

Forensic Phonetics and Sociolinguistics

P. Foulkes and J. P. French

Forensic analysis of speech and language is a relatively young discipline. A specialist academic and professional journal for the field—*Forensic Linguistics: the International Journal of Speech, Language and the Law*—was only established as recently as 1994. Appropriate strategies for applying phonetic and sociolinguistic knowledge and methodologies to legal problems are in the process of being refined, expanded, and their reliability tested.

Applications in the legal context of techniques developed in phonetics and sociolinguistics fall into four main categories: (a) deciphering the content of ‘difficult’ recordings; (b) speaker profiling; (c) speaker identification; and (d) constructing voice ‘line-ups’ in order to evaluate ear-witness testimony. These activities are discussed in turn below. More extensive introductions can be found in French (1994), Hollien (1990), Künzel (1987) and Nolan (1991, 1997).

1. Deciphering the Content of Difficult Recordings

Tape recordings, which may contain information relevant to court hearings, can be problematic or difficult to decipher for a variety of reasons. For instance, they may be contaminated by extraneous noise, which obscures the conversation of interest, or the speakers may have pathological or otherwise non-standard—perhaps foreign—patterns of pronunciation.

In relation to noisy recordings, the sound filtering and signal processing techniques developed by audio engineers can be of help in improving speech to noise ratios and thereby rendering the speech signal more audible/intelligible. However, phonetic and sociolinguistic analysis may also be crucial in arriving at a final determination of what was said.

Hirson and Howard (1994), for example, analyzed the content of the ‘black box’ flight recorder recovered from the wreck of a South African Airways aircraft which had been lost in mid-flight in 1987, with the loss of 159 lives. The black box records the last 30 minutes of activity in the aircraft’s cockpit, via a microphone housed in the ceiling. It was hoped that the spoken material would offer vital clues as to the cause of the accident. However, the recording proved to be problematic in various respects: the tape was highly degraded after having spent over a year under water; it contained only 81 seconds of speech after the first indication that there was a serious problem (a fire alarm); and the aircraft crew were multilingual, speaking English and Afrikaans. The poor signal-to-noise ratio of the recording meant that even deciphering which language was being spoken was far from straightforward. Ultimately, phonetic (audi-

tory and acoustic) analysis enabled the majority of the spoken material to be transcribed, and revealed no evidence of a bomb, or of human error on the part of any of the crew. This case also resulted in recommendations to the Federal Aviation Administration for improving the structure and positioning of flight recorders.

In certain instances the problem of interpreting the content of a recording may be extremely ‘localized,’ with only a closely circumscribed section of speech being at issue. Such cases have been referred to as ones involving ‘disputed utterances’ (French 1990). The latter study reports a case concerning a Greek-born doctor working in the UK, who had been accused of illegally prescribing drugs for profit. An investigative journalist, posing as an addict, managed to obtain a prescription for a controlled drug and secretly recorded the encounter. The dispute, at a subsequent Medical Council tribunal, concerned whether the doctor had told the journalist ‘you *can* inject those things,’ or whether he had warned ‘you *can’t* inject those things.’ The disputed word was pronounced [kan]. While such a pronunciation would clearly indicate *can* in Standard British English, it is ambiguous in other varieties, including the Greek-accented one spoken by the doctor. In order to assess which word was intended, acoustic analysis was undertaken to compare the disputed word with undisputed examples of *can* and *can’t* from the doctor’s speech. Although they were auditorily very similar, analysis of the vowel formant frequencies revealed that the disputed word more closely matched the doctor’s formant patterns for *can’t*.

2. Speaker Profiling

A combination of phonetic and sociolinguistic analysis of a voice can aid in establishing information about the speaker’s background. Speaker profiling is regularly requested in the early stages of kidnappings, where a recording of the kidnapper’s voice may be the only clue to his or her identity. Information can be gleaned about the speaker’s sex, age, regional and social background, and idiosyncrasies such as speech disorders. The strength of conclusions varies and is highly dependent on the length and quality of material, the degree of any disguise involved, and the extent of descriptive dialectological and sociolinguistic information available. Establishing a speaker profile helps in defining a ‘target’ population and thereby narrowing the search for the culprit.

One of the best known examples of speaker profiling was undertaken in the case of the serial

murderer known as the ‘Yorkshire Ripper’ (Ellis 1994). During the investigation the police received tapes from a man claiming to be the killer. Profiling by Ellis established that the man most likely came from, or had lived for a substantial period of time in, the Sunderland area of northeastern England. However, time proved that the real killer in fact came from Yorkshire, a county to the South of the area identified. The sender of the tapes, who was evidently a hoaxer, remains at large.

