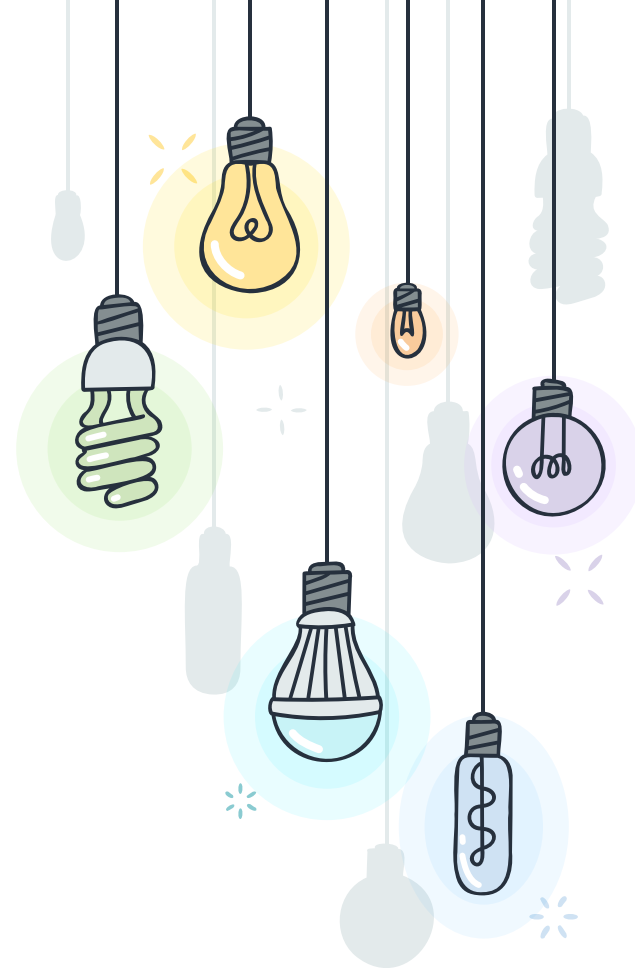


# ROZPOZNAWANIE JĘZYKA POLSKIEGO

Karolina Głowińska, Julia Kułacz, Martyna Pawlus

# \* AGENDA

1. Cel projektu
2. Silniki transkrypcji
3. Obiekt badań
4. Czynniki poprawności transkrypcji
5. Wyniki
6. Propozycje przyszłych badań

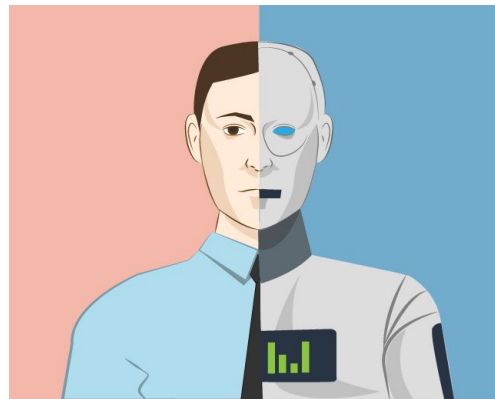
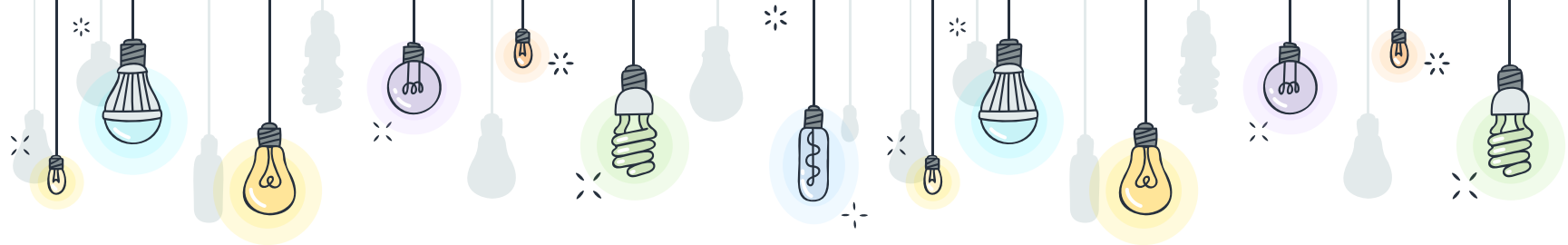


1

# CEL PROJEKTU

Czyli po co się tu zebraliśmy





# CEL PROJEKTU

Porównanie silników transkrypcji mowy w języku polskim oraz ocena wpływu różnych czynników charakteryzujących nagrania audio.

