stemming from the following research questions:

Do response type and suspects' FT or RO status affect

1) the speech rate at which the suspect produces the response?

2) the articulation rate at which the suspect produces the response?

In order to address these questions, this chapter is organized in the following manner. Section 4.2 deals with tempo studies in the literature. In the first subsection of 4.2, I survey the intersections of tempo and discourse, and tempo and emotion. In the next subsection, I look specifically at studies of police interviews and deceptive speech in which tempo has been explored. In section 4.3, I outline the methodology for this portion of the study, and offer a rationale for the metrics used, as well as an overview of the research design. Section 4.4 presents the results of the study, treating speech rate and articulation rate in their respective subsections. A correlation of the two rates and tempo profiles for response types are found in Section 4.5, along with some discussion of the findings in relation to previous work, and concluding remarks addressing the research questions.

4.2   Tempo in the literature

Speaking behaviour is studied from a number of perspectives and in the following two subsections, I provide an overview of the literature from two relevant backgrounds. First, I explore the literature on speaking rate and discourse, and speaking rate and emotion. Second, I survey the studies of police interviews and deceptive speech, in which speaking rate has been taken into consideration, often with other prosodic cues such as pause.

4.2.1 Speaking rates in discourse and emotion studies

Much the same way that pitch and pause are cues to discourse organization and the emotional condition of speakers, so too is the rate at which they speak. From a linguistic perspective, speakers accelerate and decelerate throughout a period of speaking in order to achieve a range of effects in terms of turn participation and discourse structure. From a psychological perspective, higher speaking rates are often associated with highly activated states of emotion, and lower speech rates with states that are more subdued (Kehrein, 2003; Laukka, Juslin, & Bresin, 2007; Scherer & Orshinksy, 1977). An overview of the research carried out in these two areas will be explored in this subsection.

To underscore the complexity and the extent to which speech rate can vary within individual languages and even within sentence types, van Heuvan and van Zanten (2005) show that in Dutch, Orkney English, and Manado Malay, polar questions are uttered more rapidly than statements. Furthermore, acceleration happens in different parts of the question. In Manado Malay faster speech is found closer to intonation boundaries, Orkney English questions are faster throughout, and in Dutch, the middle of a question is produced more rapidly than the outer edges. They speculate that the reason for the difference between sentence types may be due to the correlation between higher speech rates and the higher overall fundamental frequency typically associated with questions in languages throughout the world.

The type of speaking task can have an effect on speaking rates. Different

degrees of spontaneity and preparation affect how rapidly or slowly a person

speaks. Looking at French and English speakers producing spontaneous stories,

discussions about a film, a lecture, a formal speech and a written text, Barik

(1977) found an inverse relationship between speech and articulation rates. The

more prepared the speaking material, e.g., written texts and speeches, the higher

the speech rate (syll/min). Conversely, the more unprepared the speaking

material, the higher the articulation rate (syll/min, minus pauses). He suggested

that the interaction between tempo and pause is responsible for the stylistic

differences between speaking tasks, and that speakers' attention to articulatory

detail may be increased when delivering prepared speech, as opposed to

spontaneously-produced narrative.

It has been shown that changes in speaking rate occur in order to help

'contextualize' information contained in speech (Uhmann, 1992). In one study of

German speakers, accented syllables were distributed evenly throughout the

telling of foreground information (Selting, 1992). Such even distribution, causing

shortened intonation units, means that within narratives, stages of complicating

details[22] entail a perceptible acceleration of speech. In other parts of the narrative,

this pattern was not evident. Changes in speaking rate then, can signal

contextual information to the listener.

---

[22] See especially Labov (1972) for a discussion of narrative components including complication.

Closely related to the contextualization effect noted above, at least two studies have shown that 'parenthetical' information is produced at a higher speech rate than non-parenthetical information (Grosz & Hirschberg, 1992; Hirschberg & Grosz, 1992). This change again signals to the listener the difference between pertinent and peripheral information in the message.

An earlier study which brought together discourse and the mental state of speakers having a combined effect on speech rate was Essen (1949), followed by Goldman-Eisler (1956b, 1961, 1968). Both concluded that the more complex the speaking task, the greater the psychological tension experienced by the speaker, and hence the slower the speaking rate. Essen's study leads naturally into later studies that take emotion and attitude into account for speakers' variation in speaking rates.

Banse and Scherer (1996) and Johnstone and Scherer (2000) report comparably fast articulation for fear and anger, and slow articulation for sadness. They report 'some evidence' for an elevated rate of articulation for joy, although rate of speaking, as we saw with pause in the previous chapter, is more difficult to determine for joy. In a replication of Banse and Scherer (1996), Sobin and Alpert (1999) report similar findings for each fear, anger, sadness and joy with respect to speaking rate.

In a study of feigned depression and feigned sleepiness (Reilly, Cannizzaro, Harel, & Snyder, 2004), pretenders consciously decreased their

speech rate as in states of real depression or sleepiness; however, they were unable to convincingly modify their pitch range. Reilly et. al's study highlights the importance of prosody in forensic linguistic and forensic psychological research—a point I return to in Chapter 6.

Using a measure of syllables per second, Kerhein (2002) compared the speech rates of speakers engaged in a laboratory task in which participants who were unable to see each other were required to instruct each other in the construction of a Lego object. He found that speech rate increased as subjects became more excited, agitated, eager or angry, and that speech rate fell when subjects were calm, relaxed and content. Slow speech rates, however, also occurred when subjects became demotivated or "irked."

A speaker's attitude, besides emotion, may also contribute to differences in speaking rates. A mocking or mimicking message, for example, is popularly known to be conveyed by a high-pitched, whiney voice quality and lengthened syllables. Also, sarcasm as a spoken rhetorical device, has a slower overall speech rate, accompanied by greater intensity, and lower pitch (Rockwell, 2000).

The preceding survey gives the reader an idea of the range of possible factors contributing to varying rates of speaking. It is unlikely that a single factor determines speaking rate, as a speaker is simultaneously under the influence of his or her emotional state, the degree of simplicity or complexity of the speaking task, and he or she may wish the message to be delivered in a particular manner.

The notion that combined effects converge, resulting in a set of prosodic behaviours has been discussed by Kowal, O'Connell, & Sabin, (1975) and O'Connell & Kowal, 1983 and aptly called "multideterminism". The concept of multideterminism will be discussed in Chapter 6, when a rationale is given for identifying suspects' psychological states under interview conditions for forensic application.

4.2.2 Speaking rate in police interviews and deceptive speech

Taking a close look at how people say what they say has been of interest to those in law enforcement who seek investigative techniques to uncover deception. To this end, studies have been carried out both in the field and in the lab on deceptive speech. In the previous chapter on pause, a number of these studies were highlighted and will again be discussed, this time with attention to researchers' methods and findings on tempo in police interviews and deception.

It is believed that typically liars feel fear (of getting caught), excitement (at the thought of fooling someone), or guilt (for not being truthful) (Ekman, 1992), and that experiencing these emotions might contribute to a slowing of speech. Add to the heightened emotional state the complexity of creating false information and it seems unlikely that lying could be done by many without a

decrease in rate, a commonly drawn conclusion in the majority of studies looking into deceptive speech.[23] A handful of studies are touched on below.

In Vrij and Mann (2001), the comparison of truths and lies in the speech of a murder suspect included a subjective assessment of his rate of speaking. Two raters assessed whether the speech rate was slow, moderate, or fast, before and during the suspect's confession. Under both conditions, lies were consistently told at a lower rate of speech.

In a lab-based study of deception in which students belonging to either truth-telling or lie-telling groups, Vrij, Edward and Bull, (2001), examined speech rate in conjunction with other prosodic features such as pause, as well as an enormous range of non-verbal behaviour and gestures. Subjects were given questions (some in advance, and some spontaneously) regarding a vignette and were required to speak truthfully or lie. They found, quite unexpectedly, that liars tended to speak faster than truth tellers.

Not all studies have revealed that there is a speaking rate difference between truth and deception, however. Anolli and Ciceri (1997) found no significant difference in speech rate or articulation rate under the conditions of truthfulness, prepared lying and unprepared lying. Other speaker behaviours were different in the three conditions, particularly with respect to pitch, but temporal features including pause were not.

---

[23] See Vrij (2000) for a summary of studies on deceptive speech.

Authors of studies that found rate differences in truthful and deceptive conditions attribute the slower rates found in lying to speakers' increased need for planning while talking, owing to the notion of cognitive load (Goldman-Eisler, 1968). As discussed in Chapter Four on pause, greater planning requires more time, resulting in more and longer pauses, and therefore in a deceleration in the rate of speaking. In Chapter 6, the interaction between speaking rate and pause will be more closely examined.

4.3   Methodology

In this section, I discuss the methods by which speaking rate is examined in this study. I survey the various means used by previous authors in speaking rate studies and offer a rationale for the measurements used herein. The design of the study on speaking rate is described at the end of this section.

The metrics used to assess the rate of speaking can be subjective, where subjects qualitatively rate speech on a scale from slow to fast, and they can be objective, where the number of a linguistic unit is counted and divided by a period of time. A sample of these types is described below.

In some studies, raters are asked to provide a subjective assessment of speakers' speech rates. Generally, a 3 to 5 point Likert scale is used and raters indicate whether they find the speech slow, moderate or fast (Mann, Vrij, & Bull, 2002; Rockwell, 2000; Vrij & Mann, 2001). Listeners' perception of speaking rates

appears to be surprisingly uniform, as inter-rater reliability is high in each of these studies.

Objective means of measuring speech rate are more commonly used than raters' impressions. In these studies, researchers count occurrences of a particular linguistic unit within a period of time. Scollon (1981), for example, computed the 'density' of speech by counting the number of words per minute to demonstrate the differences between various types of broadcasts. The primary criticism of this approach, however, is that it makes comparison with other languages unreliable. Languages that have on average longer words in which inflection and derivation are rich, would appear to have slower speech rates than languages for which word length is typically shorter and grammatical markers are counted as words as well.

A frequently used measure for assessing the rate of speech counts the number of syllables per second, or the number of syllables divided by the length of the utterance. While this type of measure avoids the pitfalls associated with language specific word length differences, languages also differ on syllable structure such that syllable duration varies as well. Furthermore, comparison between studies can be difficult because some include pauses, and some exclude them. The number of syllables divided by the total speaking time of the task is used for a number of purposes to describe speech behaviour in what are usually, although not always, non-clinical, discourse related studies (Kehrein, 2003;

Kowal, O'Connell, & Sabin, 1975; Oliveira, 2000; Reilly, Cannizzaro, Harel, & Snyder, 2004; van Heuvan & van Zanten, 2005).

Some researchers employ a measure of segments per second or per minute to arrive at an articulation rate. This type of measure is often, although not always used in clinical studies examining pathological speech. In general, the duration of the utterance excludes pauses so that the focus is only on the actual speech itself (Crystal & House, 1990; Dankovicova, 1997; Koreman, 2006; Miller, Grosjean, & Lomanto, 1984; Robb, Maclagan, & Chen, 2004; Tsao & Weismer, 1997; Tsao, Weismer, & Iqbal, 2006; Walker & Archibald, 2006).

In this study, speaking rate is handled in two ways, defined below:

Speech Rate: syllables per second, including pauses. The number of syllables produced in the suspect's response is counted, and the total is divided by the duration in seconds of the entire utterance.

Articulation Rate: speech segments per second, excluding pauses. The speech segments in the suspect's response are counted, and the total is divided by the duration of the utterance in seconds, minus the total duration of pauses, if any.

These two measures are widely used in a great deal of work on prosody, as reported in the literature (see section 4.2 and 4.3). While they are not always studied in tandem, they will be in this study for three reasons. First, in this study, speech rate and articulation rate are equally important to examine because both

measures tell different but complementary stories about speech. Speech rate, as a measure of syllables per second throughout the duration of the entire utterance, is included because it takes into account the amount of pausing that occurs within a speaker's utterance. Articulation rate, on the other hand, provides a measure of how slowly or rapidly the speaker is moving the articulators to produce the vocalized portion of the utterance. And third, the two response categories that have exceptionally short responses, affirmative and negative, are in many cases, comprised of a single monosyllabic word: *yeah, yes, no, nope,* and other similarly short utterances such as *uh huh* and *uh uh*. A speech rate measure based on syllables per second would not be useful in that case, so articulation rate is more telling. Furthermore, in longer responses, such as relevant and irrelevant, measuring only the articulation rate would exclude any information about the amount of pausing that takes place within the response.[24] Third and most simply, I want this study to report results that might be informative to more types of research in the future, and since there does not appear to be a consensus on how rates of speaking are to be measured, this study sits on a bit of a methodological fence. Taking into account the above, not all response types lend themselves appropriately to both measures of speaking rate, and some are suitably measured both ways. Table 4.1 summarizes the design of the study with respect to the type of measurement used for the response types.

---

[24] The relationship between speaking rates and pause frequencies and durations is well studied. The interaction between these features with respect to the present corpus will be dealt with in Chapter 6.

Table 4.1   Design of the study on temporal phenomena

Temporal Phenomena	Measure	Response Types
Speech rate	syllables per second (including pauses)	relevant irrelevant (Type FT only) confession
Articulation rate	segments per second (excluding pauses)	affirmative negative relevant irrelevant (Type FT only) confession

Response types for which there is sufficient data are treated statistically. These include: Affirmative, Negative, Relevant and Irrelevant. Confessions are handled on a more qualitative basis, as has been the case in the previous chapter on pause. Table 4.2 shows the number of responses to be analysed in this portion of the study.

Table 4.2   Distribution of response types analysed for speaking rate by suspect

| | FT (first-timers) | | | | | | | | |
	FT-1	FT-2	FT-3	FT-4	FT-5	FT-6	FT-7	FT-8	Total
Affirmative	27	9	22	22	31	11	10	18	**150**
Negative	21	16	6	23	19	9	19	20	**133**
Relevant	24	20	34	22	26	12	18	18	**174**
Irrelevant	19	12		20	16	20	16	15	**118**
Confession		5		4		7			**16**
Total	**91**	**62**	**62**	**91**	**92**	**59**	**63**	**71**	**591**

| | RO (repeat offenders) | | | | | | | | |
	RO-1	RO-2	RO-3	RO-4	RO-5	RO-6	RO-7	RO-8	
Affirmative	12	18	10	41	23	11	12	19	**146**
Negative	21	14	9	18	12	11	11	10	**106**
Relevant	28	14	23	18	21	7	26	14	**151**
Confession		2		5	2				**9**
Total	**61**	**48**	**42**	**82**	**58**	**29**	**49**	**43**	**412**

4.4   Results

In the following two subsections, the results of the study on speech rate

(4.4.1) and articulation rate (4.4.2) are presented.

4.4.1   Speech rate

The speech rate of suspects' responses varies, both within and between

groups. A number of representative examples from one suspect of each type are

shown below. Recall that only responses of significant length (i.e., relevant,

irrelevant, and confessions) are measured in syllables per second and that only

FT suspects produced irrelevant responses.

Table 4.3   Speech rate samples from Suspect FT-1

Suspect FT-1	Response Type	Syll/s
I: How many beers did you have there? S: Probably about three there.	Relevant	3.85
I: What time did you leave the BAR? S: Oh I dunoo probly at uh, last time I looked it was probably after nine or something.	Relevant	3.42
I: Ok, but you remember there were people there? Or were there people there but you don't remember their names? S: **Relevant→** There were probably people there and I don't I never got introduced maybe, never knew anybody. **Irrelevant→** WITNESS usually has people coming and going through the place everyday. When I'm working they're gone, there's a big crowd of people in and out all the time.	Irrelevant	5.71

Table 4.4   Speech rate samples from Suspect RO-7

Suspect RO-7	Response Type	Syll/s
I: How were you helping him? S: Like he said he needed some money cause he had a kid and like he's got a bad habit (crack) and he didn't want anybody else to know about his bad habit.	Relevant	5.62
I: What gun? He had a gun? S: He had his gun in his brief case and he was like holding it all the time.	Relevant	6.13
I: How much a week are you spending on crack? S: Thousand bucks a week.	Relevant	5.56

These examples are typical of what was found in the corpus with respect to speech rate for one FT and one RO Suspect for these response categories. Table 4.5 shows the ranges and averages of all suspects' speech rate values by response type.

Table 4.5   Speech rate values for each suspect by response type

Suspect	Relevant				Irrelevant			
	min	max	mean	s.d.	min	max	mean	s.d.
FT-1	2.40	7.06	4.32	1.18	4.12	7.14	5.17	.93
FT-2	1.77	6.00	2.83	1.11	2.95	6.05	3.72	1.13
FT-3	2.47	7.78	4.15	1.25				
FT-4	1.21	5.00	3.31	.91	3.02	5.00	3.75	.82
FT-5	3.08	7.27	4.26	.97	3.33	4.69	4.02	.48
FT-6	1.48	3.93	2.70	.78	3.27	5.28	4.19	.78
FT-7	2.39	6.54	3.74	1.00	3.12	5.04	4.41	.55
FT-8	1.52	4.30	2.63	.79	3.27	4.59	3.42	.45
mean	2.04	5.99	3.49	1.00	3.28	5.40	4.17	.73
s.d.	.64	1.43	2.63	.79	.39	.91	0.58	.25
RO-1	2.40	6.67	4.57	1.21				
RO-2	3.08	6.57	3.98	1.04				
RO-3	2.97	8.67	5.97	1.52				
RO-4	2.70	7.50	5.00	1.32				
RO-5	2.27	8.64	4.86	1.66				
RO-6	3.53	6.83	4.95	1.16				
RO-7	3.14	7.67	5.22	1.24				
RO-8	3.75	6.82	5.38	1.20				
mean	2.98	7.42	4.97	1.29				
s.d.	.51	.85	.54	.20				

Table 4.5 shows the ranges and average speech rate values for each suspect. Speech rates vary quite widely for the majority of individuals, with a

difference of approximately 4 or 5 syllables per second between the highest and lowest for any one suspect in relevant responses. The smallest range in relevant response speech rates is shown by suspect FT-6, with a difference of 2.45 syll/sec between his minimum speech rate and his maximum. The greatest range is shown by suspect RO-5, with a difference of 6.37 syll/sec between his lowest and his highest rates.

Among the irrelevant responses, the difference between minimums and maximums is narrower, possibly because irrelevant responses tend to be produced more toward the upper ranges of natural speech rates for most speakers. The smallest range is shown by suspect FT-8, with a difference of 1.32 syll/sec between his minimum and maximum speech rates, and the largest range is shown by suspect FT-2, with a difference of 3.1 syll/sec between his minimum and maximum speech rates. The figure below illustrates the speech rate averages for relevant responses for both groups, and for irrelevant responses for FT suspects.

