Journal of Literature, Languages and Linguistics
ISSN 2422-8435 An International Peer-reviewed Journal
Vol.29, 2016

The Role of Forensic Phonetics in Legal Investigation: A Case Study of Two Speaker-Identified/Unidentified Recorded Samples

Bandar Khalil Mukattash

English Language and Literature Department, Al Balqa Applied University, PO box 19117, Al-Salt, Jordan

Abstract

The study sheds light on the growing role forensic phonetics has recently been given in legal investigation, particularly in cases of speaker identification and identity authentication. Though this field of study is relatively new, strenuous efforts have been exerted by its founders, namely John Baldwin, Peter French and Francis Nolan among others, to come up with a concrete body of principles and methodologies on which the phonetic analysis of samples under legal investigation could be built. To prove this, the study is divided into two parts. Part one traces back the study of forensic phonetics to its origins with special emphasis given to its major methods of application: analysis of recordings, speaker profiling, speaker identification and construction of voice line-ups in ear-witness testimony. The second part of the study is based on a case study of two recorded samples, one for an unknown person speaking on the phone and the other for an identified person recorded in prison. By analyzing the different occurrences of a specific word (believe) in the two samples using Praat, the study aims to show the degree to which forensic phonetics can be accurate in solving cases of speaker identification and identity authentication under legal consideration.

1. Introduction

Nearly every issue in daily life has a legal aspect to consider. Whether in courts, police stations, hospitals, the media, different cases of murder, burglary, fraud, etc. have led to investigations where efforts have been directed to uncover the real identity of the suspect, yet failed due to the lack of established parameters and methodologies integrating linguistics with the law (Siegel 78). To this end, the study of forensic linguistics has been identified by experts dealing with cases where the two fields of knowledge (language and the law) intersect. In a review of Coulthard and Johnson's (2007) introduction to the study of forensic linguistics, Turell (2008) identifies the discipline of forensic linguistics as "a well-defined, well established discipline which seeks to uncover and establish the existing interplay between linguistic and legal issues." (1) For Turell, the significance of forensic linguistics lies in its multidisciplinary nature which not only helps integrate two related fields of study but also helps consolidate them "from the point of view of both research and teaching programmes." (1)

The study of forensic linguistics is relatively new. It dates back to the fifties of the twentieth century. Coulthard and Johnson (2007) note that the term 'forensic English' was first used by F.A. Philbrick (1949) as a title for a book he wrote on a legal approach to the English language, *Language and the Law: The Semantics of Forensic English*. Twenty years later, the same term was used by Jan Svartvik (1968) in a publication titled *The Evans Statements: A Case for Forensic Linguistics*, which marked the public birth of the study of forensic linguistics as a discipline integrating the study of language and the law. Svartvik's point behind this work was to prove that the disputed parts of the suspect's (Evans) confessions were not linguistically identical with the non-disputed parts (18-22). From this point on, forensic linguistics has served as the site where efforts of linguists and law specialists, crime specialists, lawyers, policemen and judges have united to uncover controversial aspects of cases where suspicion is hard to solve by traditional means of investigation.

One branch of forensic linguistics is forensic phonetics. Forensic phonetics draws on the applications of phonetics, the scientific study of speech, in court cases. Whether it is a case of speaker recognition, tape authentification, disputed/non-disputed utterances, the police and legal bodies have always sought the expertise of phonetic practitioners and researchers to provide them with evidence on the different cases. This along with the speeding technological development our age is witnessing has increased the number of legal cases dealing with recorded speech, whether over the phone, mobile or chatting rooms, and necessitated, as a result, the need for an established study integrating phonetics and law.

II.A brief review of literature:

The establishment of the IAFL (International Association of Forensic Linguistics) in the last decade of the twentieth century has led to the establishment of other associations specializing in the use of language in legal and court cases. Now under the name of the International Association of Forensic Phonetics and Acoustics, IAFL was founded in 1991 to examine the role phonetic analysis plays in evaluating recorded speech in legal cases. The efforts of its founders, John Baldwin and Peter French (1990), Francis Nolan (1991) and others, have given rise to the study of forensic phonetics.

