DSP Innovations

TWELP 600 bps Vocoder

(version 5.0) achieves outstanding speech quality and intelligibility, unmatched by any other vocoder at this bitrate.

For Digital HF Radio and other markets.

TWELP Technology Features. The vocoder is based on newest technology of speech coding called "Tri-Wave Excited Linear Prediction" (TWELP) that was developed by experts of DSPINI.

TWELP technology is a new class of vocoders that differs from any other LPC-based vocoders by:

- advance reliable method of pitch estimation
- pitch-synchronous analysis
- advance tri-wave model of excitation
- newest quantization schemes
- pitch-synchronous synthesis

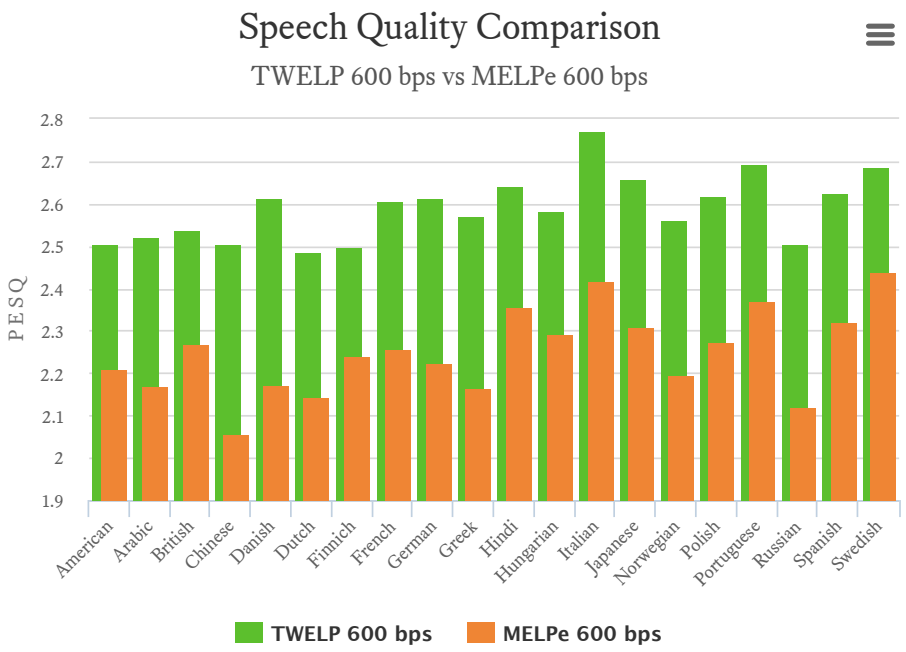
Thanks to these unique features, TWELP technology provides significantly better speech quality than other well-known technologies—including AMBE+2, MELPe, ACELP, and others—at equivalent bit rates ranging from 300 bps to 4800 bps and beyond.

Additionally, unlike other low-bitrate vocoders (such as MELPe, for example), TWELP delivers much higher quality for non-speech signals, including sirens, background music, and similar audio.

Speech Quality. This is a comparison with the MELPe vocoder, which operates at 600 bps.
The TWELP 600 bps and MELPe 600 bps vocoders were tested using the ITU-T P.50 speech base in 20 different languages.

Note:
We have updated the speech database by minimizing inter-speech pauses to eliminate their impact on the evaluation results. Therefore, the numbers obtained from the quality measurements using this updated speech database differ from those previously obtained with the original speech database, where speech pauses were not removed.

The ITU-T P.862 tool was used to evaluate speech quality in terms of PESQ scores:

