

TÀI LIỆU HƯỚNG DẪN THÍ NGHIỆM CẢM THỤ TIẾNG NÓI

1. Giới thiệu chung

Thí nghiệm cảm thụ (perception test) là bài thí nghiệm có sự tham gia con người (chuyên gia, tình nguyện viên) trong việc đánh giá chất lượng của các giọng nói khác nhau. Người tham gia cần sử dụng tai nghe chất lượng tốt trong phòng khá yên tĩnh và cần tập trung hoàn toàn vào công việc để không ảnh hưởng tới chất lượng của bài thí nghiệm.

Cuộc thi VLSP TTS 2020 có hai loại thí nghiệm cảm thụ:

- MOS Test: Có 2 bài test, bạn cần tham gia cả 2 bài test này. Với MOS Test, bạn cần nghe một đoạn tiếng nói, sau đó đánh giá bằng cách cho điểm theo hướng dẫn. Khi tham gia bài test này, bạn nên nghe đoạn tiếng nói đó 1 lần, trừ trường hợp bị lơ là chưa kịp nghe rõ thì mới nghe lại lần thứ 2.
- Transcription Test: Có 3 bài test, bạn cần tham gia ít nhất 2 trong 3 bài test này. Với Transcription Test, bạn cần nghe một đoạn tiếng nói, sau đó gõ lại text của đoạn tiếng nói đó. Bạn có thể nghe lại lần 2 nếu chưa nghe rõ hoặc câu dài. Chỉ nghe lại lần 3 khi bạn lơ là hoặc câu rất dài.

Lưu ý, bạn cần tham gia cả 2 bài MOS Test và ít nhất 2 trong 3 bài Transcription Test. Khuyến khích các bạn tham gia cả 5 bài test. Các bài test là các bài độc lập, các bạn có thể tham gia vào các thời gian khác nhau. Các bạn nên nghỉ 10-15 phút giữa các bài test để đảm bảo có những đánh giá chính xác.

Các phần tiếp theo của tài liệu này sẽ đưa ra các hướng dẫn chi tiết cho loại thí nghiệm MOS Test và Transcription Test.

2. Đăng nhập


- Sau khi được cấp tài khoản, bạn vào <https://tts.vlsp.org.vn/> và tiến hành đăng nhập.

VLSP Vietnamese Text To Speech

Đăng ký

Đăng nhập

Đăng nhập

 Đăng nhập vào tài khoản của bạn

Đăng nhập

Bạn chưa có tài khoản? [Đăng ký](#)

- Sau đăng nhập thành công lần đầu sau khi được cấp tài khoản, bạn cần thay đổi mật khẩu để tiếp tục sử dụng hệ thống.



Đổi mật khẩu

Xác nhận

3. Lựa chọn bài thí nghiệm cảm thụ

Sau khi đổi mật khẩu thành công, bạn có thể chọn thực hiện một trong 5 bài thí nghiệm VLSP TTS 2020 dưới đây. Lưu ý rằng tùy thuộc vào việc số lượng bài test bạn đăng ký thì các bài test sẽ được hiển thị tương ứng với từng tài khoản sau khi đăng nhập.

- VLSP TTS 2020 – MOS Test 1
- VLSP TTS 2020 – MOS Test 2
- VLSP TTS 2020 – Transcription Test 1
- VLSP TTS 2020 – Transcription Test 2
- VLSP TTS 2020 – Transcription Test 3

VLSP Vietnamese Text To Speech
Xin chào, **Mạnh**
Đổi mật khẩu
Đăng xuất

Danh sách các bài test bạn có thể tham gia

Tên bài test	Số lượng audio	Ngày bắt đầu	Ngày kết thúc	
No Data				

Danh sách các bài test bạn được chỉ định tham gia

Tên bài test	Số lượng audio	Ngày bắt đầu	Ngày kết thúc	
VLSP TTS – MOS Test 1	60	20 - 10 - 2020	31 - 12 - 2020	Thực hiện tiếp
VLSP TTS – MOS Test 2	75	20 - 10 - 2020	31 - 12 - 2020	Đã hoàn thành
VLSP TTS – Transcription Test 1	36	20 - 10 - 2020	31 - 12 - 2020	Bắt đầu test
VLSP TTS – Transcription Test 2	36	20 - 10 - 2020	31 - 12 - 2020	Bắt đầu test

< 1 >

Như hình trên có thể thấy:

- *Số lượng audio* là số lượng audio mà bạn cần nghe trong mỗi bài test
- *Ngày bắt đầu* là ngày bài thí nghiệm cảm thụ có thể bắt đầu. Bạn không thể bắt đầu sớm hơn ngày này.
- *Ngày kết thúc* là deadline cho việc đánh giá bài test. Quá ngày này, bạn sẽ không thấy bài thí nghiệm này xuất hiện nữa tương ứng với việc bạn không thể thực hiện bài thí nghiệm này nữa.