Figure 4.1   **Speech rate for relevant responses for both groups, and irrelevant responses for FT Suspects**

Error bars indicate one standard deviation.
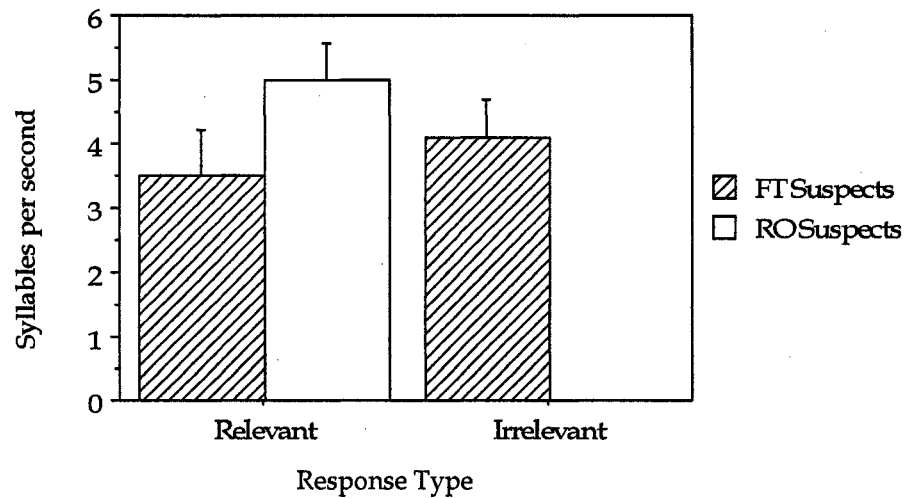
A paired t-test on the data from the seven FT suspects who produced both relevant and irrelevant responses showed a significant difference in speech rate, $t(6)=-3.532$, $p=.0123$. In addition, the two suspect groups differ in their speech rates for relevant responses. When the relevant response speech rate values for both groups were submitted to an independent samples t-test, RO suspects' speech rates were found to be significantly higher than those of FT suspects', $t(14)=-4.582$, $p=.0004$.

For those who produced them, confessions were uttered with relatively slower overall speech rates, compared to relevant responses. The table below shows three examples of confessions with their speech rates.

Table 4.6   Speech rate samples from confessions

Suspect	Confession	Syll/s
FT-6	Um I was I had my hand on her leg and I was um moving it closer to her vagina and um and then as it went on I started like on the outside I was just touching it and putting my hand back on her leg and I actually touched it for like a couple minutes and then i took my hand off and realized and then I grabbed my drink and went back to my room.	2.31
RO-5	All I'll say is, I own this one. It was more like I was testing myself.('one'=one in a series of sexual assaults)	1.83
RO-7	Well, I'd say it was five times, probably not more than ten times ('it'=exposing suspect's genitals to children in public places)	3.14

In Figure 4.2 the speech rates of confessions are tabulated for each suspect and for comparison, mean speech rate for relevant responses is included.

Figure 4.2   Speech rates of suspects' confessions

Solid dots indicate individual confession speech rates, while empty dots indicate speech rates of suspects' relevant responses.
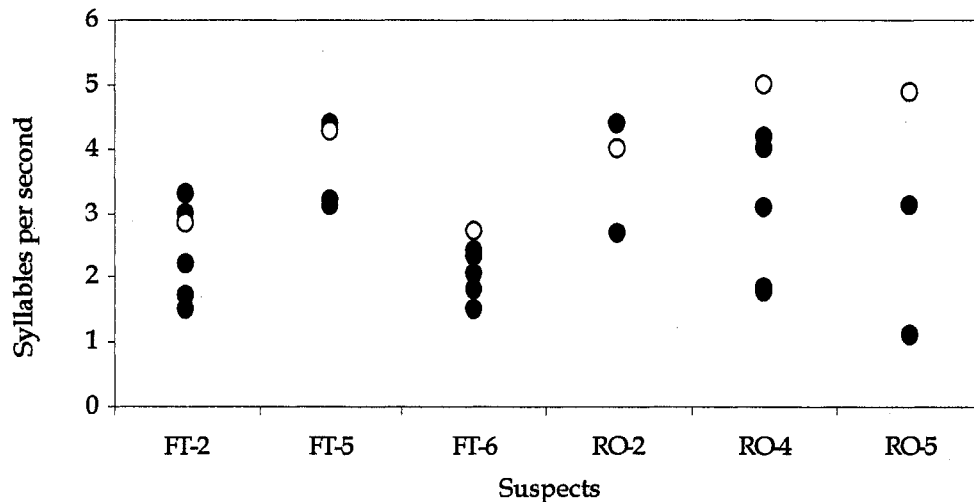
Figure 4.2 shows that suspects' speech rates in confessions are, in general, low compared to the mean of these individuals' relevant response speech rates. For suspects FT-6, RO-4 and RO-5, confession speech rates are all lower than their respective relevant speech rates. For suspects FT-2, FT-5, confession speech rates are mostly lower than suspects' respective relevant speech rates. Suspect RO-2 produced only two confession statements, so it is not possible to gauge where his confession speech rate ranks compared to the speech rate of his relevant responses.

4.4.2   Articulation rate

In this section, I look at the articulation rates measured in all responses to find that articulation rate varies according to both group and some response

types. In addition to affirmative and negative responses, examples from the previous section on speech rate are presented again below for relevant and irrelevant responses with articulation rates reported.

Table 4.7   Articulation rate samples from Suspect FT-1

Suspect FT-1	Response Type	Seg/ sec
I: You were at THE BAR that day? S: Yeah.	Affirmative	3.64
I: It's just, the opportunity was there and you took it. S: I did yeah.	Affirmative	7.41
I: Do you know VICTIM'S NAME? S: No I don't.	Negative	9.86
I: Did you ejaculate? S: No.	Negative	6.06
I: How many beers did you have there? S: Probably about three there.	Relevant	12.12
I: What time did you leave THE BAR? S: Oh I dunoo probly at uh, last time I looked it was probably after nine or something.	Relevant	9.16
I: You just told me you had sex with her. S: Maybe she raped me! (contrastive stress on *she, me*)	Irrelevant	10.00
I: Ok, but you remember there were people there? Or were there people there but you don't remember their names? S: **Relevant→** There were probably people there and I don't I never got introduced maybe, never knew anybody. **Irrelevant→** WITNESS usually has people coming and going through the place everyday. When I'm working they're gone, there's a big crowd of people in and out all the time.	Irrelevant	14.12

Table 4.8   Articulation rate samples from Suspect RO-7

Suspect RO-7	Response Type	Seg/ sec
I: Do you know where the watches are now? S: Yeah.	Affirmative	13.33
I: And when was it that you discovered that the Rolexes were fake? Was it in that apartment that night? S: Yeah.	Affirmative	11.11
I: Was anybody doing something stupid like that, intimidating people? S: No	Negative	8.96
I : Did you in any way just horse around and grab him even at any point? S: I never touched him.	Negative	10.92
I: How were you helping him? S: Like he said he needed some money cause he had a kid and like he's got a bad habit (drug use) and he didn't want anybody else to know about his bad habit.	Relevant	15.73
I: What gun? He had a gun? S: He had his gun in his brief case and he was like holding it all the time.	Relevant	15.41
I: How much a week are you spending on crack? S: Thousand bucks a week.	Relevant	18.92

Before describing the findings of within- and between-group comparisons,

the suspects' individual values are shown in order to display the range of

articulation rates. Table 4.9 reports these value

134

Table 4.9   Articulation rate values for each suspect by response type

Suspect	Affirmative				Negative				Relevant				Irrelevant			
	min	max	s.d.	mean	min	max	s.d.	mean	min	max	s.d.	mean	min	max	mean	s.d.
FT-1	3.64	14.47	2.94	7.84	4.00	9.86	1.98	6.05	6.65	16.57	2.24	10.72	10.01	14.57	12.65	1.39
FT-2	7.29	12.12	1.83	9.48	6.67	13.54	2.25	8.85	5.56	13.50	2.08	10.03	10.90	19.05	14.47	2.80
FT-3	4.17	13.95	3.77	10.11	5.96	12.81	2.46	10.18	6.20	13.70	1.69	10.97				
FT-4	5.77	16.22	3.25	11.14	9.52	13.58	2.45	11.86	7.45	15.56	1.68	11.27	5.94	11.21	10.41	1.72
FT-5	4.88	16.18	3.55	10.99	8.33	13.48	1.53	10.78	8.33	17.52	2.22	10.8	11.79	15.34	13.54	1.62
FT-6	4.76	9.09	1.23	6.87	4.44	12.50	2.52	7.43	5.97	13.54	2.29	9.77	7.95	13.69	10.57	2.05
FT-7	5.77	18.42	5.21	11.48	8.44	11.90	1.41	10.27	8.78	14.74	1.95	11.13	9.71	15.49	13.25	1.92
FT-8	7.41	13.73	2.89	8.53	10.00	18.18	3.38	15.83	6.32	15.25	2.51	9.68	11.38	13.96	12.48	0.98
mean	5.46	14.27	3.08	9.56	7.17	13.23	2.25	10.16	6.91	15.05	2.08	10.55	9.67	14.76	12.48	1.78
s.d.	1.37	2.84	1.21	1.68	2.26	2.35	.62	2.96	1.16	1.48	.29	.63	2.08	2.37	1.51	.57
RO-1	5.13	13.73	2.45	7.54	3.28	10.64	2.43	6.52	10.71	16.65	1.69	13.18				
RO-2	6.25	10.98	2.20	8.66	6.25	12.00	2.04	8.58	10.06	17.72	1.94	11.99				
RO-3	8.51	15.63	2.30	11.08	6.83	16.95	3.03	11.4	9.49	20.31	2.52	15.15				
RO-4	2.08	16.34	3.83	9.19	3.92	9.93	2.14	6.43	9.62	16.67	2.28	12.18				
RO-5	6.67	23.08	5.16	16.05	5.71	14.58	3.43	9.90	7.41	18.52	2.84	12.23				
RO-6	5.98	16.42	2.18	11.21	2.35	11.76	2.73	8.35	10.74	15.52	1.63	13.48				
RO-7	6.25	13.82	2.21	9.90	4.65	16.90	4.19	10.04	11.54	18.92	2.00	14.87				
RO-8	8.33	13.24	1.67	10.39	4.76	18.87	4.53	8.71	9.80	17.60	1.99	13.34				
mean	6.15	15.41	2.75	10.50	4.72	13.95	3.07	8.74	9.92	17.74	2.11	13.30				
s.d.	2.01	3.59	1.16	2.56	1.52	3.34	.92	1.71	1.23	1.51	.41	1.20				

135

Suspects show a tremendously wide range of articulation rates, which is unsurprising. Speaking rates vary considerably from speaker to speaker, and from utterance to utterance, within speakers, often depending on the length of the utterance (Fonagy & Magdics, 1960; Goldman-Eisler, 1956b, 1968). In spontaneously produced speech, taking into account speakers' individual speaking styles, it is expected that a "yeah" or a "no" might be elongated if it is the only word in an utterance, resulting in a slower articulation rate. Longer utterances, on the other hand, such as relevant responses, would contain grammatical words given little prosodic prominence, therefore having the effect of "speeding up" the utterance, even with slower articulation afforded to longer content words. Suspects FT-1, FT-5 and RO-1 in particular display this type of pattern.

Within response categories, suspects show a great deal of individual variation, some more than others, particularly among the types of responses that are short, i.e., affirmative and negative. Suspect FT-7's affirmative responses, for example, range from 5.77 segments per second to a very rapid 18.42. Suspect RO-8's affirmative responses, on the other hand showed less variation, with a minimum of 8.33, and a maximum of 13.24. For approximately half the suspects in each group there was a large difference between the minimum and maximum values, and furthermore, the standard deviations are quite high, indicating a great degree of variation. This trend was particularly strong in affirmative and

negative responses. Figure 4.3 shows the means articulation rates for all responses for both groups.

Figure 4.3 Articulation rates for FT and RO suspects by response type
Error bars show one standard deviation.
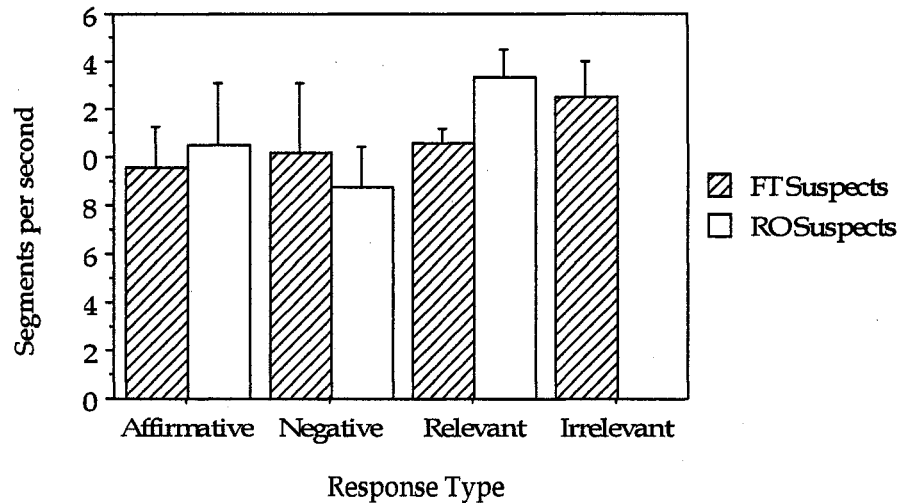
Among the relevant and irrelevant responses, the values for most suspects show considerably less variability than the affirmative and negative responses, and the standard deviations are generally lower, particularly for FTs' relevant responses. A mixed design ANOVA[25] shows an interaction between group and response type $F(2,28)=5.858$, $p=.0075$. Post hoc Tukey tests with Bonferroni adjustment yielded a significant difference (p<.0167) between affirmative and relevant, and negative and relevant, where RO suspects' articulation rates are higher. Considering the fact that affirmative and negative responses are so brief,

---

[25] Irrelevant is excluded from this ANOVA because only FT suspects produced this response type.

and relevant responses may be up to several clauses, articulation rate may not be a particularly telling result. Nonetheless, in the interest of being thorough, the comparison is reported.

In order to include FT's irrelevant responses in the analysis, the articulation rates were submitted to a one-way repeated measures ANOVA together with corresponding relevant response articulation rates. The effect of response type is significant, $F(3,15)=6.865$, $p=.0039$. Post hoc Tukey tests with Bonferroni adjustment show that irrelevant articulation rates are significantly higher than both affirmative and relevant. No significant difference is found between affirmative, negative and relevant responses for articulation rate in FT suspects.

For the six suspects who produced confessions, the articulation rates show considerable variation, although are on the whole, lower compared to relevant responses. Using the same samples for confession speech rates, the articulation rates are given below.

Table 4.10   Articulation rate samples from confessions

Suspect	Confession	Seg/s
FT-6	Um I was I had my hand on her leg and I was um moving it closer to her vagina and um and then as it went on I started like on the outside I was just touching it and putting my hand back on her leg and I actually touched it for like a couple minutes and then i took my hand off and realized and then I grabbed my drink and went back to my room.	9.09
RO-5	All I'll say is, I own this one. It was more like I was testing myself.('one'=one in a series of sexual assaults)	6.75
RO-7	Well, I'd say it was five times, probably not more than ten times ('it'=exposing suspect's genitals to children in public places)	12.35

In the following figure, individual confession articulation rates for each suspect are shown. For purposes of comparison, suspects' relevant articulation rate means are included.

Figure 4.4   Articulation rate values for suspects' confessions

Solid dots indicate individual confession articulation rates, while empty dots indicate articulation rates of suspects' relevant responses.
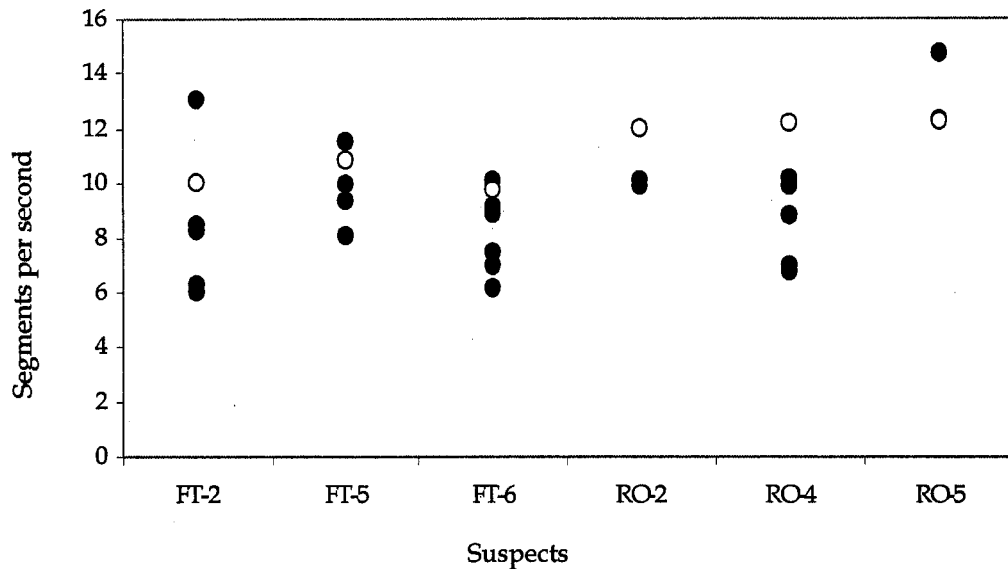
For all but one suspect, RO-5, there is a strong tendency for confessions to be produced at lower articulation rates than the average for a relevant response. Suspects RO-2 and RO-4 show articulation rates uniformly below the average for their respective relevant responses, and the articulation rates of Suspects FT-2, FT-5, and FT-6 are mostly below the average for their own relevant responses. Interestingly, the results of articulation rates for confessions parallel those of speech rates in comparison to relevant response means. See Figure 4.2, which reports speech rates along with means for relevant responses.

4.5 Tempo profiles and conclusion

The preceding sections report the results of studies on speech and articulation rates. In this section, I discuss the results in terms of how FT and RO's speech and articulation rates differ, and how speech and articulation rate correlate with one another within response types, and I draw comparisons to the findings of others in the literature. The chapter ends with a summary of the answers to the research questions listed in section 4.1.

4.5.1 Tempo profiles and discussion

Once again, a subjective visual representation is given, this time bringing together the statistically significant differences and strong tendencies from the tempo results based on values shown in Figure 4.1 and Figure 4.3. Large cones represent high values, small cones, low values, mid-sized cones represent medium range values, and flat discs represent values that vary too widely to safely summarise as high, mid or low. The pair of figures below show tempo profiles for the responses of both groups.

Figure 4.5   FT tempo features by response type, showing relative values
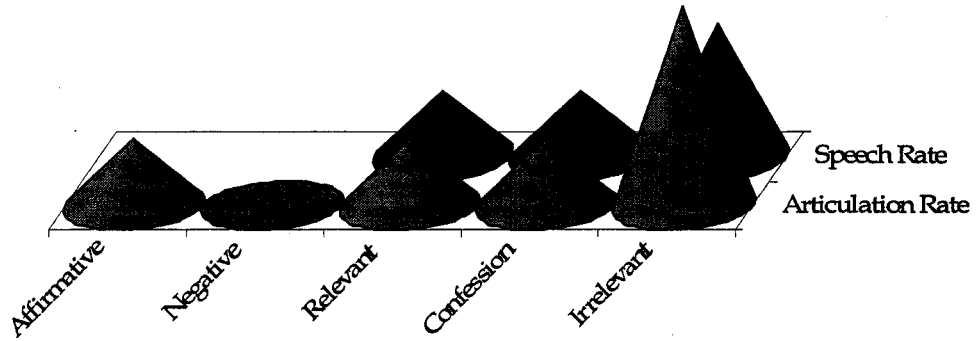
Figure 4.6   RO tempo features by response type, showing relative values
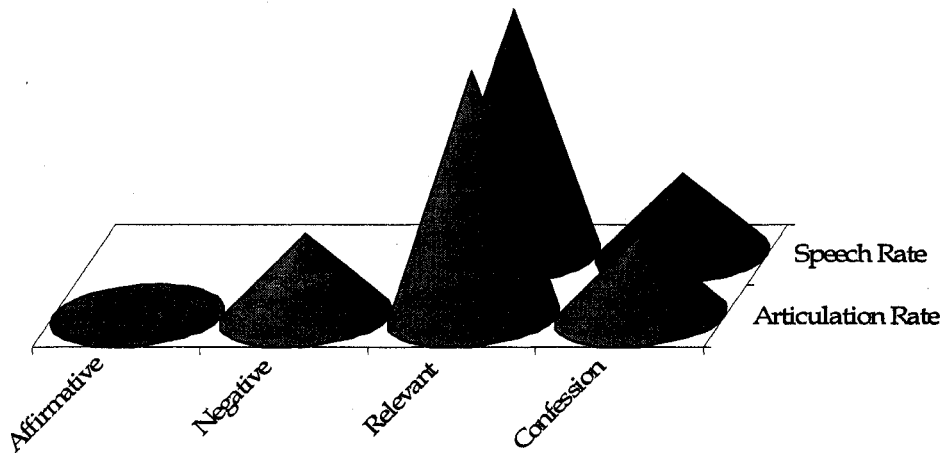
The following table summarizes the statistical results of comparisons.