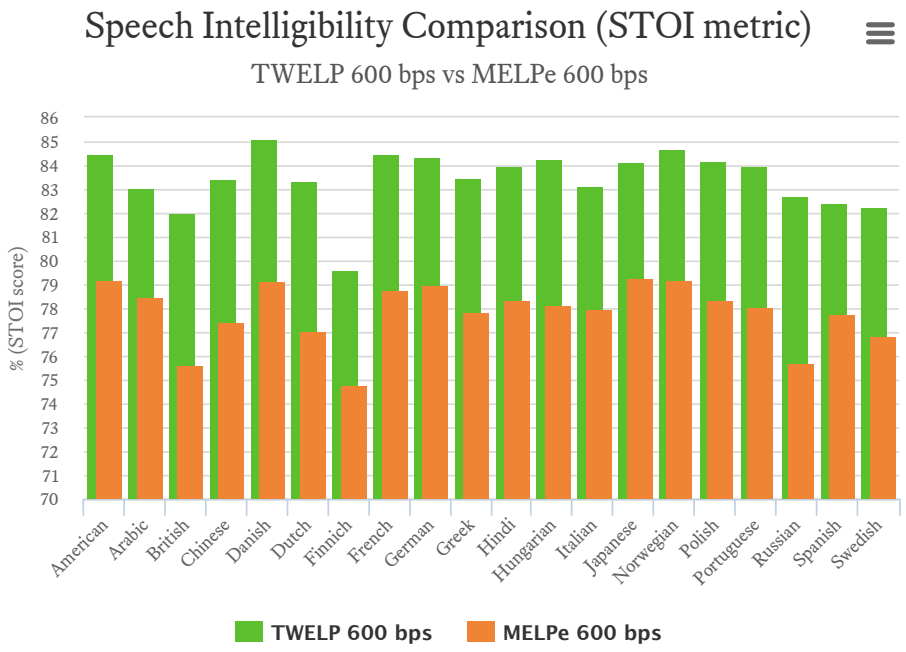


The diagram demonstrates a huge difference in speech quality between the TWELP 600 and the MELPe 600 vocoders. Exact numbers are shown in the table below.

Language	TWELP 600 bps	MELPe 600 bps
American	2.506	2.211
Arabic	2.522	2.168
British	2.540	2.270
Chinese	2.506	2.055
Danish	2.615	2.174
Dutch	2.487	2.145
Finnich	2.499	2.243
French	2.606	2.258
German	2.616	2.227
Greek	2.571	2.164
Hindi	2.642	2.358
Hungarian	2.583	2.293
Italian	2.770	2.417
Japanese	2.659	2.308
Norwegian	2.562	2.197
Polish	2.621	2.274
Portuguese	2.694	2.370
Russian	2.506	2.119
Spanish	2.627	2.322
Swedish	2.688	2.437
Average	**2.591**	**2.251**

A difference is on average 0.340 PESQ

Speech Intelligibility. Here is a comparison with the MELPe vocoder, which operates at 600 bps. The TWELP 600 bps vocoder and the MELPe 600 bps vocoder were tested using the ITU-T P.50 speech database, covering 20 different languages.
STOI (Short-Time Objective Intelligibility) and ESTOI (Extended Short-Time Objective Intelligibility) metrics were used to assess speech intelligibility:



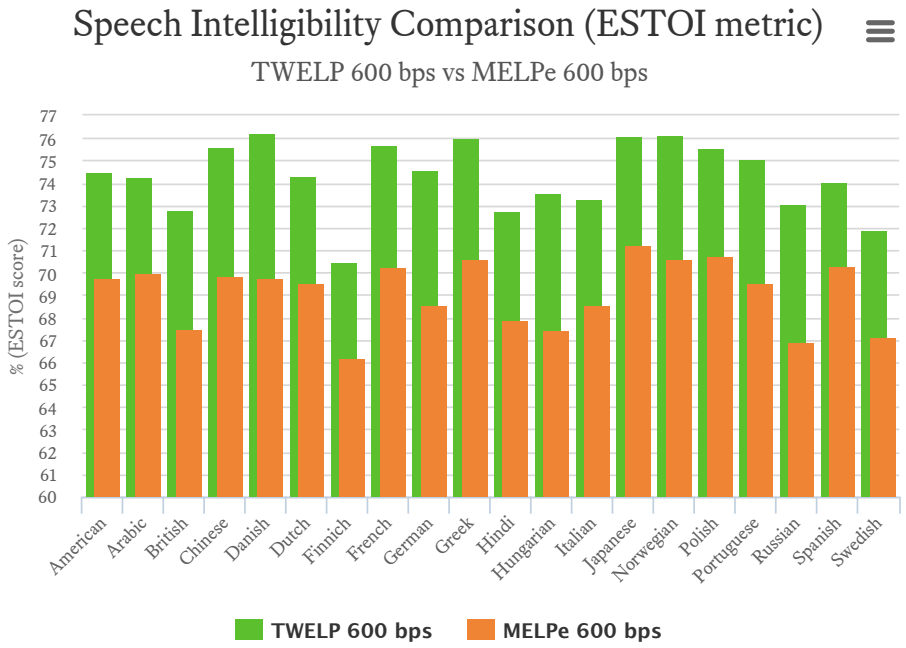
The diagram illustrates a significant difference in speech intelligibility between the TWELP 600 and MELPe 600 vocoders in the STOI metric. Exact values are provided in the table below:

Language	TWELP 600 bps	MELPe 600 bps
American	84.53	79.24
Arabic	83.07	78.50
British	82.02	75.67
Chinese	83.40	77.40
Danish	85.13	79.12
Dutch	83.37	77.04
Finnish	79.64	74.76
French	84.48	78.79
German	84.37	79.00
Greek	83.48	77.89
Hindi	83.97	78.35
Hungarian	84.27	78.14
Italian	83.15	78.03
Japanese	84.14	79.30
Norwegian	84.74	79.24
Polish	84.24	78.37
Portuguese	84.01	78.04
Russian	82.71	75.74
Spanish	82.46	77.82

Swedish	82.31	76.89
Average	**83.48**	**77.90**

A difference is on average 5.58 %

Considering that a low-bitrate vocoder is a nonlinear device that significantly distorts the spectrum of the original speech signal, the ESTOI metric provides more accurate assessments of speech intelligibility after vocoding:



The diagram shows a significant difference in speech intelligibility between the TWELP 600 and MELPe 600 vocoders in the ESTOI metric. Exact numbers are shown in the table below.

Language	**TWELP 600 bps**	**MELPe 600 bps**
American	74.48	69.76
Arabic	74.28	70.04
British	72.80	67.54
Chinese	75.65	69.85
Danish	76.22	69.81
Dutch	74.36	69.58
Finnich	70.44	66.23
French	75.68	70.28
German	74.59	68.55
Greek	76.04	70.59
Hindi	72.77	67.87
Hungarian	73.59	67.46
Italian	73.28	68.60
Japanese	76.09	71.22
Norwegian	76.13	70.59

Polish	75.53	70.79
Portuguese	75.07	69.53
Russian	73.07	66.88
Spanish	74.07	70.33
Swedish	71.91	67.17
Average	**74.30**	**69.13**

A difference is on average 5.17 %

You can see that speech intelligibility testing confirms the significant superiority of the TWELP 600 bps vocoder over the MELPe 600 bps vocoder.

You can download the P.862 and STOI/ESTOI utilities, along with all speech samples, by using the links in the 'Downloads' section at the bottom of the page, and then check all the numbers presented above.

Speech Samples (WAV-files).

A few independent experts compared the TWELP 600 bps vocoder with the MELPe 600 bps vocoder using the preference method.
All listeners preferred the TWELP vocoder, noting its more natural and less synthetic sound.

You can play and listen to short samples of the source speech, as well as the speech processed by the MELPe 600 bps vocoder and the TWELP 600 bps vocoder, using the links in the table below.

You can also download the complete set of P.50 samples as zip files for all languages simultaneously by using the links in the 'Downloads' section at the bottom of the page.