Trong “Danh sách các bài test bạn được chỉ định tham gia”, bạn có thể:

- “*Bắt đầu test*”: Bài test hiển thị nút này nếu bạn chưa bắt đầu bài thí nghiệm này. Khi nhấn vào nút này, bạn sẽ được bắt đầu đánh giá từ câu đầu tiên của bài thí nghiệm.
- “*Thực hiện tiếp*”: Bài test hiển thị nút này nếu bạn đang thực hiện bài test này nhưng chưa xong. Bạn có thể nhấn vào nút “Thực hiện tiếp” để tiếp tục bài test này. Hệ thống sẽ chuyển sang câu đầu tiên bạn chưa làm tính từ đầu danh sách.
- “*Đã hoàn thành*”: Với bài test bạn đã thực hiện, bạn sẽ không thể thực hiện thao tác gì trên đó. Vì vậy, bài test này mang tính chất liệt kê và có trạng thái là “Đã hoàn thành”.

4. MOS Test

Khi chọn một bài thí nghiệm MOS Test, màn hình chính được hiển thị như dưới đây.

VLSP

Vietnamese Text To Speech

Xin chào, Vũ Duy Mạnh

Đổi mật khẩu

Đăng xuất

Câu	Điểm
Câu 1	
Câu 2	
Câu 3	
Câu 4	
Câu 5	
Câu 6	
Câu 7	
Câu 8	
Câu 9	
Câu 10	

1

2

3

4

5

6

10 / page

Bạn đang đánh giá câu 1/ 60. Lượt nghe: 1

Nội dung:

Câu nhàu nói với Gonz bằng thứ tiếng mà Robert Jordan không hiểu

0:01 / 0:03

Đánh giá chất lượng giọng nói

☐ 5 điểm - Rất tự nhiên, giống như giọng người thật.

☐ 4 điểm - Tương đối tự nhiên, khá giống giọng người thật

☐ 3 điểm - Hơi tự nhiên, khá nhiều yếu tố nhân tạo

☐ 2 điểm - Kém tự nhiên, rất nhiều yếu tố nhân tạo

☐ 1 điểm - Rất kém tự nhiên, hoàn toàn nhân tạo

< Câu trước

Kết thúc

Câu tiếp >

Bên tay phải là phần đánh giá từng đoạn tiếng nói bạn nghe được. Nội dung tiếng nói được hiển thị cùng trên trang với đoạn tiếng nói tương ứng. Nhiệm vụ của bạn là nghe đoạn tiếng nói và đánh giá chất lượng đoạn tiếng nói tương ứng với nội dung đó. Khi tham gia bài test này, bạn nên nghe đoạn tiếng nói đó 1 lần, trừ trường hợp bị lơ là chưa kịp nghe rõ thì mới nghe lại lần thứ 2. Có 5 mức điểm bạn có thể chọn như được mô tả dưới đây, bạn chỉ có thể chọn các mức điểm này khi đã bạn nghe đoạn audio ít nhất một lần:

- **5 điểm - Rất tự nhiên, giống như giọng người thật:** Bạn hãy cho mức điểm này nếu tiếng nói nghe được rất tự nhiên, nghe giống như giọng người thu âm, không nghe thấy yếu tố nhân tạo trong đó. Có thể dùng giọng này để trao đổi, tương tác như tương tác với người.
- **4 điểm - Tương đối tự nhiên, khá giống giọng người thật:** Bạn hãy cho mức điểm này nếu tiếng nói nghe được khá tự nhiên và khá giống giọng người thu âm, có một chút yếu tố nhân tạo nhưng không đáng kể. Có thể dùng giọng này để trao đổi, tương tác được mặc dù có đôi chỗ còn chưa hoàn hảo.
- **3 điểm - Hơi tự nhiên, khá nhiều yếu tố nhân tạo:** Bạn hãy cho mức điểm này nếu tiếng nói nghe được có thể hiểu và có thể dùng để giao tiếp nhưng còn khá nhiều yếu tố nhân tạo.
- **2 điểm - Kém tự nhiên, rất nhiều yếu tố nhân tạo:** Bạn hãy cho mức điểm này nếu tiếng nói nghe được còn kém tự nhiên và có rất nhiều yếu tố nhân tạo. Để giao tiếp với tiếng nói này thì đôi chỗ còn khó nghe.
- **1 điểm - Rất kém tự nhiên, hoàn toàn nhân tạo:** Bạn hãy cho mức điểm này nếu tiếng nói nghe được rất kém tự nhiên và khó dùng để giao tiếp. Tiếng nói hoàn toàn nhân tạo và nhiều chỗ còn khó nghe.

