The International
Journal of
Speech,
Language
and the Law

IJSLL (PRINT) ISSN 1748-8885
IJSLL (ONLINE) ISSN 1748-8893

Notices

PhD Abstracts

Forensic automatic speaker recognition using Bayesian interpretation and statistical compensation for mismatched conditions

Anil Alexander
Senior Research Engineer

Clarifying Technologies Ltd.,
Unit 31, *Ddole Road*, *Enterprise Park*,
Llandrindod Wells, *Powys*, *LD1 6DF*, *UK*
anila@clarifyingtech.com

Awarding Institution: Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland

Date of Award: November 2005

KEYWORDS: AUTOMATIC SPEAKER RECOGNITION, STRENGTH OF EVIDENCE, MISMATCHED RECORDING CONDITIONS, STATISTICAL COMPENSATION TECHNIQUES

State-of-the-art automatic speaker recognition systems show very good performance in discriminating between voices of speakers under controlled recording conditions. However, the conditions in which recordings are made in investigative activities (e.g., anonymous calls and wire-tapping) cannot be controlled and pose a challenge to automatic speaker recognition. Differences in the telephone handset, in the transmission channel and in the recording devices can introduce variability over and above that of the voices in the recordings. The strength of evidence, estimated using statistical models of within-source variability and between-source variability, is expressed as a 'likelihood ratio'. The likelihood ratio is estimated using a probabilistic Bayesian interpretation which gives the probability of observing the features of the questioned recording in the statistical model of the suspected speaker's voice, given two competing hypotheses: first, that the suspected speaker is the same speaker as that on the questioned recording, and second, that the speaker heard on

 doi : 10.1558/ijsll.v14i1.145

the questioned recording is not the suspected speaker. The main unresolved problem in forensic automatic speaker recognition today is that of handling mismatch or differences in recording conditions. As such, mismatch in recording conditions has to be considered when estimating the likelihood ratio.

The research in this thesis mainly addresses the problem of the erroneous estimation of the strength of evidence due to mismatch in technical conditions of encoding, transmission and recording of the suspected speaker reference and control databases, as well as the relevant potential population databases used in a Bayesian interpretation framework. Three main directions in applying the Bayesian interpretation framework to forensic automatic speaker recognition casework are investigated. The first addresses the problem of mismatched recording conditions of the databases used in the analysis. The second concerns introducing the Bayesian interpretation methodology to aural-perceptual speaker recognition. It also involves comparing aural-perceptual tests performed by laypersons, with an automatic speaker recognition system in matched and mismatched recording conditions. The third addresses the problem of variability in estimating the likelihood ratio and several new solutions to cope with this variability are proposed.

As mentioned above, firstly, a new approach is proposed to estimate and statistically compensate for the effects of mismatched recording conditions. It uses databases recorded in different conditions to estimate parameters for scaling distributions to compensate for mismatch, called 'scaling databases'. The estimated condition-specific compensation parameters can be applied to larger databases in mismatched conditions. These scaling databases thus reduce the need for recording large databases for potential populations in each recording condition, which is both expensive and time consuming. This method uses the principal Gaussian components in the distributions pertaining to different recording conditions to estimate the value of the compensation parameters. The error in the likelihood ratios obtained after the compensation used in this approach increases with the deviation of the score distributions from the normal distribution. Detecting and compensating mismatches between databases helps to more accurately reflect the similarity or dissimilarity of the voices in the recordings. Guidelines are proposed for the creation of a database that can be used in order to estimate and compensate for mismatch, and a prototype of this database is created to validate the methodology.

Further, to investigate the second direction, the effect of mismatched recording conditions on the strength of evidence is analyzed using both aural-perceptual and automatic speaker recognition methods. The Bayesian interpretation methodology is introduced and applied to aural-perceptual speaker recognition from which likelihood ratios can be estimated. The aural-perceptual speaker recognition was performed by 90 phonetically untrained subjects. It was experimentally observed that in matched recording conditions of suspect and

questioned recordings, the automatic systems showed better performance than the aural recognition systems. In mismatched conditions, however, the baseline automatic systems showed a comparable or slightly degraded performance as compared to the aural recognition systems. Adapting feature extraction and modeling in the baseline automatic system to make it suitable for mismatched conditions resulted in comparable or better performance than aural recognition in the same conditions.