Detailed phonetic analysis may also serve to indicate the likelihood that a speaker is intoxicated, although the precise effects of alcohol on speech are not clearly understood or consistent across individuals (Künzel et al. 1992). The recorded speech of the captain of the Exxon Valdez oil tanker, which ran aground in Alaska in 1989, was investigated to assess whether he was intoxicated at the time of the accident (Hollien 1993).

More recently, sociophonetic analysis has been used as evidence to pinpoint the origins of asylum seekers in some European countries, in order to verify the legitimacy of their asylum claims (Simo Bobda et al. 1999).

3. Speaker Identification (SID)

Perhaps somewhere between 70 percent and 80 percent of all forensic cases involve SID: that is, identifying, by means of comparative phonetic testing, a person heard speaking in a criminal recording. A ‘criminal recording’ in this context may be anything from a hoax bomb warning recorded over an emergency telephone line, to a death threat left on a tape posted through the victim’s letterbox, or a surveillance recording made secretly by the police within the home of someone planning a robbery.

The commonest scenario is for a forensic phonetician to compare the questioned voice in the criminal recording with a speech sample from the suspect, in order to assess the likelihood of their being the same person. Labov (1988), for example, reports the case of a New Yorker accused of making threatening calls to Pan Am airlines. Detailed phonetic transcripts and instrumental evidence comparing the caller with the accused suggested beyond reasonable doubt that the caller was in fact from Boston, resulting in the acquittal of the accused. Many other example cases are detailed by Baldwin and French (1990), and in the journal *Forensic Linguistics*.

It is now widely accepted that there is no indelible feature of an individual’s voice. That is, there is no vocal equivalent of a fingerprint or a DNA profile, which can offer irrefutable proof of speaker identity. Research in sociolinguistics and phonetics has shown that an individual’s speech may vary as a result of many factors, including the dimensions of the vocal tract, and features acquired as a result of the

speaker’s social and regional background. Speech features may also vary markedly from situation to situation: for example, as a result of stress, conscious disguise, or the effect of speaking on a telephone. Different voices may also be affected in different ways. For instance, most people, but not all, speak more loudly when using a telephone, which results in a rise in average fundamental frequency (Hirson et al. 1994).

In casework, various segmental and suprasegmental features may be analyzed, as well as lexis and grammar. Nowadays SID usually includes both auditory and acoustic phonetic analysis, although debate on the relative merits of the two approaches can be found in the earlier literature (Kersta 1962, Baldwin and French 1990, Nolan 1990, French 1994). Much research focuses on determining which phonetic parameters of a voice are the most useful in identifying an individual speaker, and also the degree of variability individuals display along various phonetic and sociolinguistic dimensions. In all cases the strength of conclusion that can be drawn is dependent on reference information, including—where available—population statistics, from fields such as phonetics, phonology, dialectology, sociolinguistics, and speech pathology. For the majority of forensic phoneticians, and all IAFP members, the outcome of an SID analysis is an expression of opinion ranged on a scale of confidence, rather than a categorical statement that the suspect is or is not the culprit. And although either the prosecution or the defence initially engages an analyst, the IAFP promotes the view that a neutral comparison of samples should always be undertaken. Analysts therefore do not set out to try and win the case for the side which engages them.

4. Speaker Identification by Lay Witnesses

While the bulk of forensic cases involve speaker identification by linguists using tape-recorded materials, in certain cases there is no permanent record of the voice involved in the crime. Instead, the perpetrator may simply have been heard by a ‘lay’ (i.e., phonetically-naïve) witness. Examples include the receipt of obscene phone calls, or witnessing a crime committed in the dark or by a masked robber.

It is now generally acknowledged that a witness must be able to demonstrate recall of a voice in a formal test, although this has not always been the case. Milroy (1984) reports a blackmail case from Northern Ireland in which a family received telephone threats following an incident where a masked intruder had entered their home. A suspect was arrested and subsequently given a ten-year prison sentence after his voice was overheard in a police station by the family members. No formal testing established their ability to identify the suspect as the

man who had made the phone calls. Milroy's analysis of the sociolinguistic features of the suspect and the phone caller later revealed significant accent differences between them.

The most likely means of testing a witness's recall of a voice is by identifying the perpetrator of the crime in a *voice parade* or *line-up*. This consists of a set of recorded voices, including that of the suspect. In such cases forensic phoneticians may be called upon either to construct the parade itself, or to analyze the contents of the parade to ensure that no samples stand out unfairly in relation to the others, for example through a markedly different accent, voice quality, or recording fidelity. A detailed description of the construction of a voice parade for use in a rape case is provided by Nolan and Grabe (1996), while Butcher (1996) exemplifies the assessment of samples in a voice parade used to identify an armed robber.

There is at present, however, no incontrovertible method of constructing or administering a voice parade (see Broeders 1996, Hollien 1996). However, the most carefully argued and rigorously constructed set of guidelines may be found in the seminal article by Broeders and Rietveld (1995).