Table 4.11   Summary of significant and non-significant differences by group and response type

Feature	Comparison		Significance	Test
	Group	Response Type		
Speech rate	FT/RO	Rel	<.05	t^a
	FT	Rel/Irrel	n.s.	t^b
Articulation rate	FT/RO	Aff/Rel	<.05	F^c
		Rel/Neg	<.05	F^c
		Neg/Aff	n.s.	F^c

[a] independent samples; [b] paired; [c] mixed design

FT suspects maintained fairly consistent low speech and articulation rates throughout their relevant responses and confessions, with the exception of the irrelevant information they produced, which came at higher rates according to both measures. RO suspects were also consistently low in negative articulation rates, and relevant responses were produced much more rapidly. The caveat previously mentioned in section 4.4.2 is repeated here, that the articulation rates of responses as short as those in the affirmative and negative categories may not be as comparable with relevant and irrelevant categories owing to the difference in response length, even with all pauses excluded. It comes as no surprise then that the response types which it was not possible to profile were affirmative (for RO) and negative (for FT). Nevertheless, all values are reported.

That FT suspects' relevant responses are slower overall than RO suspects is suggestive of two possible effects on suspects. First, FT suspects' may be more

apprehensive than RO suspects who are familiar with the interview process. This notion is supported by the fact that when discussing irrelevant information, which is, from the suspects' point of view, safe information to reveal, FT suspects' speech rates increase, close to the rates at which RO suspects produce relevant information. Second, the results suggest that RO suspects are somewhat less apprehensive given their familiarity with being interviewed by police, and their speech rates are even somewhat higher than the average speech rate of 3.66 syll/s for North American English (Robb, Maclagan, & Chen, 2004)[26]. It is possible to speculate that, considering RO suspects' familiarity with being interviewed by police, they have thought in advance how they might respond to questions. In fact, their responses of relevant information are closer to standard speech rates in English, than are FT suspects' which are lower.

Returning briefly to the rapid rates that accompany irrelevant information, we note that previous studies have shown that parenthetical information is produced at high rates (Grosz & Hirschberg, 1992; Hirschberg & Grosz, 1992). The irrelevant information produced by FT suspects, often along with relevant information, suggests that suspects are signalling a contextual difference between pertinent and non-pertinent. Why FT suspects produce this type of information and RO suspects do not is unknown at this time. I offer some speculation in Chapter 6.

---

[26] The study by Robb et al. combined the results of several speaking rate studies to arrive at an average 220 syllables per minute rate. This figure is converted to syllables per second, to make comparison possible with the values in this study.

The correlation between speech and articulation rate for lengthy responses (i.e., relevant and irrelevant), is reported in Table 4.12, followed by a discussion of the interaction between pause and tempo.

Table 4.12   Pearson correlation coefficients for Relevant and Irrelevant tempo features

	Relevant		Irrelevant
	FT	RO	FT
	r	r	r
Speech rate/Artic rate	.793, $p<.05$	.737, $p<.05$	.029, $p=.9539$

There is a moderate to strong correlation between speech and articulation rates for both groups for relevant responses. No correlation is found between both tempo measurements for irrelevant responses produced by FT suspects. That there is no correlation suggests that pause affects the rate of irrelevant responses. When taking into account pauses, as speech rate values do, irrelevant responses and relevant responses alike have high pause ratios which prevents any correlation between speech and articulation rate. Irrelevant responses are spoken more rapidly, but the pause ratio for irrelevant responses varies to a greater degree than the speaking rates. In other words, an irrelevant response can be uttered rapidly and have a little or a lot of pausing, compared to a relevant response, which has consistently a moderate amount of pausing.

The speech and articulation rates of confessions are low in relation to the same rates for relevant responses for both groups. It is interesting to note that

throughout the literature, sadness is consistently associated with lower speaking rates (Banse & Scherer, 1996; Johnstone & Scherer, 2000; Kehrein, 2003; Sobin & Alpert, 1999). Assuming the suspect is experiencing[27] sadness or remorse (either for his victim or himself), slower speech rates might reasonably be expected to accompany confessions.

4.5.2 Conclusion

The stand-out feature of speaking rate in this study is the large ranges, both within- and between-groups. Despite this degree of variation, however, some uniformity in tempo features has been found. The research questions are repeated below, and followed by a brief summary of results.

Does response type and suspects' FT or RO status affect:

1) the speech rate at which the suspect produces the response?

2) the articulation rate at which the suspect produces the response?

The answer to both of these questions is yes, partially. FT suspects' speech and articulation rates differ significantly for relevant and irrelevant responses, with the former being produced at lower rates than the latter. Also, FT and ROs' relevant responses differ significantly, with RO suspects producing relevant responses at faster speech and articulation rates than FT suspects. Confession

---

[27] It is plausible that suspects' remorse or other emotions could be feigned, and Reilly et al. (2004) conclude that manipulating speech rate is typical of speakers feigning depression. There is no reliable method, however, to determine genuine emotion for this data set.

speech rates for both groups are consistently lower than the speech rates for relevant responses.

No differences were found in the articulation rates of affirmative and negative responses for either group, which is unsurprising, given their relative brevity. Among FT suspects, relevant responses were only slightly faster than affirmative and negative, although the results are significant, and among RO, relevant responses were produced considerably faster than affirmative and negative. Confessions for both groups are lower overall than relevant responses.

As was shown in Chapter Three with pause features, not all response types come with a unique set of tempo features. However, some significant differences do emerge.

5 PITCH

All feelings have their particular tone of voice.
-Francois de la Rochefoucauld

In this chapter, three dimensions of suspects' pitch behaviour during police interviews are examined. First, the average pitch of suspects' responses is examined to determine whether suspects' average pitch values differ significantly by response type, and by group. Second, the variability of pitch is explored, to find an overall difference between the groups wherein RO suspects' variability is much greater than that of FT suspects. However, there is no significant difference between response types. Finally, I consider the pitch range, taken as the difference between individual speakers' pitch maxima and minima to conclude that overall, FT suspects' pitch ranges more than does RO suspects'. However as with variability, there is no significant difference between response categories.

5.1   Introduction

As far back as the 17th century, the dynamic highs and lows of speakers' voices during speech have attracted scholarly interest. At the time, Walker (1781) introduced his 'theory of inflexions' consisting of five 'slides' — rising, falling, fall-

rise, rise-fall and monotone. Since then, there has been considerable work on the contribution that intonation makes to the meaning of the spoken message.

When listeners describe the ups and downs of a speaker's voice, they are referring to pitch. Changes in pitch signal a wealth of information about the speaker's message, and about the speaker himself or herself. Pitch is dynamic in terms of both interspeaker and intraspeaker ranges, and changes in pitch have been linked to numerous causes, all of which also determine variations in other prosodic phenomena such as pausing and tempo, addressed in the previous two chapters. Among the more well-studied factors surrounding pitch phenomena are emotional state (Baum, 1998; Johnstone & Scherer, 2000; Kehrein, 2003; Laukka, Juslin, & Bresin, 2007; Nilsonne, 1987, 1988; Scherer, 2003; Scherer, Banse, & Wallbot, 2001; Scherer & Orshinksy, 1977); discourse segmentation (Ford & Thompson, 1996; Grosz & Hirschberg, 1992; Hirschberg & Grosz, 1992; Hirst, 2005; Ji, 2002; Oliveira, 2000; Schleef, 2003; Selting, 1992, 1996; Swerts, 1997; Wennerstrom, 2001), and turn taking (Auer, 1996; Caspers, 2003; Ford & Thompson, 1996; Local, Kelly, & Wells, 1986; Schaffer, 1983; Schegloff, 1996; Wells & Peppé, 1996; Wennerstrom, 2001), where intonation is, in many cases, considered in conjunction with syntactic and pragmatic boundaries.

In this chapter, suspects' pitch behaviour as it occurs during questioning by police is examined in terms of average pitch, pitch variability, and pitch range. The goal of this part of the study, parallel to the previous two data

chapters, is not to identify the possible causes that underlie suspects' pitch behaviour. Rather it is to closely examine pitch values of different response types, comparing the two groups. The following research questions are answered in this chapter:

Do response type and suspects' FT (first-timer) or RO (repeat offender) status affect:

a) the average pitch of responses?

b) the variability of pitch in responses?

c) the pitch range of responses?

Examples of the manner in which pitch is dealt with are shown below. An example of the measured pitch values for two relevant responses are given[28]. Bracketed spans of speech indicate one intonation unit (IU). The definition of an IU is provided in section 5.3, along with the rationale for reporting suspects' pitch values in semitones.

---

[28] Variability on individual responses is not shown here. Rather, variability will be taken on the whole of each suspects' standard deviations on pitch averages for each response category. See section 5.4.2

Table 5.1   Pitch values of a relevant response from Suspect FT-3

Suspect FT-3			Relevant
I: So he's already there, or he's coming to the house. S: [1 No he's outside] [2 waiting for me] [3 He's on the side walk.]			
	IU 1	IU 2	IU 3
Minimum	0.7 st	-2.6 st	-1.4 st
Maximum	4.7 st	4.0 st	3.3 st
Mean	3.2 st	0.9 st	2.0 st
Range	5.4 st	6.6 st	4.7 st

Table 5.2   Pitch values of a relevant response from Suspect RO-5

Suspect RO-5		Relevant
I: What gun? He had a gun? S: [1 He had a gun in his briefcase,] [2 and he was holding it all the time.]		
	IU 1	IU 2
Minimum	-0.7 st	-2.6 st
Maximum	7.8 st	5.5 st
Mean	3.7 st	0.7 st
Range	8.5 st	8.1 st

Suspect FT-3 produced a relevant response consisting of three intonation units. His pitch went as low as -2.6 semitones in the second intonation unit, and at his highest, was 4.7 semitones. His ranges were within approximately one semitone of 5.5 semitones. Suspect RO-5's lowest pitch value also reached -2.6, and his highest, 7.8 semitones. His range was considerably greater in his response than suspect FT-3's, reaching 8.5 and 8.1 semitones in his intonation units. These two examples illustrate how pitch values were collected. Further

examples of suspects' pitch behaviour are given in the results subsections in section 5.4.

The next section of this chapter is composed of a literature review (section 5.2) focussing on previous work which has addressed the intersections of pitch and discourse, and pitch and emotion. Following that is a review of studies on police interviews and deception in which pitch has been explored. The methodology is discussed in section 5.3. The results of the studies on average pitch, pitch variability, and pitch range are reported in section 5.4. Profiles of pitch features for both groups across response types, and concluding remarks close the chapter in section 5.5 .

5.2   Pitch in the literature

The intonation of a speaker's voice reveals information about a speaker's emotional state and his or her pragmatic and conversational intensions. The purpose of this section is to review related literature on the causes of pitch variation in speakers' speech, from the broad spectrum of discourse and emotion factors, and studies on deception.

5.2.1   Pitch as a prosodic feature in discourse and emotion

Changes in pitch, pitch coupled with accent or stress, and boundaries between pitch contours are shown throughout the literature to play multiple roles in the organization of discourse, and in modifying the speaker's message to

convey pragmatic and contextual information, in addition to signalling to hearers when the speaker's turn is about to end. At the same time, pitch features are a window into the speaker's emotional state, and pitch helps to convey attitudes and intents that accompany the message. In this section, I survey studies that examine the role of pitch in discourse and pitch as an emotional cue.

An array of studies by Lehiste (1975, 1976, 1979, 1982) have shown that changes in pitch serve to partition discourse segments. In particular, she found that a new conversational topic is introduced with a wider pitch range, which narrows as the utterance comes to a close. Many others have echoed this finding, noting that utterances which on a global discourse level are the equivalent of 'paragraph initial' have the widest discourses ranges (Crystal, 1969; Swerts & Geluykens, 1994; van Donzel, 1999). Grosz & Hirschberg (1992) and Hirschberg & Grosz (1992) recognized that the divergence in pitch ranges of neighbouring intonation units frequently results in a topic change.

The notion of comparing neighbouring boundary tones was taken up earlier by Brown, Currie, & Kenworthy (1980), who concluded that low boundary tones signal the speaker's end to the topic, while non-low, i.e., rising or level, indicate that the speaker intends to hold the floor and continue the topic. Swerts & Geluykens (1994) concluded the same, although extending the discourse outcome of topic continuation boundary tones to distinguish between the speaker's intent to produce a little more speech, or a lot more speech.

153

The opening and closing of topics through intonation cues leads naturally into discussion of intonation as a cue to turn-taking. A number of studies have found language-specific intonation contours that signal turn-yielding in conversation. Speakers of Tyneside English (Local, Kelly, & Wells, 1986) and Ulster English (Wells & Peppé, 1996), for example, indicate the end of a turn with a rising pitch. Malay speakers, on the other hand, signal turn completion with falling pitch and they sustain turns with level or rising pitch (Zuraidah & Knowles, 2006). In London Jamaican English, interlocutors determine turn ends based on the contour that accompanies the tag question 'you know' (Local, Wells, & Sebba, 1985).

Ford & Thompson (1996) undertook a corpus study on American English, in which transcribed conversations were coded for prosodic, syntactic, and pragmatic completion points, in order to examine the correlation of completion points. They found that where the three converge, a turn transition (i.e., the shift to a new speaker) is most likely to occur.

In addition to revealing speakers' conversational intensions, pitch serves an extra-linguistic function in conveying speakers' emotional state, in conjunction with other prosodic cues such as pause, loudness, and tempo. Of all prosodic cues, pitch has been called the most perceptually prominent of the suprasegmentals (Banse & Scherer, 1996), and hence hearers are sensitive to the emotional content conveyed by speakers' vocal highs and lows. Emotional

categories themselves, e.g., fear, anger, sadness, joy, shame, etc, are shown to be accompanied by relatively consistent fundamental frequency (F0) patterns, confirmed by a number of lab-based studies[29] employing actors who portray emotional states. Sadness, for example, has been widely demonstrated to be accompanied by low average pitch and narrow pitch range (Banse & Scherer, 1996; Johnstone & Scherer, 2000; Scherer, 2003; Thompson & Balkwill, 2006). Many of the same studies concluded that fear is accompanied by high pitch values (Banse & Scherer, 1996; Johnstone & Scherer, 2000; Scherer, 2003), although at least one study found the opposite (Thompson & Balkwill, 2006). A survey of the same studies for anger and joy, however, points to inconclusive results. See Table 5.3 for a summary of these findings.

Table 5.3   Broad summary of some findings on emotion and pitch[30]

	Fear	Anger	Sadness	Joy
Average pitch	high[a,b,d],low[c]	low[a,b,c],high[d]	low[a,b,c,d]	low[a], high[b,c,d]
Pitch range	high[b,d]	high[d],var[b],low[c]	low[b,c,d]	high[b],low[c,d]
Pitch variability	high[a,d]	high[a,d]	low[a,d]	high[a,d]

[a]Sobin & Alpert (2002); [b] Scherer (2003); [c]Thompson and Balkwill (2006), [d]Johnstone & Scherer (2000).

---

[29] See chapter 3 for a discussion on the limitations of laboratory studies where emotion is simulated.
[30] Many other emotional states are examined in some studies. Only those emotions that appear consistently across studies are included in this summary. Other emotional states include boredom, anxiety, pride, disgust, interest, despair and distinctions are made between hot anger and cold anger.

In another study on emotion and prosody, Breitenstein, van Lancker, &

Daum (2001) presented hearers with stimuli in which pitch variation was

manipulated. Results showed that utterances with reduced pitch variation were

categorized as sad or neutral, while greater pitch variation was heard as

frightened, angry and happy.

Besides there being at least some uniformity among some of the emotional

states with respect to F0 patterns, it has also been suggested that with heightened

arousal come higher F0 values, and conversely, with subdued states come lower

F0 values. Hot and cold anger, for example, tend to have higher and lower F0

values respectively; despair and sadness pattern similarly (Banse & Scherer,

1996). The notion that there are different 'versions' of the same emotional state is

noted elsewhere in the literature as activation (Laukka, Juslin, & Bresin, 2007),

whereby differences in the degree of emotion are reflected in speakers' pitch.

In a study that examined vocal properties of feigned depression and

sleepiness, it was found that subjects were unable to alter their pitch range in a

manner consistent with genuine depression and sleepiness (Reilly, Cannizzaro,

Harel, & Snyder, 2004). The importance of this type of work to the domain of

forensic psychology highlights the importance of linguistic and/or phonetic

analyses, as depressive, and likely other disorders appear to be marked by

measurable acoustic features.[31] (Nilsonne, 1987, 1988; Stassen, Kuny, & Hell, 1998).

5.2.2 Pitch in police interviews and deception studies

In the previous two chapters, it was possible to survey a handful of studies in which pause and tempo were examined in police interviews and deception studies. I am unaware, however, of any studies that address pitch in authentic police interviews. In one study on a simulated police interview, Vrij (1995), found that lies were told with higher pitch values. However in another, no difference was found (Vrij & Winkel, 1991).

Police interviews aside, pitch is well studied in deception studies. Ekman, Friesen & Scherer (1976), showed that liars' voices displayed higher pitch, which they attribute to stress experienced while engaged in lying. The difference in pitch between truthfulness and deception was significant although they differed only by a few Hertz. The same conclusions are drawn in a variety of studies (Ekman, 1988, 1992; Ekman, O'Sullivan, Friesen, & Scherer, 1991; Streeter, Krausse, Geller, Olson, & Apple, 1977). Rockwell, Buller, & Burgoon, (1997) found the same, although their data were collected impressionistically, rather than instrumentally.

---

[31] The intersection of speech science and psychology is only slightly alluded to in this study, although it is acknowledged that this is a well-studied domain with a long tradition. To give further treatment of it goes beyond the scope of this work.

A study conducted by Anolli & Ciceri (1997) took a slightly different approach than many others in that truthful speech was compared with two different lying conditions: lying to a hearer who appeared to be convinced of the lie, and one who appeared suspicious. They found that average pitch, pitch variability, and pitch range differed significantly under all conditions and that both lying conditions showed elevated pitch values compared to the truthful condition. Furthermore, lying when the hearer was suspicious resulted in higher pitch than lying when the hearer appeared convinced. They attributed the higher pitch values to emotional arousal, with which nervous tension or fear might be expected (van Bezooijen, 1984).