Since its early emergence as a distinct study in the 1990s, forensic phonetics has dealt with court cases addressing cases such as speaker identification, tape authentication and disputes in speech. Nevertheless, the lack of common principles and methodologies have led to numerous controversies among forensic phonetics experts. Realizing the serious outcomes those controversies could lead to, Nolan (1991) insists, in "Notes and discussion: Forensic phonetics," that phonetic experts doing forensic analysis work within an integrated framework of well-defined parameters and methodological approaches to be able to put their phonetic expertise to correct use in legal cases (12-13). Phoneticians have always been involved in analyzing, comparing and contrasting evidence in court cases, but their involvement, in his words, ranged from:

> the occasional appearance of individual dialectologists in court giving evidence based largely on accents, to … a fairly well established body of practitioners and researchers, on a more broadly-based set of analyses including acoustic measurements. (Abstract, 2)

Nolan's comment shows that phoneticians' and dialectologists' involvement in legal cases ranks secondary to that of detectives and police officers. Without a concrete body of principles and methodologies to base their analyses on, their efforts could never give results accurate enough to be taken as evidence in cases of speaker identification and identity authentication.

In an attempt to evaluate the role forensic phonetic has played in advancing the legal process in the UK, Nolan (1991) refers to Baldwin and French's (1990) *Forensic phonetics*, which is the first publication describing the applied methodology and the theoretical background by which forensic phonetics works (484). Designed to introduce the study of forensic phonetics to phoneticians, the book assigns a new role to phoneticians: that of detectives. When conducted by experts in forensic studies, phonetic analysis aims to examine and evaluate the phonetic features in two or more speech samples (485), such as: stylistic differences between the samples, the way variable features are quantified, etc. Their discussion, however, shows that the two authors seem to disagree on two controversial points: whether human voice is unique, and whether or not auditory phonetics and acoustic phonetics should be integrated in forensic phonetics.

The two authors adopt different views as to the employment of human voice in proving cases of identity authentication. Baldwin sees that human voice is as unique as his DNA. French, however, chooses not to build much hope on voice uniqueness as evidence in legal cases (486). The reason can be ascribed to the fact that current technical sources have fallen short from detecting the uniqueness of human voice and this has relegated it to a posterior position compared to other types of evidence. Here, Nolan refrains from adopting either of the two authors' standpoints, emphasizing the fact that Baldwin's view is only a hypothesis, and cannot "be treated as a factual basis for practical work in forensic." (487) Similarly, the two authors differ as to the possibility of integrating the two aspects of phonetics (auditory and acoustic) in forensic work. Baldwin believes that the auditory approach is sufficient for forensic analysis. French, on the other hand, opts for a more integrated approach to forensic phonetics. Nolan's standpoint in this regard is clearer than its predecessor (the first point of disagreement); he disagrees with Baldwin, but agrees with French (487). Though he does not deny the role of auditory phonetics in offering a segmental and prosodic description of the recorded data, he underscores the potential role of acoustic phonetics in perfecting phonetic analysis. Acoustic phonetics helps the forensic phonetician accomplish his task, in the sense that it provides him with a quantitative description and complementary information on the recorded data (488-89). Using acoustic instruments, the acoustic features of the recorded samples, such as formant frequencies, durations and fundamental frequency, can be measured, analyzed and compared.