Finally, to address the third direction, the application of Bayesian interpretation to real forensic case analysis, several new solutions are proposed for the analysis of the variability of the strength of evidence for handling cases where the suspect data are limited. These include using bootstrapping techniques, statistical significance testing and confidence intervals, and multivariate extensions of the likelihood ratio.

In order for forensic automatic speaker recognition to be acceptable for presentation in the courts, the methodologies and techniques have to be researched, tested and evaluated for error, as well as be generally accepted in the scientific community. The methodology presented in this thesis is viewed in light of the criteria set out in the Daubert (USA, 1993) ruling for the admissibility of scientific evidence.

'Speak English or what?' Codeswitching and interpreter use in New York Small Claims Courts

Philipp Sebastian Angermeyer
Assistant Professor

Department of Languages, Literatures and Linguistics, York University,
S561 Ross Building, 4700 Keele Street,
Toronto, Ontario, Canada, M3J 1P3
philipp.angermeyer@nyu.edu

Awarding Institution: New York University, USA
Date of Award: May 2006

KEYWORDS: SOCIOLINGUISTICS, CODE-SWITCHING, COURT INTERPRETING, ARBITRATION, LANGUAGE IDEOLOGIES

This thesis investigates the language choices of individuals with limited English skills who participate in informal court proceedings in New York City, as well as the language use of the interpreters who assist them. Drawing on sociolinguistic and ethnographic fieldwork in three Small Claims Courts and focusing on

speakers of Haitian Creole, Polish, Russian, and Spanish, the study analyzes transcripts of recorded arbitration hearings, bringing together the fields of interactional sociolinguistics, codeswitching, and translation studies, combined with an anthropological perspective on language and law.

All interpreter-assisted litigants observed in this study (67 participants in 40 transcribed proceedings) are found to use some English in addition to speaking through an interpreter, ranging from occasional codeswitching of short phrases or insertion of English lexical items to codemixing and even predominant use of English. Comparing speakers from different linguistic dyads in a setting where the social dominance of English is institutionally enforced, the study presents a cross-linguistic analysis of codeswitching that ties microsociolinguistic phenomena of language use and interaction to the macrosociolinguistic conditions of the linguistic market (cf. Bourdieu 1991). While community-specific codeswitching patterns exist, speakers of all four languages investigated use English in ways that demonstrate their participation in the English part of the interaction and suggest accommodation to English-speaking participants (cf. Giles et al. 1991). For example, insertions of English lexical items in other language structures are often lexical repetitions of items used previously by English speakers, establishing coherence across turns made in different languages.

In contrast to earlier studies of court interpreting (Berk-Seligson 1990, Hale 2004), this thesis questions the assumption that equality before the law can be achieved by competent translation and argues that individuals who do not speak the language of the court are inherently disadvantaged. In particular, it describes several ways in which interpreter-mediated testimony is inferior to testimony given in the language of the court and it identifies language-related practices of courtroom procedure that contribute to this disadvantage. For example, the testimony of non-English speakers is always translated in consecutive mode. This causes narratives to be fragmented and leads to frequent interruptions by other participants, because translation pauses are misunderstood as turn transitions, especially by impatient arbitrators who perceive interpreted testimony as 'taking too long'. By contrast, English speakers are less likely to be interrupted during their testimony, as they typically do not pause but force the interpreter into simultaneous interpreting mode instead. Simultaneous interpreting places a higher demand on the interpreters, especially if several people speak at the same time, often making it impossible for them to translate everything that is being said in English, thus leading to a loss of information for the non-English speaker.

In addition, judges, arbitrators, and interpreters routinely discourage or prevent immigrant litigants from using English at all once an interpreter is present. It is argued that by doing so they not only prevent litigants from acquiring

symbolic capital (cf. Bourdieu 1991), but they also impede communication in the courtroom. It is shown in this thesis that some litigants are fully capable of presenting coherent testimony in English and require the interpreter's assistance mainly to ensure full comprehension of other people's talk. The practice of discouraging their use of English is based on an undifferentiated view of language proficiency that does not distinguish production and comprehension, and which is shown to also rest on a monological language ideology that favors native language use irrespective of the interactional context and which disregards various principles of intercultural communication (cf. Coulmas 1987). Moreover, when immigrant litigants do use an interpreter, they are often accused by an opposing party of 'pretending' not to speak English, showing that their language choice is always open to criticism, no matter which language they speak. As such, language choice and its social connotations may enter into so-called demeanor evidence (cf. Conley 1982), potentially influencing the legal decision-making process, and thus further disadvantaging immigrants compared to native speakers of English for whom this factor is not relevant.