While I am unaware of any studies in which lies were told at lower pitch values, a few studies found that pitch values of lies did not differ significantly from pitch values of truths (Bond, Omar, Mahmoud, & Bonser, 1990; Fiedler & Walka, 1993; Vrij & Winkel, 1991).

It is certainly possible and even highly probable that some responses produced by suspects in my corpus are deceptive, although at this time, there are no means available to me to determine deception. Nonetheless, as truthfulness and deception have been shown to influence speakers' pitch, it is necessary to identify them as potentially affecting prosodic characteristics.

5.3 Methodology

This section deals with three methodological issues. First, I discuss the notion of intonation groups, which are the linguistic units under examination in this portion of the study. Second, I provide the rationale for the types of measurements used, referring to methodological issues discussed in previous studies. Third, the design of the study on pitch phenomena is given.

Before measuring the pitch of suspects' responses, it was necessary to determine the spans of speech that would be measured. One possible method would involve measuring the high, low, and average fundamental frequency of each response. However, this approach would be inadequate to obtain the most useful characterization of suspects' pitch. If for example, a response is examined as a single intonation span with a maximum pitch of 7 semitones, and a minimum of 1 semitone, the range would be 6 semitones. A second possible method takes into account the minimum, maximum and range of all intonation units (IUs) in the response. If the same response were measured for the minimum, maximum, and range of two intonation units, where the first IU had a maximum of 7 semitones, and a minimum of 2 semitones, and the second IU had maximum and minimum pitch values of 4 and 1, then two ranges of 5 and 3 obtain, which gives a much more thorough measure of maximums, minimums, and ranges. In the interest of obtaining the most appropriate values, the second approach was chosen.

Following from the discussion above, we must also an "intonation unit" (IU). Suspects' responses that consist of a yes or no, or some other brief answer, seldom exceed one IU. Hence the maximum, minimum, and average are easily extracted using the phonetic software package Praat, version 4.4.20 (Boersma & Weeknink, 2006). However, for those responses composed of what appears to be more than one IU, it is necessary to decide how to divide an utterance into IUs. Cruttendon's definition, based on his notion of an Intonation Group (Cruttenden, 1997), is used. An example from the corpus illustrating intonation unit boundaries is provided below, followed by a simplified overview of his definition.[32]

> I: Was there a specific area that you were driving in?
> S: [1 Down in the crack-land area] [2 which is SOME STREET] [3 by the train there.]

Intonation units often, but not always, align with small syntactic units such as adverbials, nouns phrases with modifiers, or other such constituents, and almost always, they align with what would syntactically constitute a clause. Where they do not align with these categories, a pause and/or obvious pitch reset may appear. Intonation boundaries may or may not consist of a pause (Cruttenden, 1997). In the example above, the first IU is an adverbial phrase, followed by a second IU, which is a relative clause modifying the noun in the previous IU, and the third IU spans a prepositional phrase.

---

[32] For his complete definition of an intonation group, see Cruttendon (1997, pp. 68-87)

Throughout the literature, pitch is measured in a variety of ways, often by counting the number of cycles per second using the linear Hertz (Hz) scale. On this scale, Cruttenden (1997) reports average adult males' fundamental frequency ranges from 60 to 240Hz, with a mean of 120Hz, while the female range is higher on the scale, from 180 to 400Hz, with a mean of 220Hz. It is now well accepted that the manner in which we as hearers perceive sound, however, is not linear fashion, and that the higher the pitch, the greater the change in pitch there must be for a perceptible difference to be heard (Henton, 1989; Hudson & Holbrook, 1981). For this reason, the semitone scale allows comparisons to be made between speakers, since some speakers may have lower average frequencies and others higher. In this study, therefore, it was decided that the semitone scale, relative to 100Hz, would be used, which reflects pitch measurements logarithmically (see Table 5.4), in a manner that more closely mirrors human perception. Cruttenden's same male and female ranges in the semitone scale are reported for comparison in the table below.

Table 5.4   Comparison of hertz and semitone scales

	Hertz			Semitones		
	min	max	mean	min	max	mean
men	60	240	120	-8.8	15.2	3.2
women	180	400	220	10.2	24.0	13.7

Table 5.5 outlines the design of this portion of the study on pitch. Responses for which there is sufficient data to treat statistically will be handled

accordingly — these response types are affirmative, negative, relevant and irrelevant. Confessions, because there are fewer, will be handled on a more qualitative basis.

Table 5.5   Design of the study on pitch phenomena

Pausal Phenomena	Measure	Response Types
Average pitch	semitones	affirmative negative relevant irrelevant confession
Pitch variability	standard deviation of average pitch in semitones	affirmative negative relevant irrelevant confession
Pitch range	semitones	affirmative negative relevant irrelevant confession

Table 5.6 shows how many of each response type from each suspect group will be examined for pitch values. In this portion of the study, slightly fewer responses were analysed for pitch than were examined for pause and tempo, for two reasons. First, not all intonation phrases produced a clear enough pitch track from which reliable values could be extracted. Second, in cases of turn overlap, it was not possible to gather intonation phrases for the entire utterance, so those samples are excluded.

Table 5.6   Distribution of response types analysed for pitch phenomena by suspect

FT	FT-1	FT-2	FT-3	FT-4	FT-5	FT-6	FT-7	FT-8	Total
Affirmative	27	9	22	22	31	11	10	18	150
Negative	21	16	6	23	19	9	19	20	133
Relevant	24	17	34	22	26	12	18	18	171
Irrelevant	19	12		20	16	20	16	15	118
Confession		5		4		7			16
Total	91	59	62	91	92	59	63	71	**588**

RO	RO-1	RO-2	RO-3	RO-4	RO-5	RO-6	RO-7	RO-8	
Affirmative	12	18	10	41	23	11	12	17	144
Negative	21	14	9	18	12	11	11	10	106
Relevant	23	14	23	18	17	7	26	14	142
Confession		2		5	2				9
Total	56	48	42	82	54	29	49	41	**401**

There are 1003 responses in the larger set analysed for pause and tempo. The pitch study has 989, a difference of 14 responses. Suspect FT-2 had three relevant responses excluded from his sample set, suspect RO-1 had five relevant responses excluded, and suspect RO-5 lost four relevant responses, and suspect RO-7 lost two affirmative responses. In most cases, the excluded responses became perceptibly quieter toward the end, resulting in a response too quiet to measure for pitch. In the remaining cases, turn overlap interfered with collecting accurate pitch values. It is unsurprising therefore that relevant responses were more likely to be excluded, given that shorter answers would show less 'trailing off' toward termination. The total number of exclusions does not appear to be large enough to make a difference to the outcome of the study.

5.4 Results

In this section I discuss the results of pitch phenomena shown by suspects. Section 5.4.1 deals with the average pitch of responses; section 5.4.2 looks at the variability of suspects' pitch across the response types, and section 5.4.3 deals with pitch range across response types.

5.4.1 Average pitch

In this section, I present the average pitch values of suspects' responses. For each suspect in the two groups, the average pitch was recorded for each intonation unit of each response. An example of each type of response from Suspect FT-5's pitch averages is given below, with minimum, maximum and average values reported in semitones.

Table 5.7   Pitch values from Suspect FT-5

	Affirmative
I: So it's a carpeted floor with cement underneath it? S: [Yeah.]	
Minimum	3.9
Maximum	8.2
Average	5.6
	Negative
I: Was he uncomfortable at all during the day? S: [No not at all.]	
Minimum	3.2
Maximum	4.7
Average	3.7

					Relevant
I: So where was WITNESS? S: [WITNESS was right there with me.]						
Minimum						3.3
Maximum						7.1
Average						3.5
						Irrelevant
I: Who did you take him (infant victim) from? S: The, I think I took him from the from his carseat. Like we usually take him from the crib and put him in the carseat *irrel*→ [1 cause we don't have a swing right yet] [2 because we just moved] [3 like] [4 we do have a swing] [5 but it's at the mother's place.] [6 Or at the grandmother's]						

	IU1	IU2	IU3	IU4	IU5
Minimum	1.8	2.5	1.8	1.5	-1.4
Maximum	11.1	8.5	7.0	5.1	7.6
Average	3.9	3.9	5.0	3.5	4.5

The examples above illustrate the pattern that is found in the data, particularly with respect to FT suspects. While there is a great deal of variation, in general, affirmative responses show higher average pitch than both negative and relevant, As well, irrelevant responses are higher than relevant responses. Table 5.8 showsthe means for all members of both suspect groups' pitch minimums, maximums and averages.

Table 5.8   Minimum, maximum and average pitch in semitones by response type for FT

	Affirmative			Negative			Relevant			Irrelevant		
	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean
FT-1	-1.4	3.9	0.9	-1.8	2.7	-0.1	-2.1	3.5	0.3	-1.3	6.4	2.4
FT-2	1.8	6.2	4.4	-1.2	1.7	0.2	-0.8	4.5	1.6	-1.4	7.7	2.9
FT-3	-2.1	3	0.8	-1.3	2.1	0.5	-2.4	3	0.6			
FT-4	-2.3	-0.7	-1.5	-3.7	-2.4	-3	-3.7	-0.3	-1.8	-3.2	3.1	-0.6
FT-5	4.4	6.3	5.4	3.5	5	4.5	2.6	5.7	4	1.2	7.6	3.5
FT-6	1.9	6	4.1	0.2	4.1	1.7	-2.3	3.1	-0.2	-1.8	7.3	1.7
FT-7	1.8	6.3	4.5	1.4	3.9	2.8	1.1	5.9	3.2	0.8	7.4	4.3
FT-8	-0.4	2.7	0.8	-1.3	1.5	-0.1	-2	2.2	-0.2	-1.7	5	1.4
mean	0.46	4.21	2.43	-0.53	2.33	0.81	-1.2	3.45	0.94	-1.06	6.36	2.23
s.d.	2.38	2.5	2.47	2.2	2.28	2.24	2.08	2	1.91	1.54	1.72	1.6
RO-1	1	5.8	3.7	1.5	5.1	3.2	-1.6	3.5	1.1			
RO-2	-3.3	-0.9	-3.1	-4.8	-0.3	-2.2	-5.5	1.2	-2.2			
RO-3	6.4	9.2	7.9	-3.3	1.6	-0.8	-1	4.1	2.5			
RO-4	2.8	7.7	6	-2.8	3.7	0.3	-0.8	8	3.8			
RO-5	3	5.5	4.3	2.9	7.1	5.4	3.3	7.5	5.4			
RO-6	-3.5	6.6	2.6	-2.9	2.5	0.3	0.3	3.7	0.3			
RO-7	2.3	6.7	4.9	-0.1	7.1	3.9	3.4	11.1	7.2			
RO-8	6.6	9	7.9	2.8	9.6	6.1	2.5	13	8.2			
mean	1.91	6.2	4.28	-0.84	4.55	2.03	0.07	6.51	3.29			
s.d.	3.81	3.17	3.53	3	3.29	3.04	3	4.09	3.56			

It is possible to compare the two groups for Affirmative, Negative and relevant responses, recalling once again that irrelevant responses were produced too infrequently by RO suspects to be reliably analysed. According to the tables above, FT suspects' average pitch values are somewhat more consistent than ROs'. Average pitch for relevant and negative responses are overall lower, while affirmative and irrelevant responses pattern higher. For RO suspects, no discernable pattern emerges. Some suspects — RO-3, RO-4, RO-6 — show higher values for affirmative responses, and lower for negative and relevant. They are overall, however, quite inconsistent and this inconsistency is noted in the standard deviations reported in Figure 5.1.

Figure 5.1   **Average pitch values of both groups by response type**
                   Error bars indicate one standard deviation.
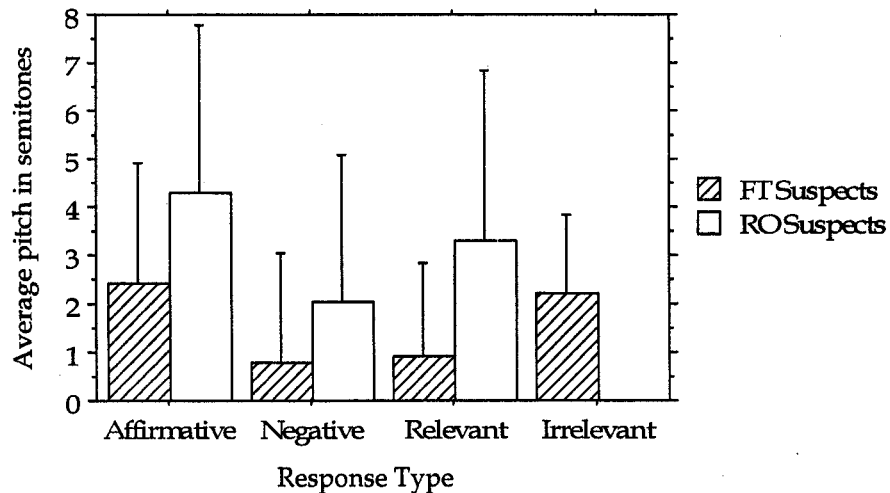
The average pitch data were submitted to a mixed design ANOVA with group and response type as factors, excluding irrelevant responses because only

FT suspects produced them. The interaction between group and response type is not significant: $F(2, 28)=.584, p=.5641$. While, FT suspects' responses appear to be produced at a lower average pitch overall compared to RO suspects, the effect of group is not significant: $F(1,14)=1.942, p=.1852$. FT suspects alone, however, for whom there is more consistency, a one-way repeated measures ANOVA for FT suspects' shows that average pitch values differ for responses, and the difference is significant $F(3,18)=7.042, p=0025$. Post hoc Tukey tests with Bonferonni adjustment show significant difference where affirmative is higher than negative, and affirmative is also higher than relevant. None of RO suspects' average pitch values differ significantly by response: $F(2,14)=2.753, p=.0981$.

For those who produced confessions, average pitch values varied widely for most suspects. The box below provides a representative example with semitone values reported for each intonation unit in the confession.

Table 5.9   Pitch values of a confession from Suspect FT-6

FT-6					Confession
I: Tell me about this time in MONTH. S: [1 I started you know], [2 playing with her hair and stuff] [3 and just you know] [4 started tryin to touch there] [5 and she started moving away.]					
	IU 1	IU 2	IU 3	IU 4	IU 5
Min	-1.7	-4.3	-2.4	-3.6	-2.4
Max	3.8	.2	3.2	0	1.1
Mean	.5	-2.0	0	-1.8	-.7

FT-6's confession, one of six he made during his interview, consists of five

intonation units in which he described a particular interaction with his child

victim. The average pitch values of his confessions do not stray far from the

mean of his average pitch for relevant responses. The figure below shows the

mean pitch values of each suspect's confessions. As a point of comparison,

suspects' relevant response average pitch values are also shown.

Figure 5.2   **Average pitch values of confessions**

         Solid dots indicate values of confession intonation units. Hollow dots show
         suspects' mean pitch for relevant responses.
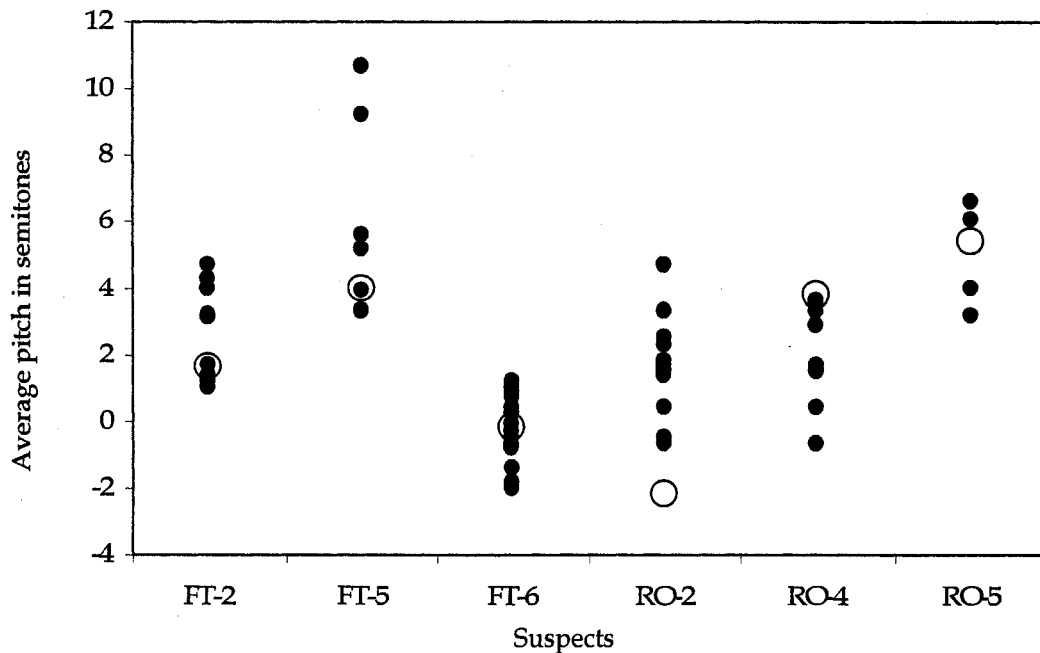
For suspects FT-2, FT-5 and RO-2, the average pitch of intonation units in

confessions is overall higher compared to the same suspects' means for relevant

responses. Suspects' FT-6 and RO-5 show relevant means in the middle range of

their confession pitch averages, and suspect RO-2's relevant mean is well below that of his confession average pitch. No discernable pattern stands out differentiating the groups, nor comparing confession average pitch with relevant response average pitch.

5.4.2 Pitch variability

Results on pitch variability are reported in this section. Pitch variability measures the amount of variation from the average pitch that suspects exhibit when responding to police questions. Variability is measured as the standard deviation of suspects' average pitch responses. See the previous section for the average pitch values upon which the results of this section are based. Table 5.10 summarizes the pitch variability values for suspects of both groups, for each response type.

Table 5.10 Pitch variability by response type for both suspect groups

Each value below represents all suspects' standard deviations of average pitch across response types.

FT	FT-1	FT-2	FT-3	FT-4	FT-5	FT-6	FT-7	FT-8	mean	s.d.
Affirm	1.3	1.3	1.8	1.5	2.4	2.4	2.9	1.4	**1.88**	0.61
Negative	1.7	1.6	1.4	1.0	1.6	1.2	2.5	1.6	**1.58**	0.44
Relevant	1.6	2.4	2.0	1.1	0.8	1.9	1.0	1.5	**1.54**	0.55
Irrelevant	1.7	1.6		1.4	0.9	2.1	2.0	1.7	**1.63**	0.40

RO	RO-1	RO-2	RO-3	RO-4	RO-5	RO-6	RO-7	RO-8	mean	s.d.
Affirm	3.3	3	3.8	4.1	0.6	2.1	3.4	3.0	**2.91**	1.11
Negative	3.6	1.9	2.0	4.1	1.3	3.0	4.7	2.7	**2.91**	1.17
Relevant	1.9	1.6	3.4	2.8	1.2	2.2	3.9	2.8	**2.48**	0.92

A comparison of the means by response type shows a clear pattern: FT suspects' pitch shows considerably less variability across all response types — under two — than RO suspects', which ranges from a group mean of 2.48 to 2.91. The figure below compares the two groups.