# 2

## SILNIKI TRANSKRYPCJI

Czyli komercyjne narzędzia do  
zamiany mowy na tekst





3

## OBIEKT BADAŃ

Czyli sety z testowanymi nagraniami  
audio



## SET 1



Fragment Kordiana - Juliusz  
Słowacki

## SET 2



Fragment odcinka Pasja Informatyki  
(YouTube) - Mirosław Zelent i  
Damian Stelmach





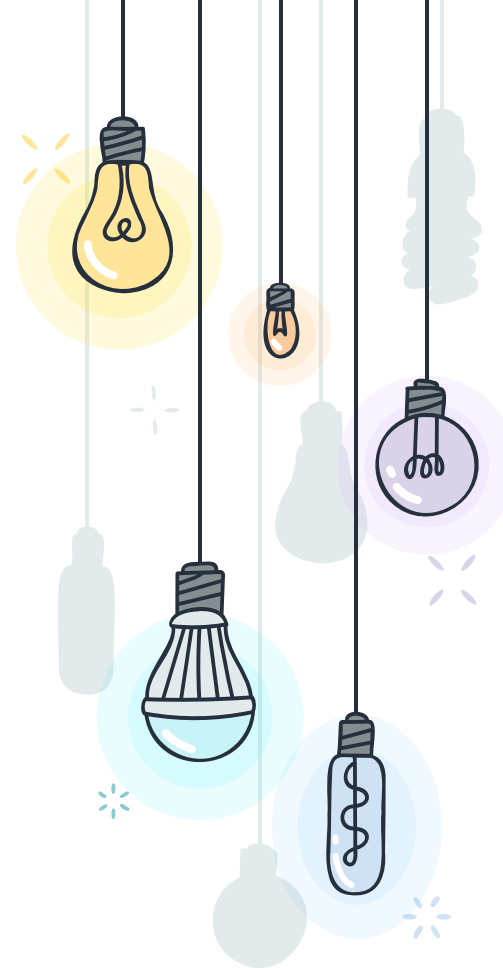
# 4

## CZYNNIKI POPRAWNOŚCI TRANSKRYPCJI

Czyli co może decydować o dobrej  
jakości zamiany mowę na tekst



OBIEKT ANALIZY	SET A	SET B
Dykcja	Audiobooki: dobra dykcja	Audiobooki: zła dykcja
Szum tła	Audiobooki: dobra dykcja	Audiobooki: dobra dykcja + szum
Mikrofon	Audiobooki: zła dykcja	Audiobooki: zła dykcja + zły mikrofon
Płeć	kobiety	mężczyźni
Ogólne porównanie setów	Set 1	Set 2



