# Pairwise comparisons of `WER` using other noise generation techniques

| | Librispeech | | | Commonvoice | | | Timit | | |
|---|---|---|---|---|---|---|---|---|---|
| | Deepspeech | Sphinx | Google | Deepspeech | Sphinx | Google | Deepspeech | Sphinx | Google |
| All VS Random | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.0039 |
| All VS Important | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.033 |
| Important VS Random | 0.001 | 0.001 | 0.001 | 0.001 | 0.0461 | **0.9** | **0.158** | **0.5847** | **0.4916** |

Table 1: P-values using One way Anova and Tukey's HSD for pairwise comparison of frame selection methods with respect to `WER` achieved. `GL` is used for noise generation.

| | Librispeech | | | Commonvoice | | | Timit | | |
|---|---|---|---|---|---|---|---|---|---|
| | Deepspeech | Sphinx | Google | Deepspeech | Sphinx | Google | Deepspeech | Sphinx | Google |
| All VS Random | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
| All VS Important | 0.001 | 0.001 | 0.034 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.0171 |
| Important VS Random | 0.001 | 0.001 | 0.0033 | **0.53** | **0.1637** | **0.9** | **0.9** | **0.7507** | **0.5239** |

Table 2: P-values using One way Anova and Tukey's HSD for pairwise comparison of frame selection methods with respect to `WER` achieved. `WN` is used for noise generation.

As can be seen from the table 2, when using WN, there is no big difference between Random and Important on Commonvoice. This is because the white noise of this group is more focused on high similarity and lacks aggressiveness when setting. So Random and Important are not much different.

| | Librispeech | | | Commonvoice | | | Timit | | |
|---|---|---|---|---|---|---|---|---|---|
| | Deepspeech | Sphinx | Google | Deepspeech | Sphinx | Google | Deepspeech | Sphinx | Google |
| GL VS WN | 0.001 | 0.001 | 0.001 | 0.001 | 0.0085 | 0.001 | 0.001 | 0.001 | 0.001 |
| OP VS WN | 0.001 | 0.001 | 0.001 | 0.001 | 0.0229 | 0.001 | 0.001 | 0.001 | 0.001 |
| OP VS GL | 0.001 | **0.9** | 0.001 | **0.9** | **0.9** | **0.9** | **0.7692** | **0.9** | **0.9** |

Table 3: P-values using One way Anova and Tukey's HSD for pairwise comparison of noise generation mehtods with respect to `WER` achieved. `Random` frames is used as the frame selection method.

| | Librispeech | | | Commonvoice | | | Timit | | |
|---|---|---|---|---|---|---|---|---|---|
| | Deepspeech | Sphinx | Google | Deepspeech | Sphinx | Google | Deepspeech | Sphinx | Google |
| GL VS WN | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
| OP VS WN | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
| OP VS GL | 0.001 | **0.7251** | 0.001 | **0.9** | **0.6951** | **0.9** | **0.9** | **0.9** | **0.9** |

Table 4: P-values using One way Anova and Tukey's HSD for pairwise comparison of noise generation mehtods with respect to `WER` achieved. `Important` frames is used as the frame selection method.

# Pairwise comparisons of `Similarity` using other noise generation techniques

Table 5: P-values using One way Anova Tukey's HSD for pairwise comparison of `Similarity` achieved by different frames selection technologies. (They are all based on the `WN` technology of noise generation part)

| | Librispeech | Commonvoice | Timit |
|---|---|---|---|
| Random VS All | 0.001 | 0.001 | 0.001 |
| Important VS ALL | 0.001 | 0.001 | 0.001 |
| Random VS Important | 0.001 | 0.0470 | **0.5438** |

Table 6: P-values using One way Anova Tukey's HSD for pairwise comparison of `Similarity` achieved by different frames selection technologies. (They are all based on the `GL` technology of noise generation part)

| | Librispeech | Commonvoice | Timit |
|---|---|---|---|
| Random VS All | 0.001 | 0.001 | 0.001 |
| Important VS ALL | 0.001 | 0.001 | 0.001 |
| Random VS Important | 0.001 | 0.0411 | **0.7929** |

Table 7: P-values using One way Anova Tukey's HSD for pairwise comparison of `Similarity` achieved by different noise generation technologies. (They are all based on the `Random` technology of frames selection part.)

|          | Librispeech | Commonvoice | Timit  |
| -------- | ----------- | ----------- | ------ |
| GL VS WN | 0.001       | 0.001       | **0.7334** |
| OP VS WN | 0.001       | 0.001       | **0.802** |
| OP VS GL | **0.4847**  | **0.8261**  | **0.3887** |

Table 8: P-values using One way Anova Tukey's HSD for pairwise comparison of `Similarity` achieved by different noise generation technologies. (They are all based on the `Important` technology of frames selection part.)

|          | Librispeech | Commonvoice | Timit  |
| -------- | ----------- | ----------- | ------ |
| GL VS WN | 0.001       | 0.001       | **0.0568** |
| OP VS WN | 0.001       | 0.001       | **0.4622** |
| OP VS GL | **0.2978**  | **0.6312**  | **0.5102** |