171

Figure 5.3 **Mean standard deviations of average pitch for both groups for each response type Error bars indicate one standard deviation.**
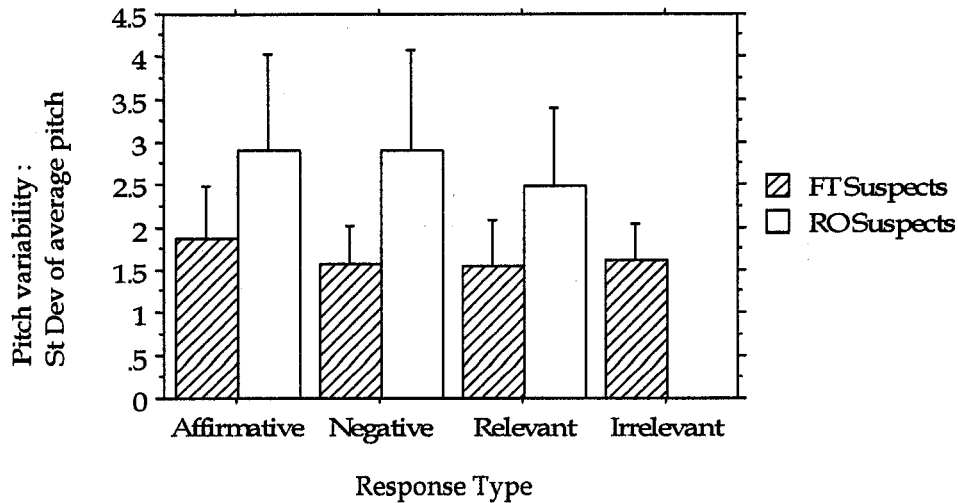
Excluding the response category irrelevant which was produced by only FT suspects, the remaining response categories, affirmative, negative, and relevant, were submitted to a mixed design ANOVA. Results show no significant interaction between group and response type: $F(2,28)=.447$, $p=.6439$. However, it is clear from the data in Table 5.10, FT suspects' variability values are lower than ROs', and a one-way between group ANOVA for each response type shows that the two groups differ significantly: affirmative, $F(1,14)=5.361$, $p=.0363$; negative, $F(1,14)=9.137$, $p=.0091$; relevant, $F(1,14)=6.142$, $p=.0266$. Within the groups, differences between the response types were not significant.

Among the variability values for suspects who confessed, suspect FT-4 shows mean relevant variability to be lower than his confession mean variability, and to a lesser extent, the same can be said for RO-6. All others show lower mean

variability in their confessions than their relevant responses. Figure 5.4 depicts these values.

Figure 5.4   **Mean pitch variability of confessions**

Diamonds indicate mean variability for relevant responses for comparison.
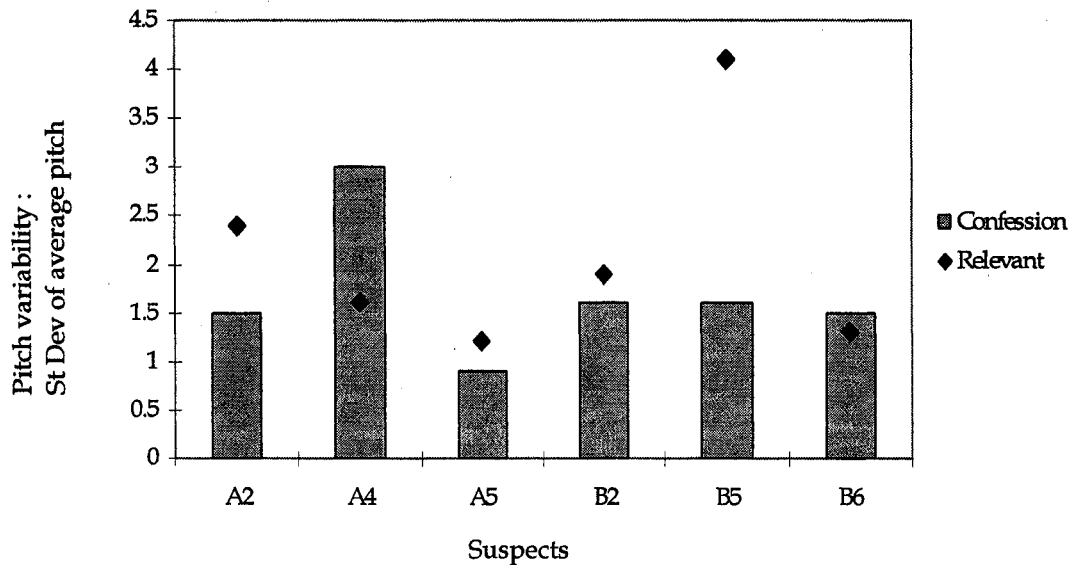
The figure above shows some tendency for relevant response pitch variability to be similar to, or higher than confession pitch variability.

5.4.3   Pitch range

In this section the results of the study on pitch range are reported. The pitch range was taken as the difference between the maximum and minimum of each intonation unit in the responses, for each suspect. The examples below, taken from Suspect FT-6's interview illustrate how pitch range was measured in semitones. Intonation units are bracketed.

Table 5.11   Pitch range values from Suspect FT-6

	Affirmative
I: Everybody were in their room sleeping and Haley was sleeping on the couch. You went out to check on Haley. S: [Yeah.]	
Minimum	-1.4 ST
Maximum	4.2 ST
Range	5.6 ST

	Negative
I: So your penis was never in her, was it. Never? S: [Never.]	
Minimum	-1.3 ST
Maximum	2.1 ST
Range	3.4 ST

	Relevant
I: So on Boxing Day she came over, you had a couple of beers. How many beers did you have? S: [I had probably like four.]	
Minimum	-4.3 ST
Maximum	6.0 ST
Range	8.3 ST

				Irrelevant
I: How, how often do you drink? S: Relevant→ For a while, I was like drinking ever day. Irrel→ [1 I actually just cleaned it all out] [2 the other day] [3 under the bed.] [4 It's quite a bit of wasted money.]				
	IU 1	IU 2	IU 3	IU 4
Minimum	-1.0	-1.6	-1.8	-2.4
Maximum	10.8	7.5	4.6	5.4
Range	11.8	.04	6.4	7.8

The table below shows the mean pitch ranges for each suspect in both groups, by response type. For suspects' mean minimums and maximums upon which these data are based, refer to Table 5.8.

Table 5.12   Pitch range in semitones by response type for both suspect groups

FT	FT-1	FT-2	FT-3	FT-4	FT-5	FT-6	FT-7	FT-8	mean	s.d.
Affirm	5.3	4.4	5	1.6	2	4.1	4.5	3	**3.7**	1.29
Negative	4.5	2.9	3.4	1.3	1.5	3.9	2.5	2.7	**2.8**	1.03
Relevant	5.6	5.3	5.4	3.4	3.1	5.3	4.8	4.2	**4.6**	0.90
Irrelevant	7.7	9.1		6.3	6.4	9.2	6.6	6.8	**7.4**	1.16

RO	RO-1	RO-2	RO-3	RO-4	RO-5	RO-6	RO-7	RO-8	mean	s.d.
Affirm	4.8	2.5	2.8	4.9	2.5	10	4.4	2.4	**4.3**	2.38
Negative	3.7	4.5	4.9	6.5	4.2	5.4	7.2	6.8	**5.4**	1.21
Relevant	5	6.7	5	8.8	4.2	3.4	7.7	10.5	**6.4**	2.29

According to the table above, the pitch ranges of FT suspects vary consistently. Affirmative response pitch ranges are uniformly wider than negative responses. Relevant responses are uniformly narrower than both negative and affirmative, and irrelevant responses are widest for all suspects. This consistency does not hold for RO suspects. While relevant responses show at least somewhat wider pitch ranges for most suspects, no consistent pattern is to be found. Affirmative responses are lower than negative responses for six suspects, and the opposite is true for the remaining two. The inconsistency is noted in the standard deviations of RO suspects, which is greater than the

175

standard deviations for FT. Figure 5.5 provides a visual comparison of the two

groups.

Figure 5.5 **Pitch range values for both suspect groups by response**
Error bars indicate one standard deviation.
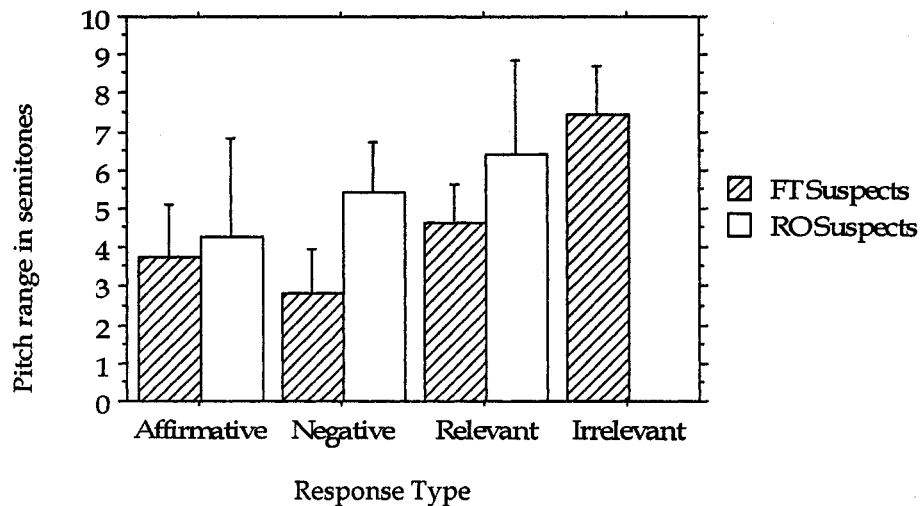
The figure above suggests that RO suspects' pitch ranges are overall wider

than FT suspects. However, a mixed-design ANOVA shows no interaction of

group with response type: $F(2,28)=1.673$, $p=.2060$, excluding irrelevant responses

which were only produced by FT suspects. There is, nevertheless, an effect of

group: $F(1,14)=7.583$, $p=.0155$. Closer inspection of within group differences

show that FT suspects' pitch ranges differ significantly across conditions,

$F(3,18)=87.410$, $p=<.0001$, and post hoc Tukey tests with Bonferonni adjustment

show significant difference in all pairwise comparisons of response types. The

same is not true for RO suspects, in which no significant differences are found

between response types: $F(2,14)=1.901$, $p=.1860$.

The suspects who produced confessions showed pitch ranges that vary

considerably within the response category. The example used for average pitch is

shown below illustrating range:

Table 5.13 **Pitch range values from a confession from Suspect FT-6**

Suspect FT-6				Confession
I: Tell me about this time in November. S: [1 I started you know], [2 playing with her hair and stuff] [3 and just you know] [4 started tryin to touch there] [5 and she started moving away.]				
	IU 1	IU 2	IU 3	IU 4	IU 5
Min	-1.7	-4.3	-2.4	-3.6	-2.4
Max	3.8	0.2	3.2	0.0	1.1
Range	5.5	4.5	5.6	3.6	3.5

Suspect FT-6's confession is composed of five intonation units in which

the narrowest range is 3.6 semitones and the widest range is 5.6 semitones. His

mean pitch range for relevant responses is 5.3 semitones, which makes this

particular example a reasonable representation of his pitch ranges in the

confession data. Figure 5.6 summarizes the findings of pitch range in confessions

for each suspect.

Figure 5.6   Pitch range values for confessions

Solid dots indicate pitch range values of individual intonation units. Hollow dots indicate mean pitch range for relevant responses.
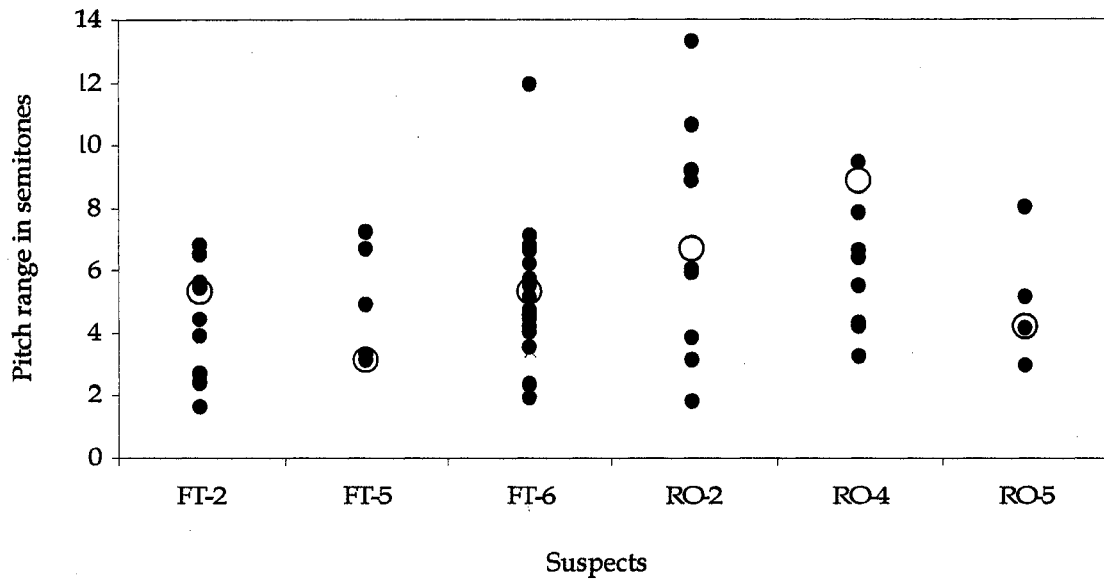
For suspects FT-6 and RO-2, relevant pitch ranges are similar to those of confessions, for suspects FT-2 and RO-4, confession pitch ranges are mostly lower than relevant pitch ranges, and for suspects FT-5 and RO-5, confession pitch ranges tend to be high in comparison to relevant range means. Certainly, no conclusions can be drawn with respect to differences between the groups.

5.5   Pitch profiles and conclusion

The previous subsections present the results of the three studies on suspects' pitch behaviour: average pitch, pitch variability, and pitch range. In the following sections, profiles of the two groups' pitch phenomena summarize the

findings, along with discussion relating to the current literature. The research questions in section 5.1 are readdressed to conclude the chapter.

5.5.1 Pitch profiles and discussion

As has been done for pause and tempo, suspects' pitch values by group and response type are drawn together to compose profiles summarizing relative pitch values taken statistically significant differences and strong tendencies shown in Figures 5.1, 5.3, and 5.5. The same schema as in previous chapters is used below in which the tallest cones represent the greatest values, short cones the lowest values and mid sized cones represent mid values. Empty fields on the figures below represent features that vary too widely to be considered high, mid, or low.

Figure 5.7   FT pitch profiles by response type, showing relative values
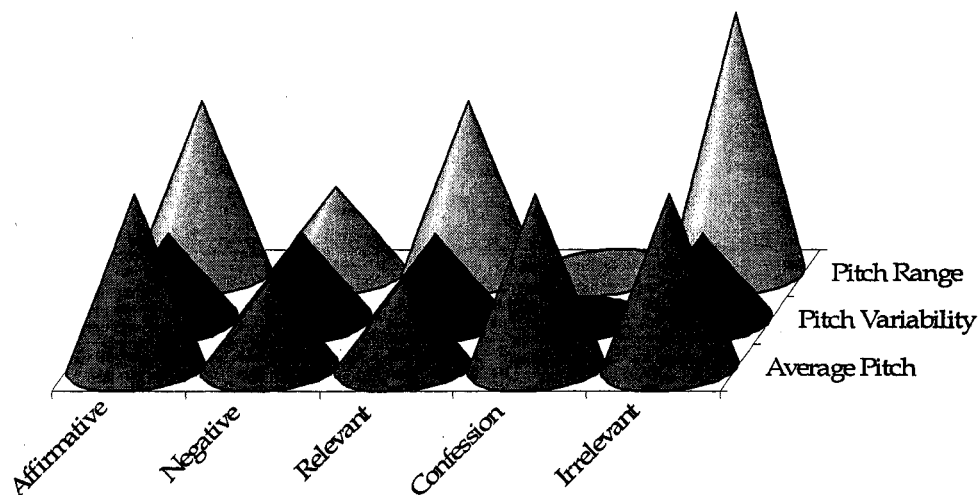
Figure 5.8 RO pitch profiles by response type, showing relative values
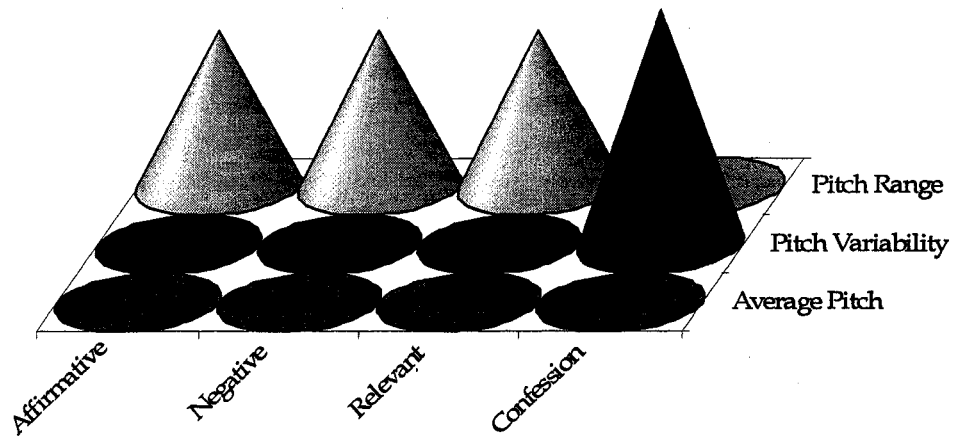
The significant and non-significant results are summarized and the following tables recap the results by statistical significance.

Table 5.14   **Summary of significant and non-significant differences by group and response type for average pitch**

Feature	Comparison		Significance	Test
	Group	Response Type		
Average Pitch	FT/RO	aff/neg	n.s.	F^a
		neg/rel	n.s.	F^a
		rel/aff	n.s.	F^a
	FT	aff/neg	<.05	F^b
		aff/rel	<.05	F^b
		aff/irr	n.s.	F^b
		neg/rel	n.s.	F^b
		neg/irr	n.s.	F^b
		rel/irr	n.s.	F^b
	RO	all	n.s	F^b

[a]mixed design; [b]one-way repeated measure

Table 5.15   **Summary of significant and non-significant differences by group and response type for pitch variability**

Feature	Comparison		Significance	Test
	Group	Response Type		
Pitch Variability	FT/RO	aff/neg	n.s.	F^a
		neg/rel	n.s.	F^a
		rel/aff	n.s.	F^a
	FT	all	n.s.	F^b
	RO	all	n.s	F^b
	FT/RO		<.05	F^c

[a]mixed design; [b]one-way repeated measure; [c]one-way between group

Table 5.16 Summary of significant and non-significant differences by group and response type for pitch variability

Feature	Comparison		Significance	Test
	Group	Response Type		
Pitch Range	FT/RO	aff/neg	n.s.	F^a
		neg/rel	n.s.	F^a
		rel/aff	n.s.	F^a
	FT	all	<.05	F^b
	RO	all	n.s	F^b
	FT/RO		<.05	F^c

[a]mixed design; [b]one-way repeated measure; [c]one-way between group

The figures above generalize the findings for pitch values separately for the two groups of suspects. The clearest observation to be gleaned from the figures is the lack of consistent pitch values reported for RO suspects, as indicated by the absence of eight of the possible twelve cones in the figure. FT suspects' profiles, on the other hand, were more consistent throughout the data. FT suspects' average pitch values differed by response type, with affirmative and relevant responses in the mid range, negative low, and confessions and irrelevant responses showed comparably high averages. For RO suspects, average pitch was the only pitch variable that could be characterized reliably, in that all suspects showed mid range pitch averages. On the whole, FT suspects' pitch variability values were low in all response types except confessions, which varied. Conversely, RO's pitch variability was variable across all response types except confessions, in which it was found to be quite high.

