

Speaker Verification

Speaker verification is a speaker recognition technology which enables speakers to use their voice to access restricted services such as banking. There are two types of speaker recognition:

- Speaker verification
- Speaker identification

Speaker verification involves accepting or rejecting the claim the speaker is who they say they are (a one-to-one comparison). While speaker identification attempts to recognise a speaker's voice from a set of known speakers (a one-to-many comparison). Both technologies require that speakers 'enrol' their voice so that the system can learn their particular speech characteristics. The enrolled speech forms a 'voiceprint' also called a template.

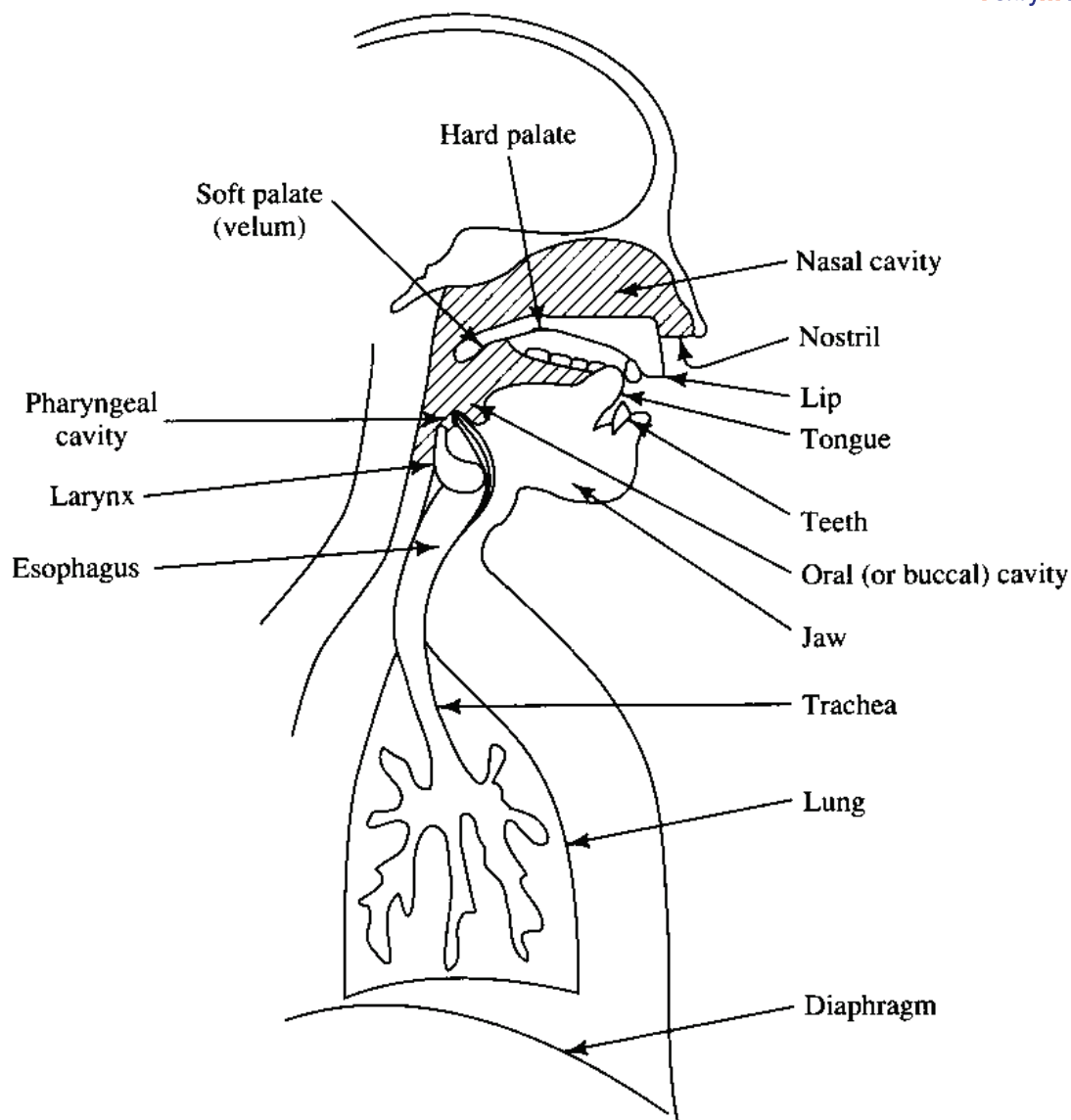
Speaker verification, rather than speaker identification, is more commonly used to authenticate a caller. An organisation may use speaker verification on its own to authenticate a caller or in combination with other authentication items, such as an account number or security questions. (see solutions page and box office)

Speaker identification may be used in law enforcement areas in a similar way as fingerprint recognition. It is also used to determine if a person has previously enrolled or not.

How does speaker recognition work?

Differences in the way sounds are produced by speakers enable the recognition system to differentiate speakers.