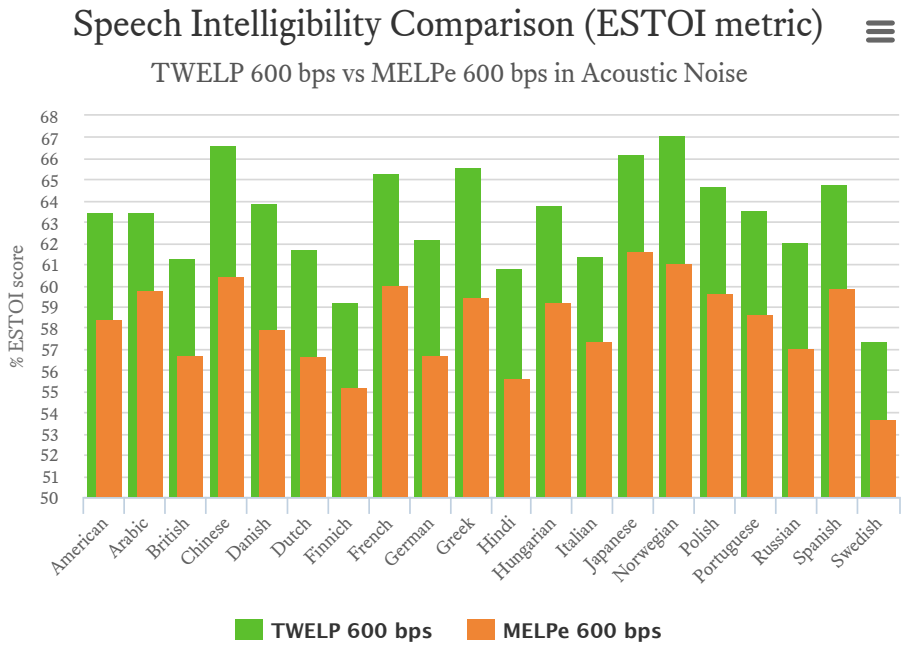
Language	**Source speech**	**MELPe 600 bps**	**TWELP 600 bps**
American	▶	▶	▶
Arabic	▶	▶	▶
British	▶	▶	▶
Chinese	▶	▶	▶
Danish	▶	▶	▶
Dutch	▶	▶	▶
Finnich	▶	▶	▶
French	▶	▶	▶
German	▶	▶	▶
Greek	▶	▶	▶
Hindi	▶	▶	▶
Hungarian	▶	▶	▶
Italian	▶	▶	▶
Japanese	▶	▶	▶
Norwegian	▶	▶	▶
Polish	▶	▶	▶

Portuguese	▶	▶	▶
Russian	▶	▶	▶
Spanish	▶	▶	▶
Swedish	▶	▶	▶

Superiority In Quality Of The Non-speech Signals. In contrast to other LBR vocoders (MELPe, AMBE+2, etc.), TWELP vocoders provide high quality of non-speech signals, including police, ambulance, fire sirens, etc. This feature in conjunction with high quality natural human-sounding of voice makes TWELP vocoders well suitable for replacement of analog radio by digital radio and also for other applications where high quality transmitting of non-speech signals is relevant along with high quality transmitting of speech signals.

Source signal	MELPe 600 bps	TWELP 600 bps
▶	▶	▶

High Robustness To Acoustic Noise. In contrast to other LBR vocoders, TWELP vocoders are highly robust to acoustic noise due to a reliable pitch estimation method and other features of TWELP technology.



The diagram demonstrates a significant difference in speech intelligibility for noisy speech (SNR = 10 dB) between the TWELP 600 and MELPe 600 vocoders, based on the ESTOI metric. Exact numbers are shown in the table below.

Language	TWELP 600 bps	MELPe 600 bps
American	63.51	58.47
Arabic	63.47	59.83
British	61.32	56.73
Chinese	66.61	60.42
Danish	63.88	57.98
Dutch	61.77	56.63
Finnich	59.25	55.25

French	65.31	60.07
German	62.24	56.79
Greek	65.56	59.48
Hindi	60.82	55.59
Hungarian	63.85	59.22
Italian	61.39	57.39
Japanese	66.21	61.64
Norwegian	67.13	61.09
Polish	64.72	59.62
Portuguese	63.59	58.66
Russian	62.02	57.10
Spanish	64.80	59.90
Swedish	57.39	53.69
Average	**63.24**	**58.28**

A difference is on average 4.96 %

Additionally, the TWELP vocoder features an NCSE (Noise Cancellation Speech Enhancement) preprocessor, which cleans the input speech signal from noise and enhances speech quality.
Below, you can listen to a short fragment of heavily noisy English speech after passing through MELPe and TWELP vocoders, with NPP (Noise Pre-Processor) and NCSE disabled and enabled, respectively.