VLSP

Vietnamese Text To Speech

Xin chào, Vũ Duy Mạnh

Đổi mật khẩu

Đăng xuất

Câu	Điểm
Câu 1	
Câu 2	
Câu 3	
Câu 4	
Câu 5	
Câu 6	
Câu 7	
Câu 8	
Câu 9	
Câu 10	

1

2

3

4

5

6

10 / page

Bạn đang đánh giá câu 1/ 60. Lượt nghe: 1

Nội dung:

Câu nhân nói với Gonz bằng thứ tiếng mà Robert Jordan không hiểu

0:03 / 0:03

Đánh giá chất lượng giọng nói

☐ 5 điểm - Rất tự nhiên, giống như giọng người thật.

☐ 4 điểm - Tương đối tự nhiên, khá giống giọng người thật

☐ 3 điểm - Hơi tự nhiên, khá nhiều yếu tố nhân tạo

☐ 2 điểm - Kém tự nhiên, rất nhiều yếu tố nhân tạo

☐ 1 điểm - Rất kém tự nhiên, hoàn toàn nhân tạo

< Câu trước

Kết thúc

Câu tiếp >

- Khi chọn xong, bạn chọn **Câu tiếp >** để ghi nhận điểm và chuyển sang câu tiếp theo. Ngoài ra, bạn có thể chọn **< Câu trước** để đánh giá lại câu trước đó. Nút “Kết thúc” sẽ được ẩn đi **Kết thúc** cho đến khi nào bạn hoàn thành tất cả các câu, nút này sẽ hiện lên cho bạn nộp bài.
- Bên tay trái là bảng ghi lại kết quả đánh giá của bạn. Ngoài ra, bạn có thể click vào các câu để đánh giá lại hoặc nhảy trực tiếp đến câu nào đó để đánh giá. VD: Khi đánh giá xong câu 5, hệ thống sẽ tự động nhảy đến câu 6. Nhưng vì một lý do nào đó, bạn muốn đánh giá lại câu 4. Lúc này, bạn chỉ cần click vào câu 4 để tiến hành đánh giá lại.
- Khi đánh giá xong tất cả các câu, nút “Kết thúc” sẽ sáng lên để bạn hoàn thành bài test như hình dưới đây

Đánh giá chất lượng giọng nói

☐ 5 điểm - Rất tự nhiên, giống như giọng người thật.

☒ 4 điểm - Tương đối tự nhiên, khá giống giọng người thật

☐ 3 điểm - Hơi tự nhiên, khá nhiều yếu tố nhân tạo

☐ 2 điểm - Kém tự nhiên, rất nhiều yếu tố nhân tạo

☐ 1 điểm - Rất kém tự nhiên, hoàn toàn nhân tạo

< Câu trước

Kết thúc

Câu tiếp >

- Bạn có thể nghe và chỉnh sửa lại đánh giá của mình trước khi nộp bài. Khi nhấn vào nút “Kết thúc”, bạn sẽ không được thay đổi cũng như xem lại bài của mình nữa.

5. Transcription Test

Khi vào bài thí nghiệm Transcription Test, màn hình chính được hiển thị như dưới đây.

- Bên tay phải là phần nhập nội dung tiếng nói mà bạn nghe được. Nhiệm vụ của bạn là nghe đoạn tiếng nói và nhập nội dung tiếng nói mà bạn nghe được. Bạn có thể nghe lại lần 2 nếu chưa nghe rõ hoặc câu dài. Chỉ nghe lại lần 3 khi bạn lơ là hoặc câu rất dài.
- Tương tự như MOS Test, bạn chỉ có thể gõ nội dung câu khi bạn đã nghe đoạn tiếng nói ít nhất 1 lần.
- Khi nhập xong, bạn chọn để **Câu tiếp >** ghi nhận kết quả và chuyển sang câu tiếp theo. Ngoài ra, bạn có thể chọn **< Câu trước** để nhập lại nội dung câu trước đó. Nút **Kết thúc** sẽ là nút được ẩn đi cho đến khi nào bạn hoàn thành tất cả các câu, nút này sẽ hiện lên cho bạn nộp bài.

- Bên tay trái là bảng ghi lại trạng thái các câu của bạn (Đã hoàn thành hay chưa). Ngoài ra, bạn có thể click vào các câu để gõ lại hoặc nhảy trực tiếp đến câu nào đó để gõ nội dung tiếng nói.
- Khi đánh giá xong tất cả các câu nút “Kết thúc” sẽ sáng lên để bạn hoàn thành bài test.
- Bạn có thể nghe và chỉnh sửa lại nội dung tiếng nói mà mình nghe được trước khi nộp bài. Khi nhấn vào nút “Kết thúc”, bạn sẽ không được thay đổi cũng như xem lại bài của mình nữa.