ZŁA DYKCJA & ZŁY MIKROFON



DOBRA DYKCJA & SZUM



ZŁA DYKCJA & DOBRY MIKROFON



5

# WYNIKI

Czyli to co nas najbardziej interesuje 🐱



# JAK PORÓWNAĆ OTRZYMANE TRANSKRYPCJE?

**WER (Word Error Rate)** - parametr dokładności tekstu

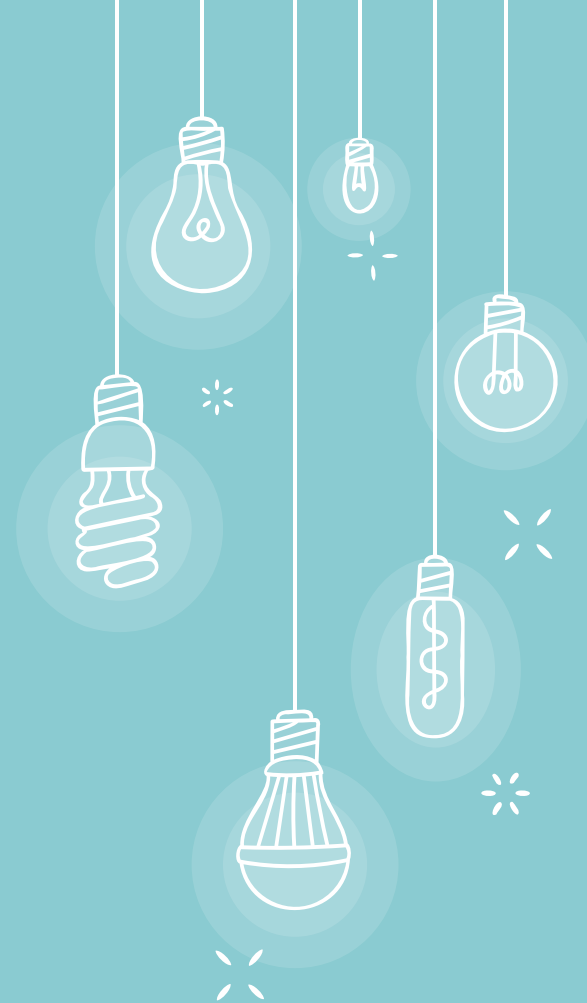
$$WER = \frac{S + D + I}{S + D + C}$$

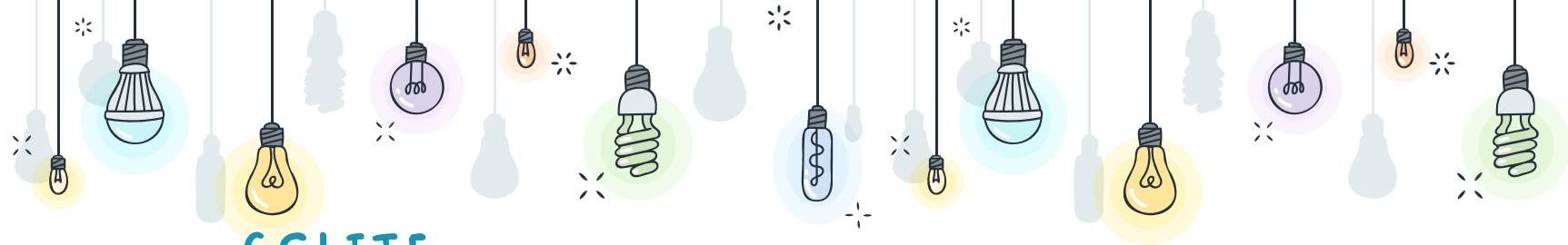
*S* - substitution

*D* - deletions

*I* - insertions

*C* - correct words





# SCLITE - NARZĘDZIE DO PORÓWNYWANIA TEKSTÓW

S

```
id: (speaker 1)
Scores: (#C #S #D #I) 198 28 1 7
REF: dzień był letni i świąteczny wszystko na świecie JAŚNIAŁO kwitło pachniało śpiewało ciepło i radość lały się
HYP: dzień był letni i świąteczny wszystko na świecie JASŁO kwitło pachniało śpiewało ciepło i radość lały się
Eval: S
```

D

```
one trochę jak walec KTÓRY miażdży ognisko choroby ale rozjeżdża również wszystko wokół TYMCZASEM człowiek to nies
trochę jak walec ***** miażdży ognisko choroby ale rozjeżdża również wszystko wokół ***** człowiek to niesan
D D
```

I

```
dziecka nie uprowadził pochłaniacz ** BOSS gangu obcinaczy palców socjopata i protegowany dziadka b
dziecka nie uprowadził pochłaniacz BO Z gangu obcinaczy palców socjopata i protegowany dziadka
I S
```



# SCLITE - PRZYKŁADOWY RAPORT

## SYSTEM SUMMARY PERCENTAGES by SPEAKER

/home/julia/Desktop/Google - Speech Recognition/1_bonda_k_m1_d1_s0_118_2.trn									
SPKR	# Snt	# Wrđ	Corr	Sub	Del	Ins	Err	S.Err	
spe	1	232	93.1	6.0	0.9	1.7	8.6	100.0	
Sum/Avg	1	232	93.1	6.0	0.9	1.7	8.6	100.0	
Mean	1.0	232.0	93.1	6.0	0.9	1.7	8.6	100.0	
S.D.	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
Median	1.0	232.0	93.1	6.0	0.9	1.7	8.6	100.0	