Testing recall of voices is complicated by research showing that witnesses' performance in recall tasks varies markedly as a result of numerous factors. These include hearing ability, the degree of familiarity with the voice, and whether exposure to the voice was active or passive (Künzel 1994).

It is widely assumed that recall of familiar voices is straightforward or even automatic. However, while it is generally true that recall of familiar voices is better than that of unfamiliar ones, even close friends and family members can show a significant degree of inaccuracy or inability to recognize voices that are well known to them (Foulkes and Barron 2000). It has also been demonstrated that memory of a voice tends to decay over time, so that it is imperative to begin constructing a line-up test as soon as possible after the witnessing of the event (Hollien 1996). In practice, though, there may be long delays as a result of failure to apprehend a suspect quickly.

5. Conclusion

There remain many controversial issues in the practice of forensic phonetics. As far as speaker identification is concerned, it is unlikely that any speaker-specific invariant parameter will ever be discovered, such that statements of speaker identity might come to emulate the strength of those derived from DNA testing or fingerprinting. That said, it is encouraging that research in relevant areas is undergoing continued growth.

See also: Discourse Analysis and the Law.

Bibliography

- Baldwin J, French J P 1990 *Forensic Phonetics*. Pinter, London
- Broeders A P A 1996 Earwitness identification: common ground, disputed territory, and uncharted areas. *Forensic Linguistics* 3: 1–13
- Broeders A P A, Rietveld A C M 1995 Speaker identification by earwitnesses. In: Braun A, Köster J-P (eds.) *Studies in Forensic Phonetics*. Wissenschaftlicher Verlag, Trier
- Butcher A 1996 Getting the voice lineup right: analysis of a multiple auditory confrontation. *Proceedings of the 6th Australian International Conference on Speech Science and Technology*. Flinders University, Adelaide
- Ellis S 1994 The Yorkshire Ripper enquiry: part 1. *Forensic Linguistics* 1: 197–206
- Foulkes P, Barron A 2000 Telephone speaker recognition amongst members of close social network. *Forensic Linguistics* 7: 180–98
- French J P 1990 Analytic procedures for the determination of disputed utterances. In: Kniffka H (ed.) *Texte zu Theorie und Praxis forensischer Linguistik*. Max Niemeyer, Tübingen
- French J P 1994 An overview of forensic phonetics with particular reference to speaker identification. *Forensic Linguistics* 1: 169–81
- Hirson A, French J P, Howard D M 1994 Speech fundamental frequency over the telephone and face-to-face: some implications for forensic phonetics. In: Windsor Lewis J (ed.), *Studies in General and English Phonetics*. Routledge, London
- Hirson A, Howard D M 1994 Spectrographic analysis of a cockpit voice recorder tape. *Forensic Linguistics* 1: 59–69
- Hollien H 1990 *The Acoustics of Crime*. Plenum, New York
- Hollien H 1993 An oilspill, alcohol and the Captain: a possible misapplication of forensic science. *Forensic Sciences International* 60: 97–105
- Hollien H 1996 Consideration of guidelines for earwitness lineups. *Forensic Linguistics* 3: 14–23
- Kersta L G 1962 Voiceprint identification. *Nature* 196: 1253–7
- Künzel H 1987 *Sprechererkennung: Grundzüge Forensischer Sprachverarbeitung*. Kriminalistik-Verlag, Heidelberg
- Künzel H J 1994 On the problem of speaker identification by victims and witnesses. *Forensic Linguistics* 1: 45–58
- Künzel H J, Braun A, Eysholdt U 1992 *Einfluß von Alkohol auf Sprache und Stimme*. Kriminalistik-Verlag, Heidelberg
- Labov W 1988 The judicial testing of linguistic theory. In: Tannen D (ed.) *Linguistics in Context*. Ablex, Norwood, NJ
- Milroy J 1984 Sociolinguistic methodology and the identification of speakers' voices in legal proceedings. In: Trudgill P (ed.) *Applied Sociolinguistics*. Academic Press, London
- Nolan F J 1990 The limitations of auditory-phonetic speaker identification. In: Kniffka H (ed.) *Texte zu Theorie und Praxis Forensischer Linguistik*. Max Niemeyer, Tübingen
- Nolan F J 1991 Forensic phonetics. *Journal of Linguistics* 27: 483–93

Language Variation & Change: Dialects and Social Groups

- Nolan F J 1997 Speaker recognition and forensic phonetics. In Hardcastle W J, Laver J (eds.) *Handbook of Phonetic Sciences*. Blackwell, Oxford
- Nolan F J, Grabe E 1996 Preparing a voice line-up. *Forensic Linguistics* 3: 74-94
- Simo Bobda A, Wolf, H G, Peter L 1999 Identifying regional and national origin of English-speaking Africans seeking asylum in Germany. *Forensic Linguistics* 6: 300-19