Sadness and remorse are conveyed by lower pitch values (Banse &
Scherer, 1996; Johnstone & Scherer, 2000; Sobin & Alpert, 1999; Thompson &
Balkwill, 2006) and it therefore seems natural to expect that suspects ought to
display pitch values associated with sadness and remorse when confessing. This
expectation is not met, however, and in fact, where pitch is concerned,
confessions were perhaps too few in the data for any pattern to emerge. At best,
with this small set, the only findings that appear consistent are a mid average
pitch for FT suspects, and mid pitch range for RO suspects.

For FT suspects, negative responses were accompanied by lower pitch
values overall and affirmative responses were in the mid ranges for average
pitch and pitch range. This result echoes the findings of earlier studies on a
smaller sets of suspects (Fadden, 2006b, to appear), in which it was shown that
when responding negatively, the question is most often one that casts negative
light on the suspect, and the response comes with lower pitch values. When
answering affirmatively, on the other hand, the question is frequently of either
neutral or positive, and pitch values are higher. It was not possible to derive any
pattern of this sort from the RO suspects.

When a suspect departs from the topic set by the investigator, it appears
that a change in pitch can occur. Irrelevant responses are mid and high for
average pitch and pitch range respectively, and are consistently higher than
relevant responses. This finding concurs well with studies that show higher pitch

associated with topic changes (Crystal, 1969; Grosz & Hirschberg, 1992; Hirschberg & Grosz, 1992; Swerts & Geluykens, 1994; van Donzel, 1999).

5.5.2 Conclusion

The goal of this chapter has been to examine pitch phenomena of first time and repeat offender suspects while responding to police. In so doing, the following research questions were answered:

Do response type and suspects FT or RO status affect:

a) the average pitch of responses?

b) the pitch variability of responses?

c) the pitch range of responses?

For FT suspects, all three pitch variables differed across response types. Affirmative and relevant responses were characterized by low to mid values; negative by low values; and confessions and irrelevant, by mid to high values, except for pitch variability, which was either too variable in the case of confessions, or low in the case of irrelevant.

For RO suspects, pitch range was the only pitch variable that was reliably profiled, although response types were not shown to differ in range.

From a forensic point of view, that RO suspects' response types vary more than FT suspects' in terms of pitch characteristics, it is clear that pitch may not be a reliable prosodic indicator when it comes to searching for vocal cues of

deception, or assessing suspects' emotional states. FT suspects, however, do show more consistent behaviour, both as a group, and within individual suspects' values across response types, and pitch values may be a more reliable indicator when detecting deception or ascribing emotion to speakers. This difference with respect to the consistency shown by the two groups is a strong indicator that any conclusions for one group are not applicable to another.

6 PROSODIC PROFILES

For by thy words thou shalt be justified,
and by thy words thou shalt be condemned.

–Jesus Christ (Matthew 12:37)

The last three chapters identified those prosodic features that are

relatively uniform within groups across response types, and those features that

are too variable to profile reliably. In this concluding chapter, I draw together the

previous three chapters' results to create prosodic profiles consisting of pause,

tempo and pitch features for each response type from both groups. This chapter

commences with an overview of the dissertation, after which, prosodic profiles

are presented and discussed as the overall findings produced by this study. I

then elaborate on the contribution of this corpus study to forensic linguistics

alluded to in Chapter One. The chapter closes with directions for future research.

6.1   Overview of the dissertation

The purpose of this dissertation has been to offer a descriptive

examination of the prosodic properties of suspects' speech as it naturally occurs

during police interviews. In Chapter One, I surveyed the literature on police

interviews, which are normally undertaken in the frameworks of Conversation Analysis and Discourse Analysis. The literature on deceptive speech was also explored, for two reasons: first, the connection between police interviews and deception is a logical and practical one since it is expected that some suspects might be quite motivated to lie, and that there would be legal implications for truthfulness and deception; and second, the methods used in deception studies provided much in the way of guidance for my own methods used in this study. The third area of research gave a very brief overview of the interaction between prosody and discourse, and prosody and emotion. The research questions were listed that would be dealt with in subsequent chapters.

Chapter Two introduced the corpus, including information on the suspects' whose speech would be studied. The technical methods by which the data were handled and the types of responses that would be examined were also defined and exemplified. Limitations on the scope of this corpus study were noted, in particular the fact that the role investigators' discourse may have on the discourse of suspects is left for future study.

Chapters Three, Four, and Five dealt with pause, tempo and pitch phenomena respectively. Each of these chapters was presented in parallel fashion. The literature was reviewed addressing the appropriate prosodic feature, followed by the specific methods by which that feature would be handled. Results were presented, and conclusions drawn relating back to

research questions. The present chapter brings together the results of the previous chapters in two ways. First, Section 6.2 summarizes the findings by response type. Second, Section 6.3 summarizes the findings according to whether results were statistically significant between and within groups, on the basis of each of the prosodic fields analysed.

6.2 Prosodic profiles

The previous three chapters addressed pause, tempo and pitch phenomena, in relation to suspects' response types and the differences between first time and repeat offenders. In this section, I readdress the research questions introduced in Chapter One, and I draw together the profiles generated with respect to the findings presented in the data chapters to create a comprehensive look at the prosodic features of suspects' speech.

Whereas in the previous chapters I handled one prosodic feature at a time, comparing response types by suspect group, and summarizing the statistically significant differences, I will now arrange the results according to response type showing both groups, as well as all measured prosodic features. The following subsections recast the prosodic profiles introduced in Chapters Three, Four and Five by response type. Recall that small cones represent low values, mid-sized cones, mid values, and large cones, high values. Flat discs represent variables that vary too widely to profile, and empty fields were not measured for reasons indicated in the methodology sections. It is important to note that the visual

profiles are dealt with subjectively on the basis of the relative values depicted in the preceding chapters, and represent both statistically significant findings, as well as reasonably strong tendencies. Confessions are included for the sake of comparison, although the obvious limitation to their conclusiveness associated with their infrequency is acknowledged.

6.2.1 Affirmative

Affirmative responses are those in which the suspect confirms the information in the investigator's prompt. The prosodic profiles of affirmative responses are shown below in Figure 6.1.

Figure 6.1   **Prosodic profiles of affirmative responses for FT and RO suspects**
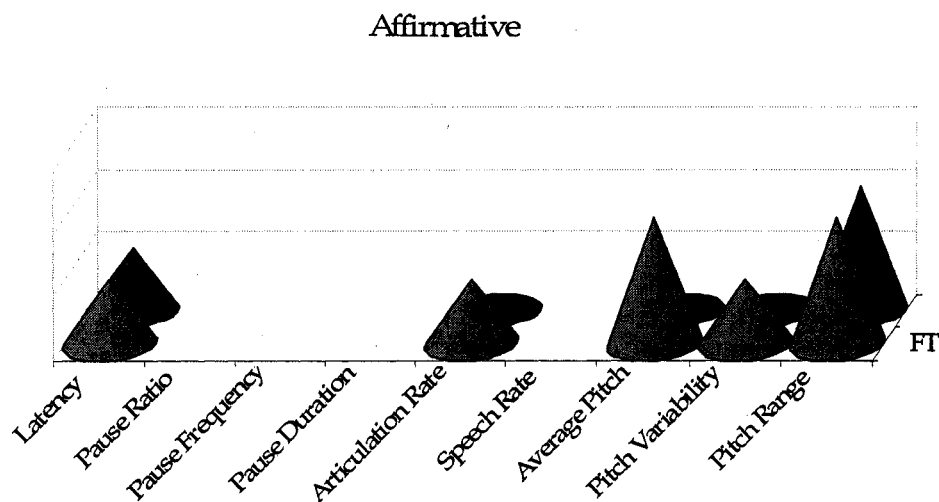
Affirmative responses show low latency for both groups of suspects. Recall that other pause values were not measured because affirmative responses were too short for pauses to occur within them. Articulation rates were low

among FT suspects, however, RO suspects' articulation rates varied too greatly to

create a profile. The second temporal feature, speech rate, measured in syllables

per second was not conducted, since affirmative responses are mostly

monosyllabic. Average pitch and pitch variability varied too widely among RO

suspects, however they showed mid pitch ranges compared to other response

types. FT suspects' pitch values presently clearer profiles, with low pitch

variability and mid average pitch and range.

Statistically significant and non-statistically differences between the

groups within the affirmative response type are summarized in the table below.

Table 6.1   Summary of statistical differences based on unpaired *t*-tests between the two
groups for Affirmative

Shaded fields indicate dimensions not measured

Prosodic Feature	Significance	Comparison
Latency	n.s.	
Pause Ratio		
Pause Frequency		
Pause Duration		
Articulation Rate	n.s.	
Speech Rate		
Average Pitch	n.s.	
Pitch Variability	<.05	FT greater than RO
Pitch Range	n.s.	

6.2.2 Negative

Negative responses are those in which suspects reject or negate the information contained in the investigator's prompt. The prosodic profiles for negative responses, shown in Figure 6.2, share some similarity with the profiles for affirmative responses.

Figure 6.2   Prosodic profiles of negative responses for both groups

Negative responses, like affirmative, show low latency values for both groups. Articulation rates for RO suspects are low, but for FT suspects, the rates varied too greatly to profile, and the opposite was true for affirmative responses. Negative response pitch values for RO suspects were the same as those for affirmative. FT suspects' pitch values were uniformly low, in comparison with pitch values for other response types.

Statistically significant and non-statistically differences between the groups within the negative response type are summarized in the table below.

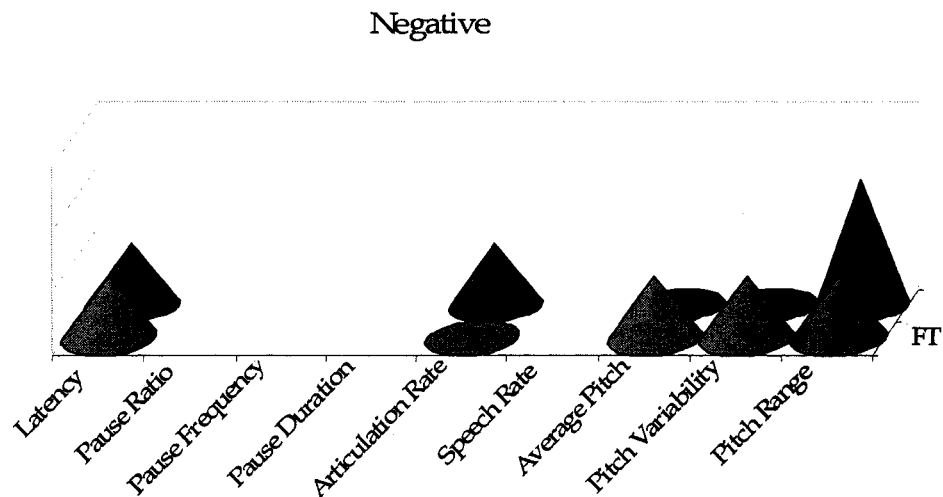Table 6.2   **Summary of statistical differences based on unpaired *t*-tests between the two groups for Negative**

Prosodic Feature	Significance	Comparison
Latency	n.s.	
Pause Ratio		
Pause Frequency		
Pause Duration		
Articulation Rate	n.s.	
Speech Rate		
Average Pitch	n.s.	
Pitch Variability	.0091	FT lower than RO
Pitch Range	.0008	FT lower than RO

6.2.3 Relevant

Relevant responses are those in which the suspect offers information related to the investigator's prompt. This response type can vary in length from a few words to a number of clauses. The prosodic profile for relevant responses is shown in Figure 6.3. Note that all prosodic features were measured for this response type.

Figure 6.3   Prosodic profiles of relevant responses for FT and RO suspects
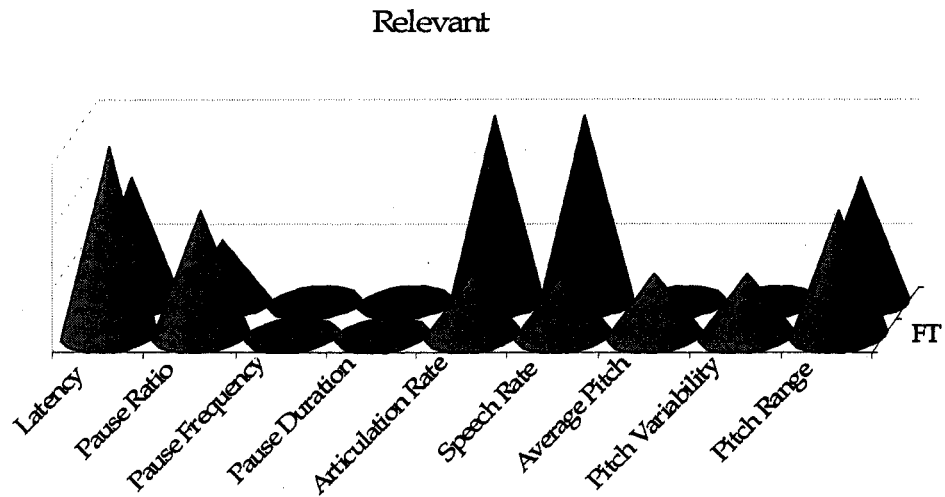
Relevant

The stand-out features that differentiate FT and RO suspects' relevant responses are speech and articulation rate. RO suspects' temporal rates are considerably higher than FTs' rates. Where pitch values are concerned, again, RO suspects' average pitch and variability vary too widely to profile, while FT suspects' averages and ranges are generally low. Both groups showed moderate range values. Pause features showed much difference between the groups, with FT suspects taking significantly more delay to before producing relevant responses, and greater pause ratios than RO suspects. RO suspects' pause ratios, frequencies, and durations were low.

Statistically significant and non-statistically differences between the groups within the relevant response type are summarized in the table below.

Table 6.3 **Summary of statistical differences based on unpaired *t*-tests between the two groups for Relevant**

Prosodic Feature	Significance	Comparison
Latency	.0388	FT greater than RO
Pause Ratio	.0422	FT greater than RO
Pause Frequency	n.s.	
Pause Duration	n.s.	
Articulation Rate	<.0001	FT less than RO
Speech Rate	.0004	FT less than RO
Average Pitch	n.s.	
Pitch Variability	.0266	FT lower than RO
Pitch Range	n.s.	

6.2.4 Irrelevant

Irrelevant responses are those that do not pertain to the investigator's prompt. They may stand alone, or they may intermingle with a relevant response. Recall that irrelevant responses were measured only for FT suspects, because this was not a response type produced frequently by RO suspects.

The differences between FT suspects' relevant responses and their irrelevant responses are apparent comparing Figure 6.3 above and Figure 6.4 below.

Figure 6.4   Prosodic profiles of irrelevant responses for FT and RO suspects
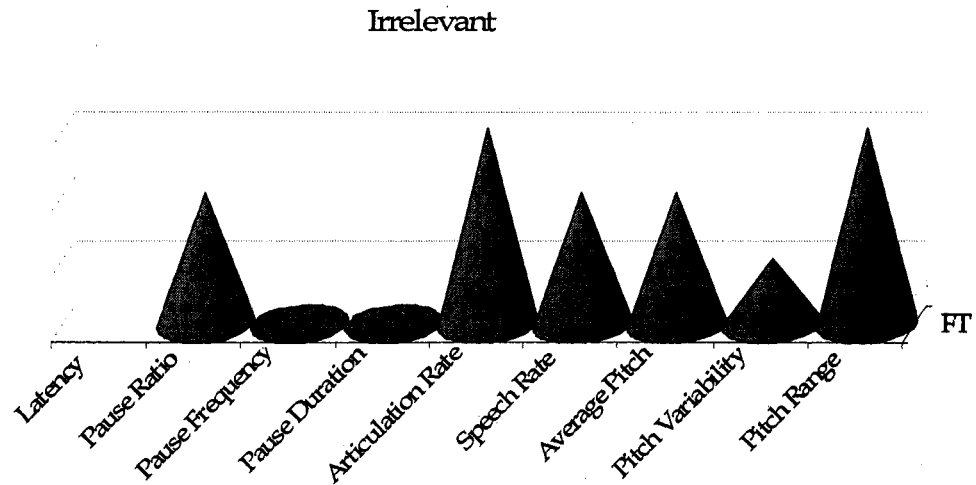
Irrelevant

FT suspects' temporal rates increased when they produced irrelevant

information, as well as average pitch and pitch range. Pitch variability remained

low, compared to relevant responses. The pause ratio remained similar, although

pause duration was somewhat higher. Pause frequency was highly varied in

irrelevant responses, as it was in relevant. Latency was unmeasured because

some irrelevant information was connected to a relevant response that went off

topic.

Statistically significant differences between irrelevant response prosodic

values and other response types are summarized in the table below.

Table 6.4   Summary of statistical differences based on paired *t*-tests between response types within the FT group

Non-significant results omitted for brevity

Prosodic Feature	Significance	Comparison
Speech Rate	.0123	Irr greater than Rel
Articulation Rate	.0196	Irr greater than Rel
	.0075	Irr greater than Aff
Average Pitch	.0084	Irr greater than Rel
	.0404	Irr greater than Neg
Pitch Range	<.0001	Irr greater than Rel
	<.0001	Irr greater than Aff
	<.0001	Irr greater than Neg

6.2.5 Confession

The last of the five profiles, confessions, were not numerous in the corpus,

however, they were analysed owing to their importance in a criminal

investigation. Of all the response types, confessions have the greatest number of

prosodic features that varied too widely to be adequately profiled, possibly due

to the sample size being quite small compared to other responses. However,

prosodic profiles based on the limited information I have, are found in Figure

6.5.

Figure 6.5   Prosodic profiles of confessions for FT suspects
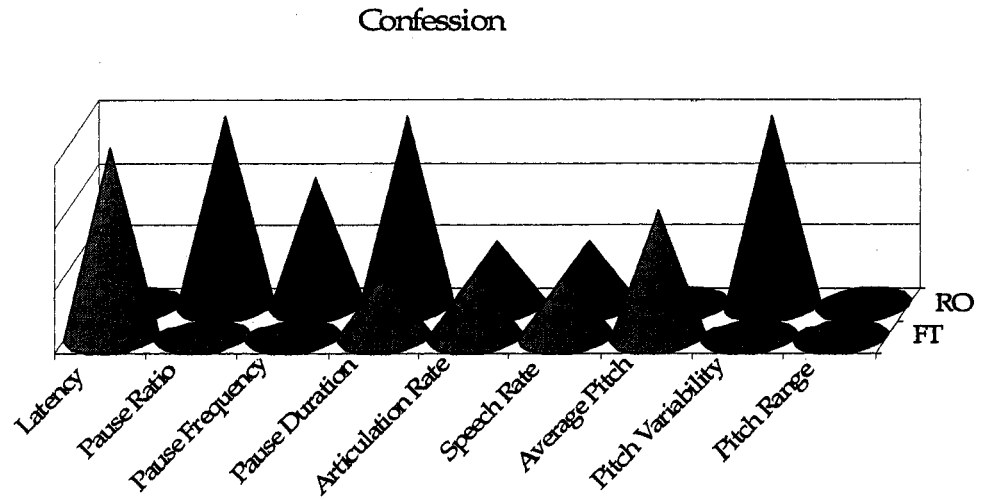
Confession

It was possible to draw only a few features out of the confession data. However, in comparison with other responses, RO suspects seemed to pause a great deal more during confessions, resulting in higher pause ratios, frequencies and durations. FT suspects consistently paused longer before starting a confession, but pause ratio and frequency varied. Pause durations tended to be low for FT suspects. Temporal values were low for both suspect groups, and for RO suspects, this marked considerable change from the pace at which they delivered relevant responses. Pitch values were not consistently profiled for either group.