Nolan adopts a more integrated approach to forensic phonetics which argues that auditory and acoustic analysis "should proceed interactively rather than successively." (488) To prove this point, he (1990) refers to a study conducted on telephone recordings where the analysis showed that while the analyzed vowels were not different in phonetic quality, they showed different formant structure: the recording of the fraudster showed a strong formant (2.5 kHz), while the defendant's samples showed a weak one (2 kHz and 3 kHz)). These findings have

led Nolan to question the reliability of auditory speaker identification in forensic phonetics. In many cases of forensic investigations, the role auditory phonetics plays has been overtly overestimated and this has resulted in unreliable findings for court consideration (492). To this end, Nolan warns against complete reliance on auditory analysis in the area of speaker identification.

As reliance on phonetic approaches to speech has evidently increased nowadays, there arose the need to turn forensic phonetics into a method-based practice that renders results with a high degree of credibility to be taken as evidence in legal cases. Foulkes and French (1990) note that different methods/applications are used nowadays by forensic phoneticians to examine speech recordings in court cases, the most important of which are the following four: deciphering the content of difficult recordings; speaker profiling; speaker identification; constructing voice line-ups to evaluate ear-witness testimony (1).

Tape authentification, the first phonetic application, is problematic and unreliable. Though a recorded tape may contain some evidence to a specific case, the difficulty faced by experts deciphering the recorded information decreases its value. Foulkes and French argue that the tape-recorded information may be disputed by external distractions. Acoustic and auditory phonetics may help in this case by transcribing the spoken conversation and scanning it for evidence; nevertheless, the attempt might be hindered by serious external noises in tape recording. Tape-recordings may be also difficult to analyze because of the different patterns of pronunciation different speakers use. In a study French (1990) conducted, a doctor-patient conversation has been analyzed to investigate in a case where a Greek doctor, living in the UK, has been accused of prescribing drugs. The different accent of the Greek doctor made it difficult for acoustic analysts to transcribe the recorded conversation. Even though the working team's comparison of the doctor's disputed word with the undisputed 'can' and 'can't' yielded close readings, they agreed, upon analyzing the vowel frequencies, that the disputed word matched the format patterns for 'can't' (1).

The second application in forensic phonetics, speaker profiling, is used in cases of kidnapping and serial killings, where information about the speaker (speaker's age, sex, social background, etc.) is needed to identify the suspect. In a case where tape-recordings were analyzed to identify the place where one serial killer lived, Stanely Ellis' (1994) analysis showed that the man lived in a particular region, but was changed later as Ellis realized that the recordings were sent by a hoaxer, not the criminal himself (Hickey 2015, 193).

Speaker identification (SID) is the third and most frequently used application in legal cases that involve speech recordings. Foulkes and French remark that "70 percent and 80

percent of all forensic cases involve SID," (2) and are thus conducted on phonetic basis. Speaker recognition is used in cases where forensic phoneticians need to identify a voice in a criminal recording, such as a death threat or a hoax call or a surveillance recording made by the police to identify a suspect. The recordings are analyzed by means of comparative phonetic tests to recognize who the recorded voice belongs to. This is usually done by comparing the recorded voice in the recorded data with a sample from the suspect, and deciding whether, and in what ways, the two recorded samples are similar or different. Nevertheless, the speaker auditory identification can be misleading since, as Foulkes and French (1900) remark, the individual's voice is not fixed as his DNA or imprint are but depends on a number of factors such as sex, age, social background, etc., and changes in accordance with the situation the individual is involved in as well. For instance, when the person is stressed, or speaking on the phone, his tone and average frequency rise to meet the demands of the situation. For this reason, recent SID analysis has been designed to encompass auditory and acoustic aspects, even though, as Nolan (1991) notes, forensic phoneticians have disagreed whether acoustic analysis and auditory analysis yield the same results in identifying speakers and locating suspects in extracts of recorded speech. Foulkes and French share this opinion as well, emphasizing the fact that the result of SID analysis is not a factual statement, but rather "an expression of opinion ranged on a scale of confidence." (3)

Speaker identification by lay witnesses, the fourth phonetic application used in forensic phonetics, is used in cases where there is no concrete recording of the criminal's/suspect's voice; instead, the criminal/suspect is only heard by some witness.