Speech is produced mainly by the vocal tract (see shaded portion of the diagram below). The vocal tract is made up of areas or cavities (nasal, oral) and articulators (tongue, teeth, lips) which produce the various sounds that languages distinguish. For example, to make the sound 'm', the lips are closed, the velum opens to allow air flow through the nasal cavity and the tongue is in a neutral position. Compare that to the sound 's', where the tongue is raised behind the palate and top teeth, air is forced between the small opening between the teeth and the tongue creating friction and airflow is through the oral cavity.

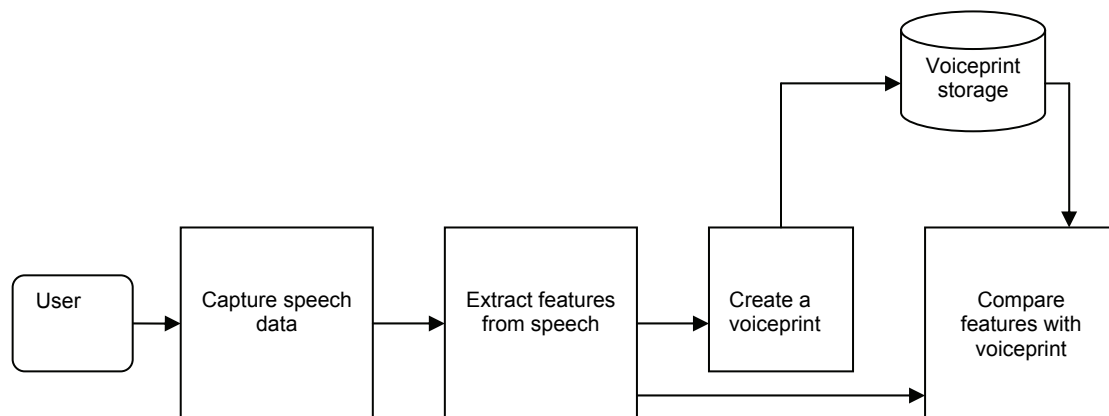


Now consider that the shape and size of the articulators, the way they are positioned and the shape of the areas that air is pushed into, are all slightly different across people. This is a physiological difference that contributes to differentiating the speech of different speakers. It is the physiology of the vocal tract that modifies the acoustic features of sound. In addition, there are behavioural aspects of people that contribute to differentiation of the acoustic sound. These aspects include voice pitch and speaking style. The physiological and behavioural differences are evident in the acoustic features that are extracted from the speech.

An interesting feature of speech is that the same utterance said by the same person at different times has similar but slightly different features. These variations can affect recognition. So for recognition to be successful a template must be created to capture these variations. This template, also known as a voiceprint, is created by the speaker 'enrolling' or saying various words, digits, phrases so that the system can recognise the speaker's speech patterns.

When a speaker attempts verification, the speech input is compared with a previously created voiceprint. The speaker is then either accepted or rejected by the system.

The basic components required in a speaker verification system are shown below.



A voiceprint consists of acoustic features extracted from the speech input and is not a recording of the speech input. A voiceprint is typically around 20kb in size.

How does it compare to speech recognition?

The main difference between speaker recognition and speech recognition is that speaker recognition attempts to recognise the speech pattern of a particular speaker. It requires a speaker to enrol a voiceprint for verification to occur. Whereas speech recognition aims to recognise the speech of all speakers that access an application. A speech recognition model then is based on speech data collected from hundreds to thousands of speakers.

Both are similar in that they use pattern matching techniques.

How does verification relate to biometrics?

Biometrics is a general term referring to a wide range of measures of biological or behavioural data.

A voiceprint created for speaker verification is a biometric. A biometric uses the physical and/or behavioural characteristics that are unique to an individual to either establish or confirm the identity of that individual.

Other biometrics are fingerprint, iris and face recognition. A unique feature of voiceprints is that they combine both physical and behavioural features into a single voiceprint which is unlike the other biometrics.

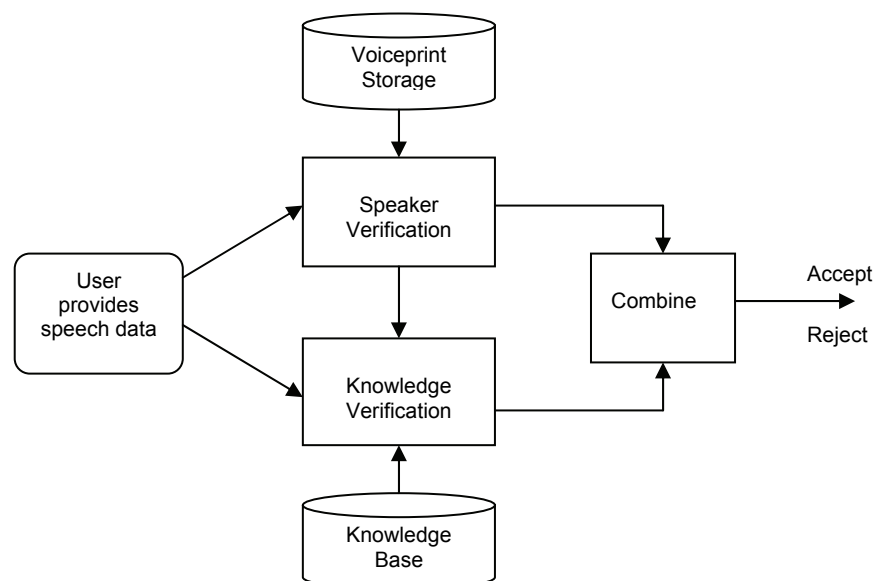
Why is speaker verification used?