6.2.6   Discussion of prosodic profiles

It is apparent from the preceding figures that FT suspects are somewhat more homogenous in their prosodic behaviour as a group than RO suspects, who

show more prosodic features that vary too widely to be consistently profiled.

Table 6.5 compares the number of features profiled for each group based on

statistically significant differences and reasonably strong tendencies, for response

types produced by both groups.

Table 6.5   Summary of profiled features by response type for FT and RO suspects

	Affirmative		Negative		Relevant		Confession	
	FT	RO	FT	RO	FT	RO	FT	RO
Latency	✓	✓	✓	✓	✓	✓	✓	✓
Pause Ratio					✓	✓		✓
Pause Freq						✓		✓
Pause Dur					✓	✓	✓	✓
Artic Rate	✓			✓	✓	✓	✓	✓
Speech Rate					✓	✓	✓	✓
Avg Pitch	✓		✓		✓		✓	
Pitch Var	✓		✓		✓			✓
Pitch Range	✓	✓	✓	✓	✓	✓		

Taking into account the most frequently occurring response types,

affirmative, negative and relevant, FT suspects are more consistent in their

prosodic features, having 17 profiled variables across these three response types.

RO suspects, on the other hand, show 12 profiled variables across the same

response categories. While RO suspects' confessions show more consistency in

their prosodic features as a response type, at this stage, I would not bestow as

much confidence in the results, owing to their small sample size.

198

The reason for the difference in the number of profiled features between groups is open to speculation. It is tempting to suggest that suspects for whom this is a first experience, are exercising more control over their speech and as a result show lower temporal rates, and lower pitch values that might be considered 'subdued' in an effort to control fear or other negative emotion. RO suspects, on the other hand, may be exhibiting more varying, natural ranges in prosodic behaviour. I reiterate, however, this is only an educated guess.

Returning to the research questions in Chapter One, I deal with pause, tempo and pitch individually, drawing on results sections from the previous three chapters, and highlighting the strongest findings.

1) Do suspect group and response type have an effect on

   a) the latency (i.e., delay) between the investigators'
   prompts and suspects' responses?

   b) the pause to speech ratio of responses?

   c) the frequency of pauses in responses?

   d) the duration of pauses in responses?

The data did show some differences in pausal behaviour, both within groups across response types, and between groups. Of the four pause features examined, latency was the most reliably profiled feature. Overall, FT suspects took longer than RO suspects before producing relevant information and

confessions. Both groups took more time before producing relevant responses than when producing affirmative and negative responses, although RO suspects took less time than FT suspects. RO suspects showed low pause frequency and duration, resulting in lower pause ratios in relevant responses, lower in fact than their confessions.

2) Do suspect group and response type have an effect on

a) the speech rate of responses?

b) the articulation rate of responses?

FT suspects showed considerably lower speech and articulation rates than RO suspects when producing relevant responses. FT suspects' irrelevant responses were accompanied by much higher articulation rates, and somewhat higher speech rates compared to relevant responses.

3) Do suspect group and response type have an effect on

a) the average pitch of responses?

b) the pitch variability of responses?

c) the pitch range of responses?

Of the three prosodic features, pitch presented the least consistent profiles. Among FT suspects, irrelevant responses showed higher average pitch and pitch range than relevant responses. While on the whole, FT suspects' responses were characterised by lower pitch values than RO suspects, there were too many

response types with widely varying values for reliable profiles to be created. I address this lack of consistency in the next section.

6.3 Relevance to forensic linguistics and potential application

Forensic studies that seek to uncover linguistic behaviour for purposes of examining legal discourse, detecting deception, or attributing emotional states, often take into account linguistic data, in particular phonetic or prosodic information. The majority of forensic linguistic work thus far focussing on deception and emotion is conducted in laboratories and on case studies in the form of forensic reports, which in my mind constitute two thirds of inquiry. I address the strengths and limitations of these here, and identify where the other third, corpus studies such as this one, forge new paths in the application of linguistics to forensic work.

First, few forensic linguistic studies focussing on vocal properties of deception and emotion are carried out on spontaneously produced, naturally occurring speech, therefore the majority of what we know to date comes from laboratory studies. Indeed the results of such studies, while authors readily acknowledge the limitations of induced emotional states or deceptive speech, have brought to light the issues facing the field and perhaps importantly, the empirical methods by which such studies can be conducted. While there is still much need for laboratory work to continue, the move toward more study of natural speech in forensic contexts is now due.

Second, case studies found in published forensic reports offer to the forensic linguistic community, the practitioners' eye-view into real-world linguistic application, opening the door for debate on best practice in a field that is still relatively new. Case studies, however informative in the manners just described, are representative of a single subject, not necessarily applicable to a larger population. Not to understate their strength, however, they are but one third of inquiry.

Corpus studies, of which the present study is an example, round out the final third of inquiry. Only coming into their own in recent decades with the advent of computational methods at our disposal, they offer to forensics and other areas, the power to examine a large amount of data, extracting from anything ranging from slight tendencies to strong patterns. Thus far, not many studies have been conducted on a corpus of police interviews, making this and other corpus work by Heydon (2005), and the forthcoming dissertation of Haworth (in prep) for example, relatively novel contributions.

I depart now from the discussion of where this dissertation fits in the tri-part model of inquiry, to discuss the relevance of this dissertation to forensic linguistics. What has been demonstrated in this dissertation is that prosodic properties of suspects' speech are dependent on two important factors: the type of information contained in the message, and whether the suspect is an experienced interviewee (classified herein as repeat offender, RO) or an

inexperienced one (first-timer, FT). This means that if a study is to address

phonetic differences between truth and lie, or genuine emotional disturbance or

false, the methodology must take into account at least three points: whether that

utterance is expected to be produced with many or few pauses, and long or short

delay; whether it can be expected to be delivered at a slow or fast pace; and

whether a high or low pitch is expected to accompany it. It is therefore worth

speculating that because different speaking tasks appear to have different

prosodic properties associated with them, these differences may contribute to the

contradictory findings in the deception studies summarized in Chapter One, in

Table 1.1 (much more thoroughly described in Vrij, (2000).)

A final word on the usefulness of prosody in forensic linguistics is

somewhat of a caution with respect to pitch data. In this particular study, that

pitch is the most variable prosodic property of all may be an artefact of this

particular corpus, or it may be that from one speaker to the next, pitch varies

more than pause or tempo and is therefore a less reliable speech property by

which to assess speaking tasks, which would be consistent with the findings of

Reilly et al., (2004).

To a lesser extent, I believe this study might also be considered a baseline

for prosodic behaviour in the interview room, in that when asked, none of the

suspects reported or displayed intoxication, and none reported the use of any

psycho pharmaceutical drugs. This is not to say that none of the suspects were

under the influence of narcotics, but if any were, the effects may have been overridden by means taken in the various portions of the study.

6.4   Future research

The present work takes into account what suspects say and how they say it. In recent years, forensic linguists and forensic psychologists have taken strides to uncover the properties of deceptive speech under the condition of police interviews. My goal, in response to that work, has been to step back and take a wide angle snapshot of suspects' vocal behaviour overall, given the relatively limited responses that police questioning elicits. Having hopefully achieved this on a relatively modest number of suspects, I offer the following related foci of study in this section.

As alluded to above, identifying other suspect groups would certainly be a natural next step, given that the two basic groups here exhibited marked differences between them. The issue of age was overlooked in this dissertation with some suspects as young as early 20s and some in their 60s. Socio-economic status was largely controlled, although by no doing of my own. Fewer high-earning individuals seem to find themselves in the interview room and when they do, they are probably not without superior legal representation, so that course of study might prove more difficult to pursue. Furthermore, this dissertation has focused exclusively on speakers of Western Canadian English. Speakers of other varieties of English, and speakers of other languages would

undoubtedly show different prosodic profiles, were parallel studies conducted on a corpus of interviews with such suspects. The groups into which suspects can be categorized, then, show no limits.

The speech of investigating officers has been completely overlooked in this study, although others have taken up this task in discourse analyses and conversation analysis. That I did not explore investigators' speech was, quite simply, by design, owing to the fact that it was not possible to give adequate treatment to both investigators and suspects in a study of this size. Nevertheless, what investigators ask and how they ask it would also be an obvious next step in a robust examination of the language of police interviews. Much the same way that this dissertation took advantage of response categories, investigators prompts too, can be categorized (Hill, 2003), and the same prosodic variables might be examined.

Prosody aside, investigators' contribution to dialogue would make for fascinating study. One notion that comes to mind is the set of strategies used to calm suspects who become agitated. Observationally speaking, an agitated suspect is apt to go off topic, and the investigator's role is to continue to elicit only information relevant to the investigation. Successful calming strategies appear to include the investigator lowering his voice, continued use of the suspect's first name, minimization of what inflamed the suspect and so forth. Investigators also use incremental steps toward eliciting information from

suspects. In other words, an interviewer will seldom start with the most pressing issues, but rather he or she will meander toward the targeted information. Training manuals (Inbau, Reid, & Buckley, 2001, for example) offer explicit instruction on appropriate use of this technique, but a linguistic analysis (i.e., discourse and vocal properties) has thus far not been conducted.

Another area which was not addressed in this dissertation, but that may be of interest to researchers of prosody more so than forensics, is the interaction between the different prosodic elements. Previous studies have found strong correlations between lower speech rates and longer pauses, as well as higher pitch values associated with faster speech, among other complex interactions. The interactions between pause, tempo, and pitch would be worth examining as well, in the interest of a more thorough treatment of suspects' prosody.

All of these other research areas aside, it is clear from the results found in this study that patterns are to be found in suspects' speech, and that they differ according to whether a suspect is new to the interview room, or an old hand, and that the type of information they offer can, in many cases, be produced with particular set of vocal, (i.e., prosodic) properties. Taking this study as a starting point, the outlook for a linguistic contribution to deception detection and the interface between language and emotion is positive.

REFERENCES

Adams, S. (2002). Communication under stress: indicators of veracity and deception in written narratives. Unpublished Ph.D. dissertation: Virginia Polytechnic Institute and state University.

Ainsworth, J. (in press). Curses, Swearing, and Obscene Language in Police-Citizen Interactions: Why Lawyers and Judges should Care. *Serie Informes, Institutue Univesritari de Linguistica Aplicada.*

Anolli, L., & Ciceri, R. (1997). The voice of deception: Vocal strategies of naive and able liars. *Journal of Nonverbal Behaviour, 21*(4), 259-284.

Auer, P. (1996). On the prosody and syntax of turn-continuations. In E. Couper-Kuhlen & M. Selting (Eds.), *Prosody in Conversation* (pp. 57-100). Cambridge: Cambridge University Press.

Baldwin, J. (1993). Police Interview techniques: establishing truth or proof. *British Journal of Criminology, 33*(3), 325-352.

Baldwin, J. (1994). Police Interrogation: What are the rules of the game? In D. Morgan & G. Stephenson (Eds.), *Suspicion and Silence: the Right to Silence in Criminal Investigations* (pp. 66-76). London: Blackstone Press.

Banse, R., & Scherer, K. (1996). Acoustic Profiles on Vocal Emotion Expression. *Journal of Personality and Social Psychology, 70*(3), 614-636.

Barik, H. C. (1977). Cross-linguistic study of temporal characteristics of different speech materials. *Language and Speech, 20,* 116-126.

Baum, K. M. (1998). Perception of Emotion: Measuring Decoding Accuracy of Adult Prosodic Cues Varying in Intensity. *Journal of Nonverbal Behavior, 22*(2), 89-107.

Benneworth, K. (2006). Repertoires of paedophilia: Conflicting descriptions of adult-child sexual relationships in the investigative interview. *The International Journal of Speech, Language and the Law, 13*(2), 189-211.

Blaauw, E. (1995). *On the perceptual classification of spontaneous and read speech.* Utrecht: LED.

Boersma, P., & Weeknink, D. (2006). Praat: doing phonetics by computer (Version 4.4.20).

Bond, C. F., Omar, A., Mahmoud, A., & Bonser, R. (1990). Lie detection across cultures. *Journal of Nonverbal Behavior, 14,* 189-205.

Boomer, D. (1965). Hesitation and grammatical encoding. *Language and Speech, 8,* 148-158.

Breitenstein, C., van Lancker, D., & Daum, I. (2001). The contribution of speech rate and pitch variation to the perception of vocal emotions in a German and an American sample. *Cognition and Emotion, 15*(1), 57-79.

Brown, G., Currie, K., & Kenworthy, J. (1980). *Questions of Intonation.* London: Croom Helm.

Caspers, J. (2003). Local speech melody as a limiting factor in the turn-taking system in Dutch. *Journal of Phonetics, 31,* 251-276.

Chafe, W. (1989). Some things that narratives tell us about the mind. In B. Pelligrini (Ed.), *Narrative Thought and Narrative Language.* Hillsdale: Lawrence Erlbaum.

Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speech. *Cognition, 84,* 73-111.

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

Crystal, D. (1969). *Prosodic Systems and Intonation in English.* London: Cambridge University Press.

Crystal, T., & House, A. S. (1982). Segmental durations in connected speech signals: Preliminary results. *Journal of the Acoustical Society of America, 72,* 1574-1585.

Crystal, T., & House, A. S. (1988). A note on the variability of timing control. *Journal of Speech, Language, and Hearing Research, 31,* 497-502.

Crystal, T., & House, A. S. (1990). Articulation Rate and the duration of syllables and stress groups in connected speech. *Journal of the Acoustical Society of America, 72,* 705-716.

Dankovicova, J. (1997). The domain of articulation rate variation in Czech. *Journal of Phonetics, 25,* 287-312.

Deakin, T. J. (1988). *Police Professionalism: The Renaissance of American Law Enforcement.* Springfield, Illinois: Thomas.

Devillers, L., Vasilescu, I., & Vidrascu, L. (2004, March 23-26, 2004). *F0 and pause featuers for anger and fear detection in real-life spoken dialogs.* Paper presented at the Speech Prosody 2004, Nara, Japan.

Drew, P., & Heritage, J. (1992). Analyzing talk at work: an introduction. In P. Drew & J. Heritage (Eds.), *Talk at Work: Interaction in Institutional Settings* (pp. 3-65). Cambridge: Cambridge University Press.

Eckman, P., Friesen, W. V., & Scherer, K. (1976). Body movement and voice pitch in deceptive interaction. *Semiotica, 16*, 23-27.

Ekman, P. (1988). Lying and nonverbal behaviour: theoretical issues and new findings. *Journal of Nonverbal Behavior, 12*, 163-176.

Ekman, P. (1992). *Telling lies: Clues to deceit in the marketplace, politics and marriage.* New York: Norton.

Ekman, P., O'Sullivan, M., Friesen, W. V., & Scherer, K. (1991). Face, voice, and body in detecting deceit. *Journal of Nonverbal Behavior, 15*, 125-135.

Essen, O. v. (1949). Sprechtempo als Ausdruck psychischen Geschehens. In *Zeitschrift für Phonetik* (pp. 317-341).

Fadden, L. (2006a). *Can't stand the heat: Suspects' attention deflecting strategies during police interviews.* Paper presented at the Northwest Linguistics Conference, Simon Fraser University.

Fadden, L. (2006b). *The prosody of suspects' speech during police interviews.* Paper presented at the Speech Prosody, 3rd International Conference, Dresden, Germany.

Fadden, L. (2007). Quantitative and Qualitative Analyses of Police Interviews with Canadian Aboriginal and Non Aboriginal Suspects. In K. Kredens & S. Goźdź-Roszkowski (Eds.), *Language and the Law: International Outlooks.* Frankfurt am Main: Peter Lang.

Fadden, L. (to appear). *Prosodic profiles: a comparison of first time and repeat offenders.* Paper presented at the 2nd International Association of Forensic Linguistics, Barcelona, Spain.

Fairbanks, G., & Hoaglin, L. W. (1940). An experimental study of the durational characteristics of the voice during the expression of emotion. *Speech Monographs 7*, 85-90.

Fiedler, K., & Walka, I. (1993). Training lie detectors to use nonverbal cues instead of global heuristics. *Human Communication Research, 20*, 199-223.

Finkelman, D., & Grisso, T. (1996). Therapeutic Jurisprudence: from idea to application. In D. B. Wexler & B. J. Winick (Eds.), *Law in a Therapeutic Key* (pp. 587-595). Durham, NC: Carolina Academic Press.

Fisher, R. P., & Geiselman, R. E. (1992). *Memory Enhancing Techniques for Investigative Interviewing. The Cognitive Interview.* Springfield, IL: Thomas.

Fisher, R. P., Geiselman, R. E., & Amador, M. (1989). Field test of cognitive interview: enhancing the recollection of actual victims and witnesses of crime. *Journal of Applied Psychology, 74*, 722-727.

Fonagy, I., & Magdics, K. (1960). Speed of utterance in phrases of different lengths. *Language and Speech, 4,* 179-192.

Ford, C. E., & Thompson, S. A. (1996). Interactional units in convesration: syntactic, intonational, and pragmatic resourses for the management of turns. . In E. Ochs, E. Schegloff & S. Thompson (Eds.), *Interaction and Grammar* (pp. 134-184).

Fox Tree, J. E. (2002). Interpreting Pauses and Ums at Turn Exchanges. *Discourse Processes, 34*(1), 37-55.

Frenkle, R. (1990). Talking in Interviews: a dispreference for patient-initiated questions in physician-patient encounters. In G. Psathas (Ed.), *Interaction Competence* (pp. 231-262). Washington, D.C: University Press of America.

Geiselman, R. E., Fisher, R. P., MacKinnon, D. P., & Holland, H. L. (1985). Eyewitness memory enhancement in the police interview: cognitive retrieval mnemonics versus hypnosis. *Journal of Applied Psychology, 70,* 401-412.

Goldman-Eisler, F. (1956a). The determinants of the rate of speech and their mutual relations. *Journal of Psychsomatic Research, 2,* 137-143.

Goldman-Eisler, F. (1956b). The determinants of the rate of speech output and their mutual relations. *Journal of Psychosomatic Research, 2,* 137-143.

Goldman-Eisler, F. (1961). The distribution of pause durations in speech. *Language and Speech, 4,* 232-231.

Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in Spontaneous Speech.* New York.: Doubleday.

Granhag, P. A., & Strömwall, L. A. (2002). Repeated Interrogations: Verbal and Non-verbal Cues to Deception. *Applied Cognitive Psychology, 16,* 243-257.

Grosjean, F., & Dechamps, A. (1973). Analyze des variables temporelles du francais spontané. *Phonetica, 28,* 191-226.

Grosz, B., & Hirschberg, J. (1992). Some intonational characteristics of discourse structure. In *Proceedings of the International Conference on Spoken Language Processing.* Banff.

Gudjonsson, G. H., Sigurdsson, J. F., Asgeirsdottir, B. B., & Sigfusdottir, I. D. (2006). Custodial interrogation, false confession and individual differences: A national study among Icelandic youth. *Personality and Individual Differences, 41*(1), 49-59.

Gudjonsson, G. H., Sigurdsson, J. F., Asgeirsdottir, B. B., & Sigfusdottir, I. D. (2007). Custodial interrogation: What are the background factors associated with claims of false confession to police? *Journal of Forensic Psychiatry & Psychology, 18*(2), 266-275.