The most reliable way to test the recall of voices by the witness is by using a voice parade or line-up, which is a set of recorded voices encompassing that of the suspect's. Forensic phoneticians are usually asked to construct these parades and to analyze them to make sure that nothing stands out in the set of recorded voices. Nolan and Grabe (1996), for example, have described in detail a voice parade to be used in cases of rape (74-94). Of course, one should not forget the shortcoming of this method in testing witnesses' recalls. Foulkes and French (1990) draw the reader's attention to this point by arguing that recalling heard voices is affected by the witness' hearing

ability, age, and his familiarity with the voice he heard.

III.Two Cases of Speaker Identified and Unidentified Recorded Samples:

A) Statement of the task required:

Two recorded samples, one for an unknown person speaking on the phone, and the other for a specific person recorded in prison, will be compared to prove whether or not they belong to the same speaker.

B) Brief description of the material:

The first recording, which was recorded in the police station, is 5 minutes 53 seconds long and has the pitch 103.25005911024596. It will be referred to as the known sample (K). The second sample, a recording on a mobile phone, is 1 minute 10 seconds long and has the pitch 123.23205431900911 Hz. It will be referred to as the unknown sample (U). The quality in both recordings seems to be clear though there are some words in the unknown sample which are not as clear as the words in the known sample.

The two samples will be analyzed using Praat with special emphasis given to one specific word (believe) which occurs in the two samples. The different occurrences of the word will be compared by getting the formants for the start of the long vowel and the formants for the end of the long vowel for each of them, which will then be represented in a chart for the start and the end of the vowel to be able to make a better comparison. After looking at the results in both charts, a conclusion will be made as to whether the two samples belong to the same speaker or not.

C) Analysis:

The phonetic transcription of the word (believe) /bili:v/ shows two vowel sounds: /i/ and /i:/. The long vowel sound shares the phonetic features of frontness and tongue height with the short vowel sound in the word, yet analyzing the frequency formants of the long vowel sound, rather than the short vowel sound,would help more in arriving at a more accurate answer as to whether the two recorded samples belong to the same speaker. The reason can be ascribed to the feature of length that marks the vowel sound /i:/ compared to its short counterpart. First, the feature of tongue position, which produces a higher rate of formant frequency in high vowels compared to low vowels, will show more clearly in a long, high vowel sound like /i:/ than in a short, long vowel sound. Second, the feature of frontness/backness of the tongue, which gives a higher rate of formant frequency in front vowels compared to back vowels, will likewise appear more clearly in a long, front vowel like /i:/ than in a short, front vowel.