Speaker verification is used to authenticate an individual and is described as one of three elements in authentication methods. These elements are:

- something you know (password, secret question/answers)
- something you have (passport, drivers licence)
- something you are (biometric)

Organisations, private and government, have authentication requirements which vary. Some require one element, others two to three depending on what the security requirements are for a particular service. For example, to change an address of a bank account, three security questions may be asked (eg date of birth, mailing address, mother's maiden name), while to get an account balance, two security questions may be asked.

The diagram below shows a caller verifying a voiceprint and knowledge ('something you know') in which the scores from each are combined and a decision is made to accept or reject the caller.



The advantage of a voiceprint is that it combines both the 'something you know' and the 'something you are'. This means that less authentication questions are required. In addition, it is more permanent. Often the 'something you know' element which includes pins, passwords, and secret questions/answers are shared, lost, forgotten, and so have to be replaced. Some organisations require pin resets every week which means that these aspects of this element involve time, money and inconvenience. In

addition, security can be compromised by people writing pins down to help with remembering them.

Security questions are asked by employees that work in the organisations you transact with. As you are likely to transact with a number of different organisations, you are sharing your personal information with a number of people. Typically these people work in call centres which may or may not be located in the country where the organisation is based. The risk of identity fraud increases the more times you pass on personal details to others.

With voiceprints, authentication is possible without the need to interact with people. Thus the appeal of voiceprints is that they enable more secure transactions over the telephone from the point of view of both the customer and the organisation.

What are the security issues with voiceprints?

Previous privacy commissioner, Malcolm Crompton (2003), quotes Richard E Norton of the International Biometric Industry Association (IBIA) who says:

'Simply put, it's getting harder and harder to preserve personal privacy without using biometrics. It's a misperception that biometrics somehow compromise privacy; in actuality, they are the best way to lock up a record and ensure that an identity cannot be stolen. Biometrics are designed to give the user total control over who has access to his or her information, and provide a clear audit trail if someone tries to obtain data from a record. Which would consumers rather have - a system like we have now, with your name, social security number, birth date, address and phone number available to anyone who has PIN, password, or "hacked" access to customer records, or a system that prevents a record from being penetrated unless it's unlocked through biometric verification? Privacy advocates are on thin ice here, especially when they claim that a record can be compromised or stolen. A biometric cannot be reverse-engineered to find out who you are, and it cannot be used to link records together - in fact, the technology by definition prevents it. Finally, you can't be an impostor by using someone's biometric; the template is dynamic, and the data is encrypted. Biometrics raise the bar against fraud and abuse at no cost to privacy.'

There is general agreement that biometrics can benefit individuals and society, and also that they could have privacy enhancing capabilities. However, it is recognised that attention to privacy is necessary giving consideration to what advantages and risks would be associated with the use of biometrics, and building in privacy protection and privacy enhancement into biometric software and hardware.

Malcolm Crompton (2003) *Biometrics and Privacy: The End Of The World as We Know It Or The White Knight Of Privacy?* Paper presented at the 1st Biometrics Institute Conference, 20 March 2003.

Links

Biometrics Institute (Australia) www.biometricsinstitute.org

Biometric Consortium www.biometrics.org

Glossary

Authentication	The process whereby an individual's identity is confirmed, either by verification or by identification.
Biometric	Where biological and/or behavioural characteristics which distinguish one person from another are used to recognise the identity, or to verify the claimed identity, of a person who has enrolled.
Capture	Obtaining a biometric sample
Enrolment	The process whereby biometric samples are collected from a person, from which a template is created and then stored with references associating it to that person.
Match	A comparison of a match template and a previously stored reference template and a score showing the degree of similarity or correlation between the two.
Match template	Data extracted from a claimant's biometric sample containing a biometric measurement of a claimant which is used for comparison against a previously stored reference template.
Privacy	The right of an individual, group, or institution, to control, edit, manage, and delete information about themselves and decide when, how, and to what extent that information is communicated to others.
Template	A representation of the biometric measurement of a person which is used by a biometric system for biometric matches.
Verification	The process whereby a comparison is made between a stored reference template and a match template based on a claimed identity.
Voice biometric	A biometric technology which is based on the unique characteristics of the acoustic information found in the voice of a speaker.
Voiceprint	A representation of a voice biometric measurement of a person which is used by a speaker verification system for matching.