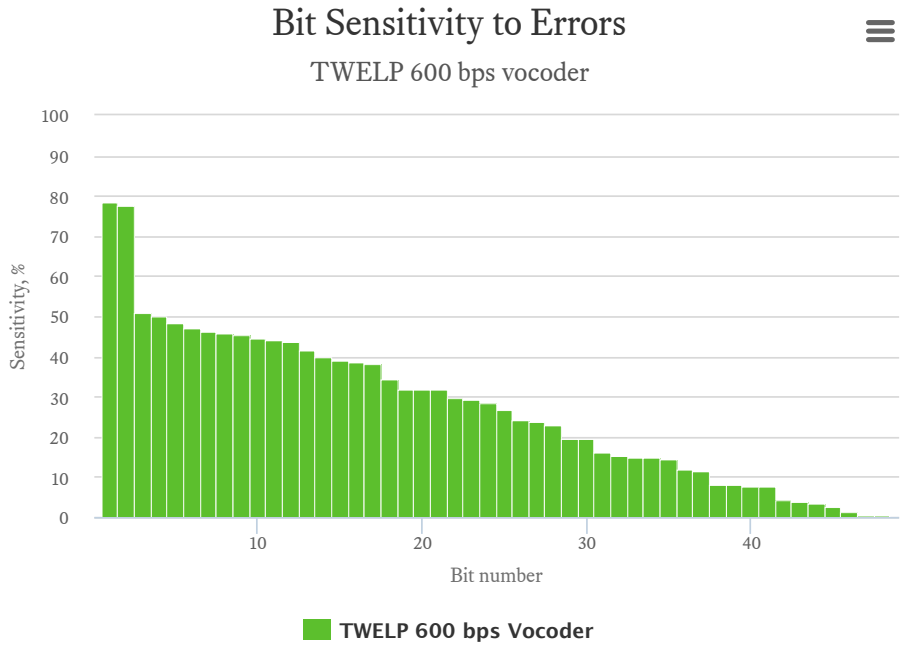
NPP/NCSE	Input speech (SNR=10dB)	MELPe 600 bps	TWELP 600 bps
Disabled	▶	▶	▶
Enabled	▶	▶	▶

The NCSE integrated into the TWELP vocoder is described in more detail on the webpage for our standalone product, 'NCSE-AGC Preprocessor'.

High Robustness To The Channel Errors.
The TWELP technology offers highly efficient speech compression by eliminating redundancy while preserving excellent quality and intelligibility. To enhance robustness against transmission errors, we provide specialized versions called **TWELP Robust**.
These vocoders are based on an effective Joint Source-Channel Coding approach. Each vocoder is equipped with a custom-designed FEC, tailored to its specific characteristics and operational conditions.
TWELP Robust vocoders provide high speech quality simultaneously in noisy channel as well as in noiseless channel. FEC can operate with "soft decisions" as well as with "hard decisions" from a modem. "Soft decisions" mode provides much better robustness in comparison with the "hard decisions" mode.
For all users of our non-robust vocoder versions, we offer the following recommendations.

The diagram below illustrates the sensitivity of bits at the output of the vocoder to communication channel errors.
Essentially, the diagram shows by what percentage speech quality is reduced when a specific bit is distorted. The first bits in order cause catastrophic distortions, while the latter bits have significantly less impact on

quality.



We strongly recommend using FEC (Forward Error Correction) with unequal protection of the bits in strong accordance with their sensitivity to errors and utilizing 'Soft Decisions' decoding. This will provide the highest robustness of the vocoder against errors in the channel.

Additional Functionalities. The following additional functionalities are developed by DSPINI and integrated into TWELP vocoders:

- Noise Cancellation Speech Enhancement (NCSE)
- Automatic Gain Control (AGC),
- Voice Activity Detector (VAD),
- Discontinuous Transmission (DTX),
- Tone Detection/Generation (Single tones and Dual tones). The tones are transmitted by the vocoder facilities.

Note:
The Tone Detector/Generator functionalities are not integrated into the code by default but can be added free of charge upon request.

Each functionality has unique features, performance and characteristics, providing significant superiority over any well-known implementations on the market.