Gudjonsson, G. H., Sigurdsson, J. F., Bragason, O. O., Einarsson, E., & Valdimarsdottir, E. B. (2004). Confessions and denials and the relationship with personality. *Legal and Criminological Psychology, 9*(1), 121-133.

Gudjonsson, G. H., Sigurdsson, J. F., & Einarsson, E. (2004). The role of personality in relation to confessions and denials. *Psychology, Crime & Law, 10*(2), 125-135.

Gudjonsson, G. H., & Williamson, T. (2006). The psychology of interrogations and confessions. In *Investigative interviewing: Rights, research, regulation.* (pp. 123-146). Devon, United Kingdom: Willan Publishing.

Gudjonsson, G. H., Young, S., & Bramham, J. (2007). Interrogative suggestibility in adults diagnosed with attention-deficit hyperativity disorder (ADHD). A potential vulnerability during police questioning. *Personality and Individual Differences, 43*(4), 737-745.

Haworth, K. (2006). The dynamics of power and resistance in police interview discourse. *Discourse & Society, 17*(6), 739-759.

Haworth, K. (in prep). *An Analysis of Police Interview Discourse and its role in the Judicial Process.* University of Nottingham.

Hegedüs, L. (1953). On the problem of pauses in speech. *Acta Linguistica Academiae Scientiarum Hungaricae, 3,* 1-36.

Hellerman, J. (2005). Syntactic and Prosodic Practices for Cohesion in Series of Three-Part Sequences in Classroom Talk. *Research on Language and Social Interaction, 38*(1), 105-130.

Henton, C. G. (1989). Fact and Fiction in the description of female and male pitch. *Language and Communication, 9,* 299-311.

Heydon, G. (2004). Establishing the structure of police evidentiary interviews with suspects. *Speech, Language and the Law, 11*(1), 1350-1771.

Heydon, G. (2005). *The Language of Police Interviewing.* New York: Palgrave MacMillan.

Hieke, A. E., Kowal, S., & O'Connell, D. (1983). The trouble with "articulartory pauses". *Language and Speech, 26,* 203-214.

Hill, M. (2003). Identifying the source of critical details in confessions. *Forensic Linguistics, 10*(1), 1350-1771.

Hird, K., & Kirsner, K. (2002). The Relationship between Prosody and Breathing in Spontaneous Discourse. *Brain and Language, 80*(3), 536-555.

Hirschberg, J., & Grosz, B. (1992). Intonation features of of local and global discourse structure. In *Proceedings of the DARPA Workshop on Spoken Language Systems* (pp. 441-446): Arden House.

Hirst, D. J. (2005). Form and function in the representation of speech prosody. *Speech Communication, 46*, 334-347.

Holmberg, U., & Christianson, S. (2002). Murderers' and sexual offenders' experiences of police interviews and their inclination to admit or deny crimes. *Behavioral Sciences and the Law, 20*, 31-45.

Hudson, A., & Holbrook, A. (1981). A study of the reading fundamental vocal frequency of young black adults. *Journal of Speech and Hearing Research, 24*, 197-201.

Inbau, F. E., Reid, J., & Buckley, J. (2001). *Criminal Interrogation and Confessions* (4th ed.). Gaithersburg: Aspen Publishers.

Jefferson, G. (1984). On stepwise transition from talk about a trouble to inappropriately next-positioned matters. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action: Studies in Conversation Analysis*. Cambridge: Cambridge University Press.

Ji, S. (2002). Identifying Episode Transitions. *Journal of Pragmatics, 34*(9), 1257-1271.

Johnson, A. (2006). Police Questioning. In K. Brown (Ed.), *The Encyclopedia of Language and Linguistics* (2nd ed., Vol. 9, pp. 661-672). Oxford: Elsevier.

Johnstone, T., & Scherer, K. (2000). Vocal communication of emotion. In M. Lewis & J. Haviland-Jones (Eds.), *Handbook of Emotion* (2nd ed., pp. 220-235). New York: Guilford.

Kassin, S. M., Goldstein, C. C., & Savitsky, K. (2003). Behavioral confirmation in the interrogation room: On the dangers of presuming guilt. *Law and Human Behavior, 27*(2), 187-203.

Kassin, S. M., & Gudjonsson, G. H. (2004). The Psychology of Confessions. *Psychological Science in the Public Interest, 5*(2), 33-67.

Kassin, S. M., & Neumann, K. (1997). On the power of confession evidence: an experimental test of the "fundamental difference" hypothesis. *Law and Human Behavior, 21*, 469-484.

Kehrein, R. (2003). The Prosody of Authentic Emotions. *Sprache - Stimme - Gehor, 27*(2), 55-61.

Koreman, J. (2006). Perceived speech rate: the effects of articulation and speaking style in spontaneous speech. *Journal of the Acoustical Society of America, 119*(1), 582-596.

Kowal, S., & O'Connell, D. (1987). Some temporal aspects of stories told while or after watching a film. *Bulletin of Psychonomic Society, 25*(5), 364-366.

Kowal, S., O'Connell, D., & Sabin, E. (1975). Development of Temporal Patterning and Vocal Hesitations in Spontaneous Narratives. *Journal of Psycholinguistic Research, 4*(3), 195-207.

Labov, W. (1972). *The transformation of experience in narrative syntax.* Philadelphia: University of Pennsylvania Press.

Lalljee, M. G., & Cook, M. (1969). An experimental investigation of the function of filled pauses in speech. *Language and Speech, 12,* 24-28.

Larson, J. A. (1932). *Lying and its detection: A study of deception and deception tests.* Chicago: University of Chicago Press.

Laukka, P., Juslin, P., & Bresin, R. (2007). A dimensional approach to vocal expression of emotion. *Cognition and Emotion, 19*(5), 633-653.

Lehiste, I. (1970). *Suprasegmentals.* Cambridge: MIT Press.

Lehiste, I. (1975). The phonetic structure of paragraphs. In A. Cohen & S. Nooteboom (Eds.), *Structure and Process in Speech Perception* (pp. 195-206). Berlin: Springer-Verlag.

Lehiste, I. (1976). Suprasegmental features of speech. In J. Lass (Ed.), *Contemporary Issues in Experimental Phonetics* (pp. 225-239). New York: Academic Press.

Lehiste, I. (1979). Perception of sentence and paragraph boundaries. In B. Lindblom & S. Ohman (Eds.), *Frontiers of Speech Communication Research* (pp. 191-201). London: Academic Press.

Lehiste, I. (1982). Some phonetic characteristics of discourse. *Studia Linguistica, 32*(6), 117-130.

Levin, H., Silverman, I., & Ford, B. (1967). Hesitations in children's speech during explanation and description. *Journal of Verbal Learning and Behavior, 6,* 560-564.

Levy, E. (1999). *Examination of witnesses in criminal cases* (4th ed.). Scarborough, ON: Carswell.

Local, J., Kelly, J., & Wells, B. (1986). Towards a phonology of conversation: Turn-taking in Tyneside English. *Journal of Linguistics, 22,* 411-437.

Local, J., Wells, B., & Sebba, M. (1985). Phonetic aspects of turn delimination in London Jamaican. *Journal of Pragmatics, 9,* 309-330.

Lounsbury, F. G. (1954). Transitional probability, linguistic structure, and systems of habit-family hierarchies. In C. E. Osgood & T. A. Sebeok (Eds.), *Journal of Abnormal and Social Psychology* (Vol. 49, pp. 93-101).

Maclay, H., & Osgood, C. E. (1959). Hesitation Phenomena in Spontaneous English Speech. *Word, 15*(19), 19-44.

Malecot, A., Johnston, R., & Kizziar, P. A. (1983). Syllabic rate and utterance length in French. *Phonetica, 26,* 235-251.

Mann, S., Vrij, A., & Bull, R. (1998). *Telling and detecting true lies.* Paper presented at the Eighth Annual Meeting of the European Association on Psychology and the Law, Cracaw, Poland.

Mann, S., Vrij, A., & Bull, R. (2002). Suspects, Lies, and Videotape: An Analysis of Authentic High-Stake Liars. *Law and Human Behavior, 26(3),* 365-376.

Maynard, D. W. (1991). Interaction and asymmetry in clinical interaction. *American Journal of Sociology, 97,* 448-495.

Miller, J., Grosjean, F., & Lomanto, C. (1984). Articulation rate and its variability in spontaneous speech: a reanalysis and some implications. *Phonetica, 41,* 215-255.

Milne, R., & Bull, R. (1999). *Investigative Interviewing: Psychology and Practice.* Chichester: Wiley.

Newbury, P., & Johnson, A. (2006). Suspects' resistance to constraining and coercive questioning strategies in the police interview. *The International Journal of Speech, Language and the Law, 13(2),* 213-240.

Nilsonne, A. (1987). Acoustic analysis of speech variables during depression and after improvement. *Acta Psychiatrica Scandinavica, 76(3),* 235-245.

Nilsonne, A. (1988). Speech characteristics as indicators of depressive illness. *Acta Psychiatrica Scandinavica, 77(3),* 253-263.

O'Connell, D., & Kowal, S. (1983). Pausology. In W. Sedelow & S. Y. Sedelow (Eds.), *Computers in Language Research 2* (pp. 221-301): Mouton.

Oliveira, M. J. (2000). *Prosodic Features in Spontaneous Narratives.* Unpublished PhD Dissertation, Simon Fraser University.

Osada, N. (2003). Analysis of pause occurrence in three kinds of modified speech: public address, caretaker talk, and foreigner talk. *Journal of Pan-Pacific Association of Applied Linguistics, 7(1),* 77-123.

Pearse, J., Gudjonsson, G. H., Clare, I. C. H., & Rutter, S. (1998). Police interviewing and psychological vulnerabilities: Predicting the likelihood of a confession. *Journal of Community & Applied Social Psychology, 8(1),* 1-21.

Peräkylä, A., & Vehvilänen, S. (2003). Conversation Analysis and the Professional Stocks of Interactional Knowledge. *Discourse & Society 14(6),* 727-750.

Rebollo Couto, L. (1997). Le rythme en espagnol et en portuguese. *Traveaux de l'institut de phonetique de Strasbourg, 27,* 34-37.

Reilly, N., Cannizzaro, M. S., Harel, B. T., & Snyder, P. J. (2004). Feigned depression and feigned sleepiness: a voice acoustical analysis. *Brain and Cognition, 55*, 383-386.

Ressler, R. K., & Burgess, A. W. (1988). Interviewing convicted murderers. In *Sexual Homicide*. New York: Lexington

Robb, M., Maclagan, M., & Chen, Y. (2004). Speaking rates of American and New Zealand varieties of English. *Clinical Linguistics and Phonetics, 18*(1), 1-15.

Rockwell, P. (2000). Lower, Slower, Louder: Vocal Cues of Sarcasm. *Journal of Psycholinguistic Research, 29*(5), 483-495.

Rockwell, P., Buller, D. B., & Burgoon, J. K. (1997). Measurement of deceptive voices: comparing acoustic and perceptual data. *Applied Psycholinguistics, 18*, 471-484.

Sabin, E. (1976). Pause and rate phenomena in adult narrative. *Science, Saint Louis University*(78), 35-53.

Sabin, E., Clemmer, E. J., O'Connell, D., & Kowal, S. (1979). A pausological approach to speech development. In A. W. Siegman & S. Feldstein (Eds.), *Of Time and Speech*. Hillsdale, New Jersey: Lawrence Erlbaum.

Sacks, H., Schegloff, E., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50*(4), 696-735.

Schaffer, D. (1983). The role of intonation as a cue to turn taking in conversation. *Journal of Phonetics, 11*, 243-257.

Schegloff, E. (1996). Turn organization: One intersection of grammar and interaction. In E. Ochs, E. Schegloff & S. Thompson (Eds.), *Interaction and Grammar* (pp. 52-133). Cambridge: Cambridge University Press.

Scherer, K. (1989). Vocal correlates of emotion. In A. Manstead & H. Wagner (Eds.), *Handbook of Psychophysiology: Emotion and Social Behavior* (pp. 165-197). London: Wiley.

Scherer, K. (2003). Vocal Communication of emotion: a review of research paradigms. *Speech Communication, 40*, 227-256.

Scherer, K., Banse, R., & Wallbot, H. (2001). Emotion Inferences from Vocal Expression Correlate across Languages and Cultures. *Journal of Cross-Cultural Psychology, 32*(1), 76-92.

Scherer, K., & Kappas, A. (1988). Primate vocal expression of affective state. In D. Todt, P. Goedeking & D. Symmes (Eds.), *Primate Vocal Communication* (pp. 171-194). Berlin: Springer.

Scherer, K., & Orshinksy, J. S. (1977). Cue utilization in emtion attribution from auditory stimuli. *Motivation and Emotion, 1*, 331-346.

Schleef, E. (2003). Prosody and Narrative Structure in Varieties of Low German and Alemannic. *Journal of German Linguistics, 15*(4), 325-357.

Scollon, R. (1981). Tempo, Density, and Silence: Rhythms of Ordinary Talk. *Center for Cross Cultural Studies. Fairbanks, University of Alaska.*

Selting, M. (1992). Intonation as a contextualization device: case studies on the role of prosody, especially intonation in contextualizing story telling in conversation. In A. Auer & A. D. Luzio (Eds.), *The Contextualization of Language.* Amsterdam: Benjamins.

Selting, M. (1996). On the interplay of syntax and prosody n the consitution of turn-constructional units and turns in conversation. *Pragmatics, 6*(3), 371-388.

Shepherd, E. (1991). Ethical Interviewing. *Policing, 7,* 42-60.

Shipp, T., & Hollien, H. (1972). Speech frequency and duration measures as a function of chronologic age. *Journal of the Acoustical Society of America, 51*(111), 155-159.

Shuy, R. W. (1997). *The Language of Confession, Interrogation, and Deception.* Thousand Oaks: Sage.

Siegman, A. W., & Pope, B. (1966). Ambiguity and verbal fluency in the TAT. *Journal of Consulting Psychology, 30,* 239-245.

Slobogin, C. (1996). Therapeutic Jurisprudence: five dilemmas to ponder. In D. B. Wexler & B. J. Winick (Eds.), *Law in a Therapeutic Key* (pp. 763-793). Durham, NC: Carolina Academic Press.

Sobin, C., & Alpert, M. (1999). Emotion in Speech: The Acoustic Attributes of Fear, Anger, Sadness, and Joy. *Journal of Psycholinguistic Research, 28*(4), 347-365.

Spinos, A.-M. R., O'Connell, D., & Kowal, S. (2002). An Empirical Investigation of Pause Notation. *Pragmatics, 12*(1), 1-9.

Stassen, H., Kuny, S., & Hell, D. (1998). The speech analysis approach to determining onset of improvement under antidepressants. *European Neuropsychopharmacology, 8,* 303-310.

Stenström, A.-B. (1984). *Questions and Responses in English Conversation.* Lund: Lund University Press.

Streeter, L. A., Krausse, R. M., Geller, V., Olson, C., & Apple, W. (1977). Pitch changes during attempted deception. *Journal of Personality and Social Psychology, 24,* 12-21.

Swerts, M. (1997). Prosodic features at discourse boundaries of different strength. *Journal of the Acoustical Society of America, 101*(1), 514-521.

Swerts, M., & Geluykens, R. (1994). Prosody as a marker of information flow in spoken discourse. *Language and Speech, 37,* 21-43.

Tauroza, S., & Allison, D. (1990). Speech Rates in British English. *Applied Linguistics, 11,* 90-105.

ten Have, P. (1991). Talk and institution: A reconsideration of the 'asymmetry' of doctor-patient interaction. In D. Boden & D. Zimmerman (Eds.), *Talk and Social Structure.* Cambridge: Polity Press.

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

Tsao, Y.-C., & Weismer, G. (1997). Interspeaker variation in habitual speaking rate: Evidence for a neuromuscular component. *Journal of Speech, Language, and Hearing Research, 40,* 858-866.

Tsao, Y.-C., Weismer, G., & Iqbal, K. (2006). Interspeaker Variation in Habitual Speaking Rate: Additional Evidence. *Journal of Speech, Language, and Hearing Research, 49,* 1156-1164.

Uhmann, S. (1992). Contextualizing relevance: on some forms and functions of speech rate changes in everyday conversation. In A. Auer & A. D. Luzio (Eds.), *The Contextualization of Language* (pp. 297-336). Amsterdam: Benjamins.

van Bezooijen, R. (1984). *Characteristics and recognizibility of vocal expression of emotion.* Dordrecht: Foris Publications.

van Donzel, M. (1999). *Prosodic aspects of information structure in discourse.* The Hague: Holland Academic Graphics.

van Heuvan, V., & van Zanten, E. (2005). Speech rate as a secondary rosodic chariteristic of polarity questions in three languages. *Speech Communication, 47,* 87-99.

van Heuven, V., & van Zanten, E. (2005). Speech rate as a secondary Prosodic characteristic of polarity questions in three languages. *Speech Communication, 47,* 87-99.

Vrij, A. (1995). Behaviour correlates of deception in a simulated police interview. *Journal of Psychology: Interdisciplinary and Applied, 129,* 15-29.

Vrij, A. (2000). *Detecting lies and deceit: The psychology of lying and the implications for professional practice.* Chichester, UK: Wiley.

Vrij, A., Edward, K., & Bull, R. (2001). Stereotypical Verbal and Nonverbal Responses while Deceiving Others. *Personality and Social Psychology Bulletin, 27*(7), 899-909.

Vrij, A., & Mann, S. (2001). Telling and Detecting Lies in a High-stake Situation: the Case of a Convicted Murderer. *Applied Cognitive Psychology, 15,* 187-203.

Vrij, A., & Winkel, F. W. (1991). Cultural patters in Dutch and Surinam nonverbal behaviour: an analysis of simulated police/citizen encounters. *Journal of Nonverbal Behavior, 15,* 169-184.

Walker, J. (1781). *Elements of Elocution* (Vol. English Linguistics 1500-1800, No 148.): Reprinted by The Scholar Press, Menston (1969). .

Walker, J., & Archibald, L. (2006). Articulation rate in pre-school children: a 3-year longitudinal study. *International Journal of Language and Communication Disorders, 41*(5), 541-565.

Walton, D. (2003). The interrogation as a type of dialogue. *Journal of Pragmatics, 35,* 1771-1802.

Wells, B., & Peppé, S. (1996). Ending up in Ulster: prosody and turn-taking in English Dialects. In E. Couper-Kuhlen & M. Selting (Eds.), *Prosody in Converstation* (pp. 101-130). Cambridge: Cambridge University Press.

Wells, G. (1993). Reevaluating the IRF sequence: A proposal for the articulation of theories of activity and discourse for the analysis of teaching and learning in the classroom. *Linguistics in Education, 5,* 1–37.

Wennerstrom, A. (2001). *The Music of Everyday Speech*: Oxford University Press.

Williamson, T. M. (1993). From interrogation to investigative interviewing: strategic trends in police questioning. *Journal of Community & Applied Social Psychology, 3,* 89-99.

Wrightsman, L. S., & Kassin, S. M. (1993). *Confessions in the Courtroom*. Thousand Oaks, CA: Sage.

Ylanne-McEwen, V. (2004). Shifting Alignment and Negotiating Sociality in Travel Agency Discourse. *Discourse Studies, 6*(4), 517-536.

Zuraidah, M. D., & Knowles, G. (2006). Prosody and turn-taking in Malay broadcast interviews. *Journal of Pragmatics, 38,* 490-512.