Finally, the institutional norms for court interpreting require interpreters to translate 'verbatim', i.e. to speak in the voice of the person whose talk they translate. Drawing on analyses of transcription excerpts, the study shows that strict adherence to this rule frequently leads to miscommunication in arbitration hearings, as it leaves interpreters unable to identify speaker and hearer roles when these cannot be inferred from the context (cf. Goffman 1981, Wadensjö 1998). Some interpreters are found to avoid such misunderstandings by adopting a translation style that accommodates to immigrant litigants, for example by using reported speech when translating from English. In a quantitative analysis of reference choice in translation, the study identifies several factors that constrain this variation in interpreting style and relates the findings to differences in the interpreters' understanding of their own role in the proceedings (see also Angermeyer 2005a, 2005b).

As an ethnograpy of Small Claims Courts in New York City and with a detailed analysis of the arbitration hearing as interactional genre, this thesis also contributes to research on 'informal justice', Alternative Dispute Resolution, and encounters of laypeople with the US legal system (Abel 1982, Merry 1990, Conley and O'Barr 1990), adding a multicultural and multilingual dimension. Highlighting the experiences of non-English-speaking claimants, the study describes the types of cases that are brought by them, the processes that they have to go through in order to have their case heard, and, drawing on analyses of transcription excerpts, investigates litigants' expectations from the court, their sense of entitlement, and their understanding of the law. This is contrasted with a discussion of the attitudes of arbitrators towards translation and language choice and towards the proceedings in general, relating the

findings to research on intercultural communication (e.g. Gumperz 2001) as well as on the ideologies of legal professionals (e.g. Conley and O'Barr 1990, Philips 1998, Haviland 2003).

The findings of this study have implications for legal policy regarding interpreter use, and they also contribute to sociolinguistic theories of language choice, addressing the social significance of speaking English or not speaking it, exploring the role of interpreters as cultural intermediaries, and examining the role of the legal system in maintaining social inequalities in multilingual settings.

This material is based upon work supported by the National Science Foundation under Grant No. BCS-0317838.

References

Abel, R. (ed.) (1982) *The Politics of Informal Justice*. New York: Academic Press.

Angermeyer, P. (2005a) Who is 'you'? Polite forms of address and ambiguous participant roles in court interpreting. *Target*: *International Journal of Translation Studies* 17: 203–226.

Angermeyer, P. (2005b) Who is 'I'? Pronoun choice and bilingual identity in court interpreting. In *Selected Proceedings from NWAV 33*. University of Pennsylvania Working Papers in Linguistics Volume 11(2): 31–44.

Berk-Seligson, S. (1990) *The Bilingual Courtroom*: *court interpreters in the judicial process*. Chicago: University of Chicago Press.

Bourdieu, P. (1991) *Language and Symbolic Power*. Cambridge, MA: Harvard University Press.

Conley, J. (1982) The law. In W. O'Barr (ed.) *Linguistic Evidence*: *language*, *power*, *and strategy in the courtroom* 41–49. New York: Academic Press.

Conley, J. and O'Barr, W. (1990) *Rules versus Relationships*: *the ethnography of legal discourse*. Chicago: University of Chicago Press.

Coulmas, F. (1987) Why speak English? In K. Knapp, W. Enninger and A. Knapp-Potthoff (eds) *Analyzing Intercultural Communication* 93–107. Berlin/New York/Amsterdam: Mouton de Gruyter.

Giles, H., Coupland, N. and Coupland, J. (1991) Accommodation theory: communication, context, and consequence. In H. Giles, J. Coupland and N. Coupland (eds) *Contexts of Accommodation*: *developments in applied sociolinguistics* 1–68. Cambridge/Paris: Cambridge University Press/Maison des Sciences de l'Homme.

Goffman, E. (1981) *Forms of Talk*. Philadelphia: University of Pennsylvania Press.

Gumperz, J. (2001) Contextualization and ideology in intercultural communication. In A. Di Luzio, S. Günthner and F. Orletti (eds) *Culture in Communication*: *analyses of intercultural situations* 35–53. Amsterdam/Philadelphia: John Benjamins.

Hale, S. (2004) *The Discourse of Court Interpreting*: *discourse practices of the law*, *the witness and the interpreter*. Amsterdam/Philadelphia: John Benjamins.

Haviland, J. (2003) Ideologies of language: reflections on language and U.S. law. *American Anthropologist* 105: 764–774.