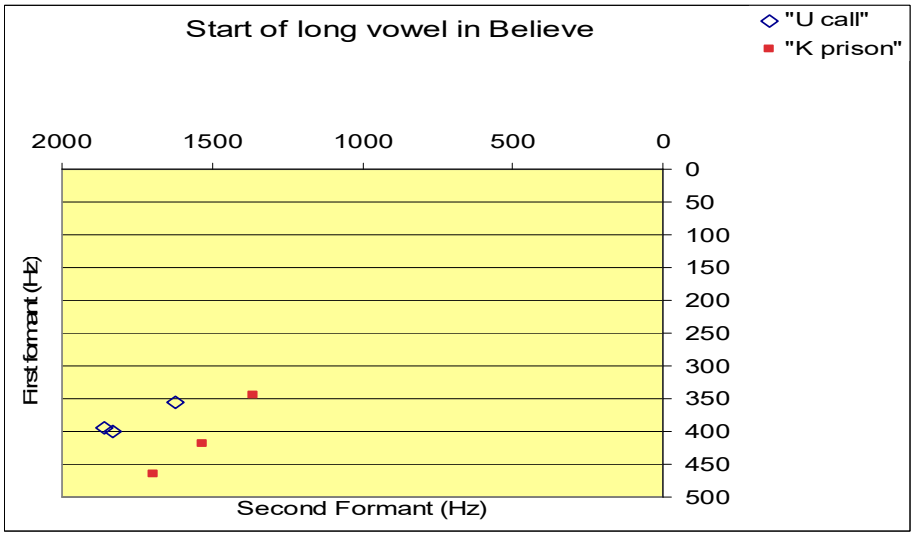
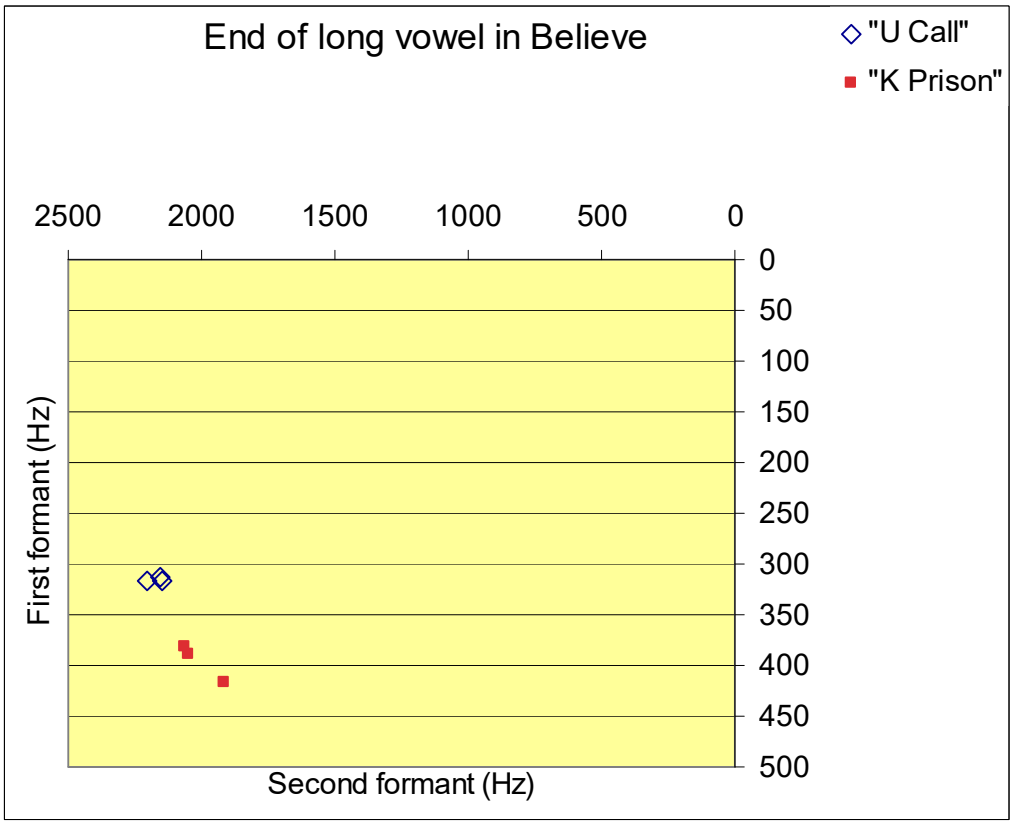
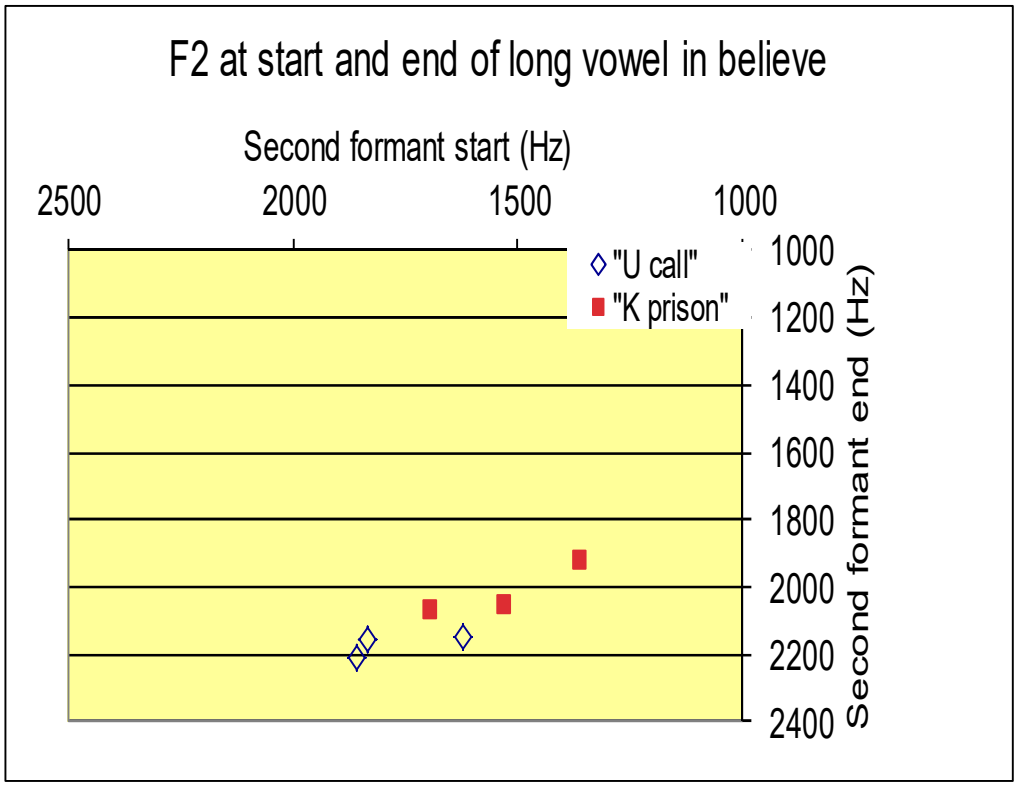