Technical Characteristics And Resource Requirements:

Technical characteristics

Bit Rate (bps)	Algorithm	**Frame size** (ms)	**Algorithmic delay (including frame size)** (ms)	**Sampling rate** (kHz)	**Signal format**	**Bit stream format**
600	TWELP	80	100	8	Linear 16-bit PCM	48

Additional functionalities

Name	**Functionality**	**Technical characteristics**	
		Name	**Value**
AGC	Automatic Gain Control	Control range:	0 ... +42 dB
NCSE	Noise Canceller - Speech Enhancer	SNR increasing	20 dB
		Speech quality improvement	> 0.1 PESQ
Tone Detector	Single/Dual tones detection	In accordance with international standards	
Tone Generator	Single/Dual tones generation	Special generator, kept continuity of signal (phase and amplitude of signal of previous frame)	
DTX	Discontinuous Transmission	Reduces bit rate down to 110 bps in pauses between active speech regions	
VAD	Voice Activity Detection	High reliability even with pink noise at an SNR < 0 dB.	
CNG	Comfort Noise Generation	Type of noise	"white"
		Level	- 60 dB

The NCSE and AGC integrated into the TWELP vocoder are described in more detail on the webpage for our standalone product, 'NCSE-AGC Preprocessor'.

Resources for ARM Cortex-M4 platform

Module	**MIPS* peak**	**Memory** (KBytes)				
		Program	**Data**			
			Constants	**Channel**	**Heap**	**Stack**
Voice Encoder	53	42	201	4.7	5.7	1.0
NCSE	6.0					
AGC	0.5					
Voice Decoder	19					
Voice Encoder + Voice Decoder	72					
Total	78.5					

Resources for TI's C64 DSP platform

Module	**MIPS* peak**	**Memory** (KBytes)				
		Program	**Data**			
			Constants	**Channel**	**Heap**	**Stack**
Voice Encoder	19	80	201	4.7	5.7	1.0
NCSE	2.6					
AGC	0.3					
Voice Decoder	5.3					
Voice Encoder + Voice Decoder	24.3					
Total	27.2					

Resources (estimated) for TI′s C55 DSP platform

Module	MIPS* peak	Memory (KBytes)				
		Program	Data			
			Constants	Channel	Heap	Stack
Voice Encoder	32	26	201	4.7	5.7	1.0
NCSE	6.2					
AGC	0.4					
Voice Decoder	13					
Voice Encoder + Voice Decoder	45					
Total	51.6					

* DSPINI continues optimization of the TWELP algorithm and code in order to minimize computational complexity of the vocoder.

Software Integrity and Security. DSPINI guarantees the ABSOLUTE integrity of its software, free from any undocumented features, undeclared capabilities, or hidden functions. Our customers can be assured that none of our software/code contains any secret features or functionalities concealed from the user. If necessary, we are ready to provide the source code of our software products for appropriate certification. Moreover, our software is available in source code form—you simply need to purchase the appropriate license to use it.

Guarantee And Support. DSPINI guarantees a quality and accordance of all technical characteristics of the product to requirement of current specifications. Testing and other method of quality control are used for guarantee support.

Any Platforms. DSPINI can port this vocoder software into any other DSP, RISC or general- purposes platform inshort time: 1-2 months.

Licensing Terms. To use the vocoder, customer should obtain a license from DSPINI only.

Customization. The vocoder can be customized under any specific requirements- other bit rate, frame size, any other robustness to channel errors, etc. Please contact with us for details.

Prospects. DSPINI is impoving and developing continuously a set of new vocoders with range from 300 bps up to 9600 bps, based on TWELP technology.

Related Software. This vocoder may be effectively used in a bundle with other DSPINI′s products:

- Linear and acoustic echo cancellers,
- Multichannel noise cancellers (including two-microphone adaptive array),
- Wired or radiomodems for any types of channels and bitrates,

- Other products.

Downloads:

- Datasheet (pdf)
- ITU-T P.50 source speech samples (zip)
- MELPe 600 bps speech samples (zip)
- TWELP 600 bps speech samples (zip)
- P.862 and STOI/ESTOI utilities
- PC-evaluation package (zip) — on request
- User's Guide document (pdf) — on request