Merry, S. (1990) *Getting Justice and Getting Even*: *legal consciousness among working-class Americans*. Chicago: University of Chicago Press.

Philips, S. (1998) *Ideology in the Language of Judges*: *how judges practice law*, *politics*, *and courtroom control*. Oxford: Oxford University Press.

Wadensjö, C. (1998) *Interpreting as Interaction*. London/New York: Longman.

Examining jurors' discursive exchanges related to mitigating factors during capital jury deliberations

Desiree Cassar
Visiting Assistant Professor

Department of Psychology, *Claremont McKenna College*,
850 Columbia Avenue, *Claremont*, *CA 91711, USA*
desireecassar@hotmail.com

Awarding Institution:	The Graduate School and University Center of The City University of New York, USA
Date of Award:	May 2005

KEYWORDS: MITIGATING FACTORS, CAPITAL JURIES, DELIBERATIONS, CONTENT ANALYSIS

In capital trials, juries are asked to determine whether a defendant is guilty or not guilty. In some states, juries are also asked to determine a defendant's sentence in a separate sentencing phase of the trial. In this phase, juries are typically presented with aggravating (i.e., arguments for death) and mitigating (i.e., arguments for life) factors, and are given judicial sentencing instructions to guide their sentencing decisions. To apply these instructions jurors must understand the vocabulary used in such instructions.

Bowers (1995) and others have argued that the legal terms 'aggravation' and 'mitigation' are unfamiliar and confusing to many jurors (Haney 1995). Moreover, Capital Jury Project (CJP) researchers have found that some jurors mistakenly consider mitigators as aggravators, often described as the conversion of mitigators into aggravators (Bentele & Bowers 2001; Haney 1995). For instance, jurors may interpret a defendant's mental illness (a mitigator) as a factor likely to increase the defendant's future dangerousness (an aggravator) and consequently vote for a death sentence.

Mitigators are typically presented by attorneys as causes of a defendant's criminal behavior in the hope that the jury will show mercy and render a life sentence. Understanding how individuals characterize the causes of others' behaviors is the domain of attribution theory (Jones & Nisbett 1972). Psychologists have demonstrated that individuals tend to attribute dispositional causes rather than situational causes to others' negative behaviors. Mitigators such as the defendant's deep religious beliefs may be characterized as dispositional; mitigators such as the defendant's neglectful family life may be characterized as situational. Attribution theory served as the foundation of the present study.

Previous research on jurors' understanding of mitigating factors has not included jury deliberations (Bentele & Bowers 2001; Luginbuhl & Howe 1995). Although CJP interview questions offer a glimpse into the deliberation process, jurors' responses represent individual jurors' recollections of their deliberations. In this thesis, a content analysis was carried out to capture jurors' discursive exchanges related to mitigating factors during the deliberation process and to challenge the assumption that jurors understand the concept of mitigation. Additionally, the study served to uncover whether fellow jurors during deliberation assist those who misunderstand legal instructions and the concept of mitigation.

The current research is a secondary analysis of Patry (2001), an experiment investigating the effects of sentencing instructions on capital mock juror decision making. A videotrial based on the sentencing phase of *Buchanan v. Angelone* (1998) in which the defendant was convicted of murdering his family was presented to mock jurors prior to deliberating. The videotrial was manipulated for heinousness of the crime (high or low), and the presence or absence of various mitigators such as the defendant's history of emotional abuse, prior criminal record, and pattern or revised Virginia sentencing instructions.

In the present research, 108 videotaped mock jury deliberations ranging from a few minutes to 40 minutes ($M = 26.8$, $SD = 11.3$) were transcribed. Juror statements were subjected to a coding scheme designed to capture the extent to which jurors discussed mitigators, converted mitigators into aggravators, and assisted fellow jurors.

A clustered coding scheme of seven categories (structural, thesis, attribution, juror assistance, legal instructions, sentence, and extra-legal issues) was employed with various codes in each category. The 'thesis' category included codes related to statements about mitigators and aggravators. The 'attribution' category included codes related to situational and dispositional mitigators, and whether the statements favored or disfavored the defendant. If a statement referred to the absence of emotional support for the defendant in response to past emotional abuse, for example, it was coded as a situational mitigator favoring the defendant. The 'juror assistance' category included codes such as jurors' explanations to fellow jurors about mitigation. The frequency of codes was the unit of analysis and the dependent measure for all analyses was the jury

sentence. To control for varying deliberation lengths, codes were reported as a percentage of all coded statements.