Chart one shows an analysis of the first and second formants at the start of the long vowel /i:/ in the two recorded samples. The first formant of the known sample started with a low rate of frequency (300Hz) and reached almost (500Hz). The start of the long vowel in the unknown sample marked a frequency rate close to that of the known sample in the first two readings of the first formant, respectively 360 Hz and 400 Hz, though the reading shows some difference between this sample and the other, respectively 410 Hz and 460 Hz. As for the frequency rate of the second formant, the readings of both samples show a relatively high rate which is expected with front vowel sounds. The known sample records almost 1650 Hz and even though the unknown sample recorded a little higher rate, 1850 Hz, the two readings are close enough to tell that both samples belong to the same speaker.



The second chart gives findings which support the same conclusion drawn from the first chart. The readings of the first formant at the end of the long vowel /i:/ in the two samples are relatively close, though not with the same degree that is noticed at the start of the long vowel in the first chart. While the frequency rate of the first formant of the known sample ranged between 400 Hz and 450 Hz, that of the unknown sample recorded a lower rate ranging between 310 Hz and 330 Hz. This finding is also clear in the frequency rates recorded in both samples for the second formant. Here, however, the range of the frequency rates for the two samples is not as extensive as the others; thus while the second formant of the known sample ranged between 1700 Hz and 2000 Hz, it ranged between 2000 Hz and 2200Hz in the unknown sample. It is also worth mentioning that these readings, though slightly different between the two samples, fit within the expected frequency rates of both formants for a front high long vowel.



The third chart combines the readings of frequency rates of the second formant at the start and the end of the long vowel sound in both samples. As can be seen, the frequency rate of the known sample ranges between 1900 Hz and 2100 Hz. The unknown samples has a relatively higher frequency rate ranging between 2100 Hz and 2200 Hz, yet it remains close enough to the readings in the known sample, indicating for a third time that the possibility of ascribing both samples to the same speaker is the most probable conclusion to draw.

