

```
bundle = torchaudio.pipelines.WAV2VEC2_ASR_BASE_960H
```

예제에는 해당 모델을 사용하여 예셋에서 음성을 인식하는 과정을 거친다.

하지만, 본인이 가져온 **batman.mp3**파일은 기존에 존재하는 예셋보다 발화가 빠르기 때문에 인식 성능이 크게 저하되는 문제가 있다.

인식 상태:

```
<class 'torchaudio.models.wav2vec2.model.Wav2Vec2Mod
S|WHAT|YOU|PROPOSIN|IT|SIMPLE|KILL|THE|BAT|MAN|
class labels' ('=' '!' 'E' 'I' 'A' 'O' 'N' 'L'
```

원본 대사:

So What are you proposing?

It's simple. Kill the Batman.

따라서 해당 문제를 해결하기 위해 모델을 다음과 같이 바꾸자 해결되었다.

WAV2VEC2\_ASR\_BASE\_960H -> WAV2VEC2\_ASR\_LARGE\_960H

```
<class 'torchaudio.models.wav2vec2.model.Wav2Vec2Mod
SO|WHAT|ARE|YOU|PROPOSING|IT'S|SIMPLE|KILL|THE|BAT|MAN|
class labels' ('=' '!' 'E' 'I' 'A' 'O' 'N' 'L'
```