When comparing favoring and disfavoring dispositional mitigators across conditions, a higher percentage of statements reflecting disfavoring mitigators (M = 8.67, SD = 5.77) than favoring mitigators (M = 3.83, SD = 4.67), $t(107)$ = -6.35, p = .001, r^2 = .41, was obtained. Less than one percent of the coded statements represented conversion or assistance.

Based on this finding, a logistic regression analysis was conducted to test heinousness, instructions, disfavoring dispositional mitigators, inconsistency of mitigators as predictor variables and sentence as the outcome variable. A test of the full model with the four predictors against a constant-only model was statistically reliable, χ^2 (4, N = 108) = 26.16, $p < .001$, R^2 = .28, indicating that the set of predictor variables distinguished between life and death sentences (52 life and 56 death sentences). Prediction success was notable, with 80.4% of life sentences and 69.2% of death sentences correctly predicted, for an overall success rate of 75%. The Wald criterion indicated that heinousness, z = 7.76, $p <$ .005, and the percentage of statements of dispositional disfavoring mitigators, z = 4.5, $p < .03$, reliably predicted mock jury death sentences. An odds ratio of .28 indicates that heinousness was the most influential predictor of death sentences. An odds ratio of 1.09 reflected little change in the likelihood of voting for death on the basis of a one unit change in percentage of statements of dispositional disfavoring mitigators.

The content analysis of deliberating capital mock juries identified heinousness, an aggravator typically presented in capital cases, as an influential predictor of death sentences. Although conversion of mitigators into aggravators and juror assistance occurred infrequently, the analysis of jurors' discursive exchanges uncovered how mitigators categorized as dispositional may serve to disfavor the defendant, thereby increasing the likelihood of a death sentence. This information is pertinent to defense attorneys in understanding how juries deliberate mitigating factors, and to the legal system in its pursuit of constitutional death penalty statutes.

Cases cited

Buchanan v. Angelone, 118 S. Ct. 757 (1998).

References

Bentele, U. and Bowers, W. J. (2001) How jurors decide on death: guilt is overwhelming; aggravation requires death; and mitigation is no excuse. *Brooklyn Law Review* 66: 1011–1090.

Bowers, W. J. (1995) The capital jury project: rationale, design and preview of early findings. *Indiana Law Journal* 70: 1043–1102.

Haney, C. (1995) Taking capital jurors seriously. *Indiana Law Journal* 70: 1223–1232.

Jones, E. E. and Nisbett, R. E. (1972) The actor and the observer: divergent perceptions of the causes of behavior. In E. E. Jones et al. (eds) *Attribution*: *perceiving the causes of behavior* 79–94. Morristown, NJ: General Learning Press.

Luginbuhl, J. and Howe, J. (1995) Discretion in capital sentencing instructions: guided or misguided. *Indiana Law Journal* 70: 1161–1181.

Patry, M. W. (2001) Exploring jury decision making in death penalty cases. Unpublished doctoral dissertation. Graduate College at the University of Nebraska (Lincoln), Nebraska.

An acoustic study on disguised voices

Cuiling Zhang
Associate Professor

Department of Forensic Science, *China Criminal Police University*,
Tawan Street NO. 83, *Huanggu District*, *Shenyang*, *China*, *110035*
cuilingzhang@yahoo.com.cn

Awarding Institution: Nankai University, China
Date of Award: June 2005

KEYWORDS: FORENSIC PHONETICS, FORENSIC SPEAKER IDENTIFICATION, DISGUISED VOICE, ACOUSTIC ANALYSIS, AUTOMATIC SPEAKER IDENTIFICATION

It is a worldwide growing tendency that perpetrators are inclined to disguise their voices in order to conceal their identities, especially in cases of threatening calls, extortion, kidnapping and even emergency calls to the police. Various disguise techniques create obstacles for forensic speaker identification. The voice can be distorted by both extrinsic and intrinsic factors ranging from noisy environments, transmission channels, inferior recording equipment, illness, highly fluctuating emotions, and imitation. Voice disguise refers to the falsification of the speaker's natural voice for the purpose of concealing their identity and can be classified as deliberate or non-deliberate and electronic or non-electronic. Some disguises can fool the listener's ear successfully and produce great variation of parameters such as pinched nostrils, whisper and falsetto, which makes spectrogram comparison difficult or even impossible.

