

# customer care solutions from Nuance



## Nuance VocalPassword™ :: Liveness Detection using Voice Biometrics (Intra - session voice variation) - Overview



VocalPassword™

### Glossary

**Text-Dependent** – where a speaker enrolls a specific pass phrase or a set of pass phrases, and is prompted to repeat one or more pass phrases during verification.

**Text-Independent** – where enrolment and verification is based on unconstrained natural speech.

**Text-Prompted** – where enrolment is based on audio samples that contain isolated atoms of speech (for example, digits or words) or combinations of them, and verification is based on prompting for random combinations of these atoms.

### The Challenge

One of the main risks that text-dependent and text-prompted voice biometrics systems face is the risk of fraudsters using voice recordings of legitimate speakers. These can be acquired by interception or “Vishing” (Voice Phishing).

### The Solution

The method suggested in this document enables mitigating this risk by performing intra-session voice variation liveness testing, as described below.

1. The speaker enrolls a specific pass phrase as in standard text-dependent voice biometrics.
2. During verification, the speaker is prompted to repeat one pass phrase.
3. The verification utterance is evaluated against a voiceprint trained from the enrollment audio, and a similarity score or decision (accept/reject) is generated by the system.
4. Following the initial verification phase, the speaker is prompted to repeat a random sentence or sentences. Due to accuracy considerations, these sentences should be selected at random from a closed set of sentences.
5. Using text-independent voice biometrics technology, the audio collected in the first verification phase (step two above) and in the second verification phase (step four above) are compared and a similarity score or decision is produced. This comparison is performed by training a text-independent voiceprint using the audio collected in step two and evaluating whether the voice recorded in step four comes from the same speaker. For enhanced accuracy, this operation is performed both ways, by training a text-independent voiceprint using the audio collected in step four and evaluating whether the voice recorded in step two comes from the same speaker. The similarity scores of both comparisons is then combined for increased confidence.
6. Using Speech Recognition Technology or the available utterance validation functionality in VocalPassword™ version 7, the audio recorded in step four (which is a random sentence or set of sentences) goes through an utterance validation process that checks whether the speaker indeed repeated the requested sentence or sentences. A score/decision regarding this match is produced.
7. A liveness detection score is produced by combining, factoring-in, or performing arithmetical calculation on the outcomes (scores/decisions) of steps three, five and six above.

**The Implications**

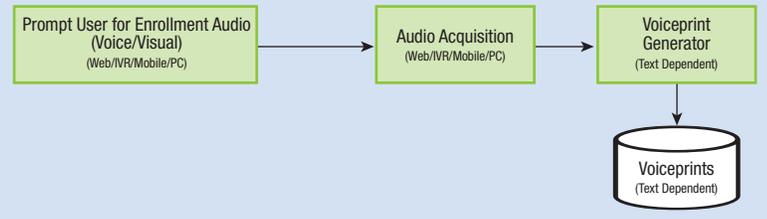
1. In order to achieve the optimal accuracy over the text-independent engine, a background model should be created for every sentence. This process also “trains” the sentence model used for utterance validation. The operational call-flow can be designed to enable collecting audio and enabling calibrating new sentences as the system is used.
2. The accuracy of the liveness detection mechanism is greatly affected by the amount of audio collected. While a rough indication can be obtained using a single sentence, it is recommended to allow for additional prompts to enhance accuracy when necessary. As the call duration will be longer, it is recommended to consider using this method for high-risk transactions only. The liveness detection score is provided in addition to the text dependent verification score and can be used as an additional factor in the overall decision.
3. The VocalPassword system includes an easy-to-use Web based application for defining the sentences used (Dictionary) and to map each sentence to the relevant background model.

The diagram outlines VocalPassword’s liveness detection mechanism implemented in VocalPassword V7. Specific APIs were added to this version enables easy and straight forward implementation of this functionality.

**Intra-Session Voice Variation – A Method for Liveness Detection**

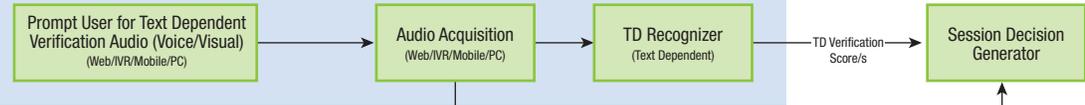
**Enrollment**

**Text Dependent Enrollment**



**Verification**

**Text Dependent (TD) Verification**



**Liveness Detection – Text Independent (TI)**