D) Conclusion in terms of consistency and distinctiveness:

As shown in the charts, the formants in both the K sample and the U sample are close to each other, and this indicates that there is a close similarity in the pronunciation of the vowel at the start and the end of the word in the two samples. The second chart shows a similarity between the samples at the end of the vowel as well. Moreover, the final chart proves this point by showing a very close relation between the two samples in the first formants, which also indicates they might belong to the same speaker.

Taking into account that the pitch in both recordings is close falling between (103 to 123) and by studying the formants in both the K sample and the U sample, the comparison between the two samples can be given a moderately distinctive result from a scale of 5 (where 1 is a not distinctive and 5 is exceptionally distinctive), which shows more probability that the two samples belong to the same speaker than to two different speakers, even though a definite result, as most forensic phoneticians agree, is hard to render.

V. Conclusion:

It can be noticed from the account and the analysis above that forensic phonetics plays an integral role in providing and evaluating evidence for legal cases which involve recorded data. Different methods are used to examine the available data; nonetheless, all methods are channeled to achieve the same ends. In this study, the two samples have shown similarity in terms of formant structure. Yet, since no current proof has yet been advanced as to the uniqueness of voiceprints for every individual, an absolute result is out of question at this level; what we can arrive at is a relative result comparing the two samples in question. Forensic phoneticians are

not mathematicians. They are analysts, and their analysis is not supposed to fit one specific category. It is as Foulkes and French remark in their article on forensic phonetics and sociolinguistics, "not a factual statement, but rather an expression of opinion ranged on a scale of confidence." (3)

Bibliography

Baldwin, J. & French, P. (1990). *Forensic Phonetics*. London: Pinter Publishers.

Coulthard, M. & Johnson, A. (2007) *An Introduction to Forensic Linguistics: Language in Evidence*, New York: Routledge.

Dallasarra, C. et al. (2010). Speaker identification: Effects of noise, telephone, bandwidth, and word count on accuracy. *Journal of the Acoustical Society of America*. 138, 2393.

Ellis, S. (1994). The Yorkshire Ripper enquiry: Part I. *International Journal of Speech Language and the Law* 1.2, 197-206.

Foulkes, P. & French, P. (1990) "Forensic Phonetics and Sociolinguistics". *Sunandaa*, 1-4. http://_www-users.york.ac.uk/~pf11/foulkes&**french**.pdf

Hollien, H. (1977). Authenticating tape recordings. In R.W.de Core & J.S.Jackson (Eds.), *Proceed Carnahan Conference: Crime Countermeasures*. Lexington: KY, 19-23.

---. (1990). *Acoustics of Crime*. New York: Plenum.

---. (1992). Noisy tape recordings in forensics. *ESCA Proceed, Speech Processing in Adverse Conditions*, 167-170.

Jillian, H (2013). *The Acoustics of Crime: The New science of forensic phonetics*. New York: Plenum.

Mobile, B.C. et al. (2009). Parametric study on speaker identification biometric system using dormant analysis. *Journal of the Acoustical Society of America*. 135, 2530-2535.

Nolan, F. (1991) "Notes and discussion: Forensic phonetics". *J. Linguistics*, 483-493. http://www.jstor.org/stable/4176127

---. (2003). A recent voice parade. *Forensic Linguistics* 10, 277-291.

Nolan, F. & Grabe, E. (1996). Preparing a voice lineup. *International Journal of Speech Language and the Law* 3.1, 74-94.

Rose, P. (2003). *Forensic Speaker Identification*. London: Taylor and Francis.

Olsson, J. (2008). *Forensic Linguistics: An Introduction to Language, Crime and the Law*. London: Bloomsbury Academic.

Turell, T. (2008) Review of An Introduction to Forensic Linguistics: Language in Evidence. *Atlantis*, 155-160.