This raises questions about whether disguised voices can be correctly attributed to the speaker. For example, can speaker identification be used for

disguised voices? To what extent can people change their natural vocal habit to adopt different types of disguise? How much success can be achieved by speakers disguising their voice to conceal their identities? Do the speakers adopt the same or different strategies to change their voice for disguise and what are those strategies? How great is inter-speaker variation when they disguise their voices? Which parameters are stable and robust for resisting disguise and which are vulnerable? How does voice disguise affect the performance of speaker identification? Could we find a way to make parameter compensation of normal voices from disguised voices? This research was developed in an effort to identify and answer these questions.

This thesis investigates nine voice disguises which are commonly used by offenders in real casework by auditory and acoustic analyses. The voice disguises used in this thesis include raised and lowered pitch, fast and slow speech, whisper, pinched nostrils, masking on mouth, and the use of bite blocks and objects in the mouth, etc. The features of disguised voices, their acoustic representation, the parameter changes, their effect on speaker identification and some regression models of parameters are summarized based on comparisons with normal (non-disguised) voices.

Eleven male students from China Criminal Police University participated in the study. All speakers were asked to read ten sentences which have been found to commonly occur in kidnapping cases using their normal voices and subsequently nine disguised voices in a quiet room. All speech was analyzed by aural perception by *Praat* computer software acoustically. The parameters such as mean pitch, duration, intensity, formant frequency and long term average spectrum were measured and statistically analyzed.

Firstly, acoustic investigation and comparisons were made for one particular speaker's normal voice with his nine disguised voices because he seemed more skilful in disguising than other speakers. The aim was to see how easily the voice can be changed and to what extent the vocal organ can change its model to adapt different disguise patterns. It was found that voice disguise produces greater speech variation than normal voices, which can be detected through both auditory perception and speech spectra. Generally, the auditory perception is more sensitive for detecting variation than spectrum analysis. However, the effect of voice disguise is limited. On the one hand physiological anatomy and the modulating ability of speech organs confine the magnitude of voice change; on the other hand voice disguise lacks stability and continuity so that the rootedness of speech habit makes it difficult to keep one's voice disguised for a long time. Of course, this is related to the skill of the individual applying the disguise.

Secondly, the nine disguised voices from all eleven speakers were separately analyzed aurally and acoustically. The acoustic investigation, statistical analysis

and comparison with the normal voice were made. Features of disguised voices, their variations compared to normal voice, the stability and robustness of parameters, the effect of parameter change on the performance of speaker identification and some regression models of parameters for deducing the normal level from disguised voice such as pitch, intensity and duration are summarized. The results show that auditory perception is not always consistent with acoustic performance. Some perceptions have no corresponding acoustic correlates at all. The changes of parameters are not proportional. The tendency and extent of parameter change differ with different disguise patterns. Some parameters are highly sensitive to voice disguise such as pitch and intensity while others such as formants are not. Regression models can help to deduce the normal level of parameters from disguised voices with reasonable success. For each voice disguise pattern, speakers show different tendencies and extent of parameter change which indicates individuality and also greater inter-speaker variation. This is beneficial for speaker identification because voices can be more easily distinguished as inter-speaker variation increases.

Thirdly, an effect investigation of the nine disguised voices on an Automatic Speaker Recognition (ASR) System was made. Each disguised voice was used to make comparisons with all non-disguised voices in the database to make automatic speaker identification and verification. The performance of the system was rated and the effect of disguised voices on Automatic Speaker Recognition was evaluated. The investigation revealed that the system's performance in speaker identification is reduced significantly for disguised voices. The effect of disguising on the rate of automatic speaker recognition varies with different disguise techniques and also varies among speakers. In the nine disguising patterns, masking on mouth and whisper has the greatest effect on ASR, where correct recognition rates (CRR) were zero, chewing gum, lowered pitch and pinched nostril were next with CRRs of 45%, 55% and 65% respectively. For the other three disguising patterns, the effect did exist but was weak with recognition rates all above 85%. Moreover, it seems that some speakers are more easily recognized than others even for the same disguising pattern. This may result from their idiosyncratic voice quality and disguising skill.

In summary, although the voice can be changed and distorted significantly through disguise, it is still possible to identify the speaker. However, it should be noted that the overall performance of speaker identification is dependent on the speaker's disguising skill and the extent of voice distortion. This thesis is significant for forensic speaker identification because it offers some strategies to recognize and identify disguised voices. The thesis also provides data and support for automatic speech recognition and speaker identification.