# JAK PRZEDSTAWIĆ ZEBRANE WYNIKI?

**Test ANOVA** - jednoczynnikowa analiza wariancji

**P-value** - prawdopodobieństwo testowe

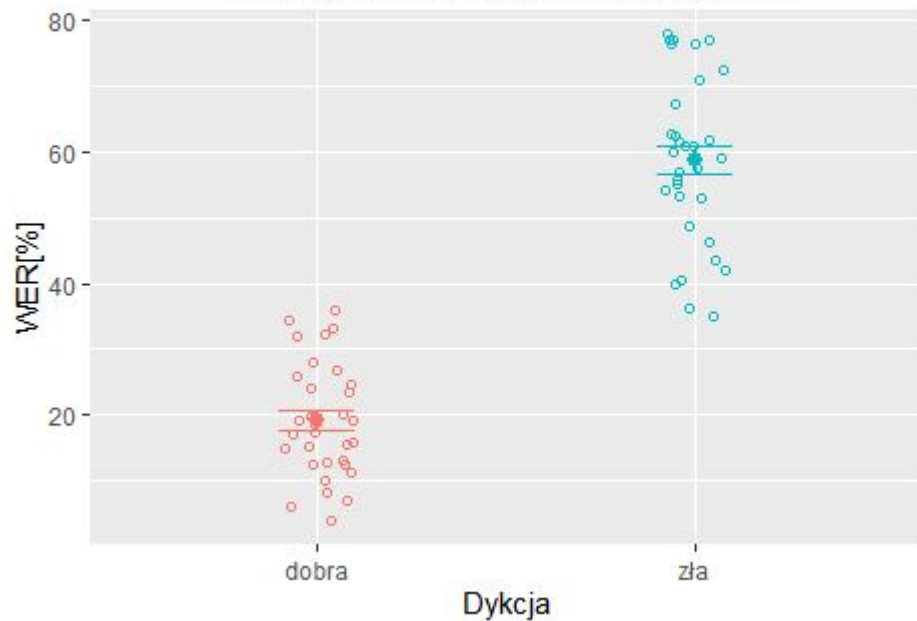
```
ANOVA table  
Df Sum Sq Mean Sq F value Pr(>F)  
Dykcja      1    25121    25121    209.3 <2e-16 ***  
Residuals   62     7440      120  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
ANOVA table  
Df Sum Sq Mean Sq F value Pr(>F)  
Szum.tła    1    1704    1704.3    11.1 0.00146 **  
Residuals   62    9518     153.5  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



# \* DYKCJA

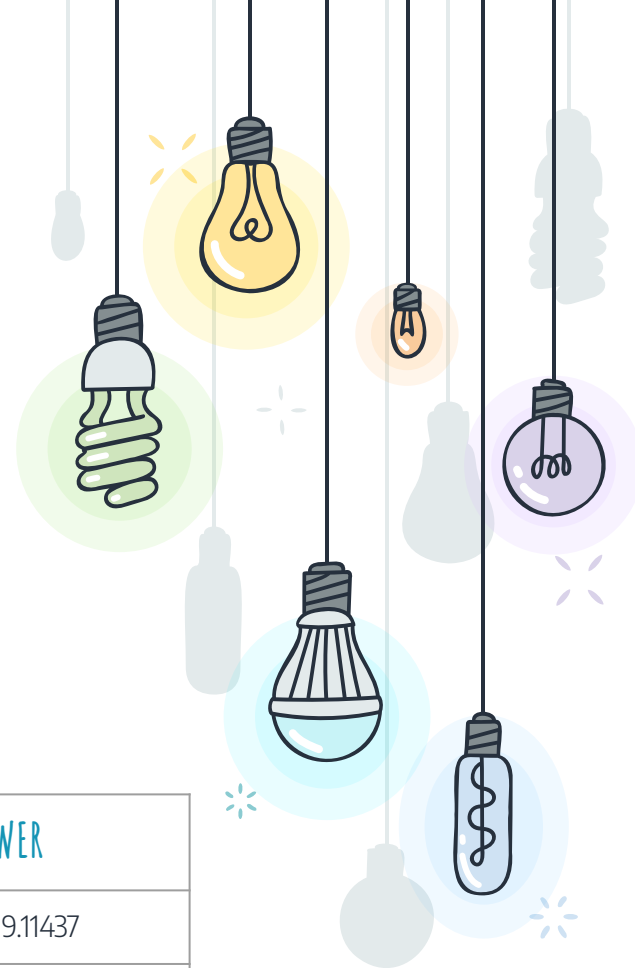
Zestawienie wartości WER  
dla dobrej i złej dykcji lektora (set 1)



Dykcja

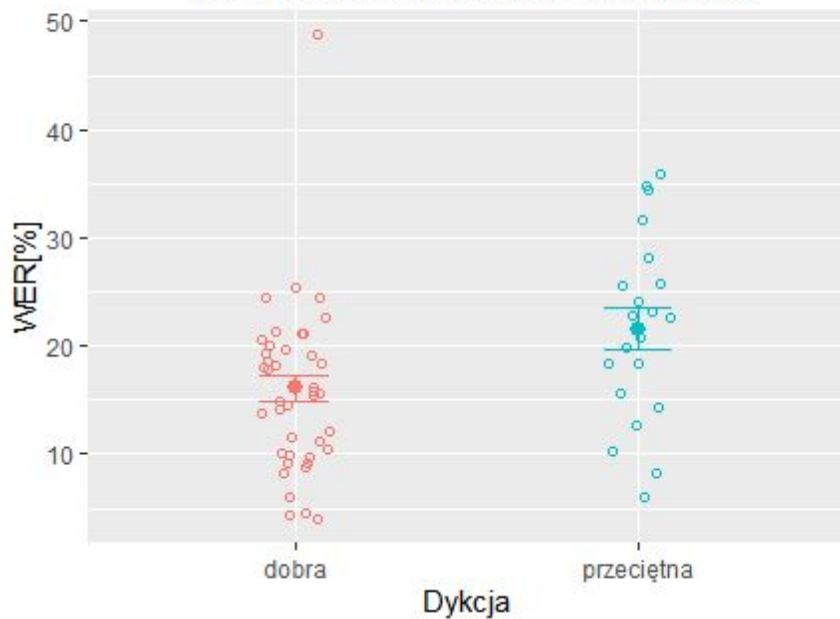
dobra  
zła

DYKCJA	WER
dobra	19.11437
zła	58.73844



# \* DYKCJA

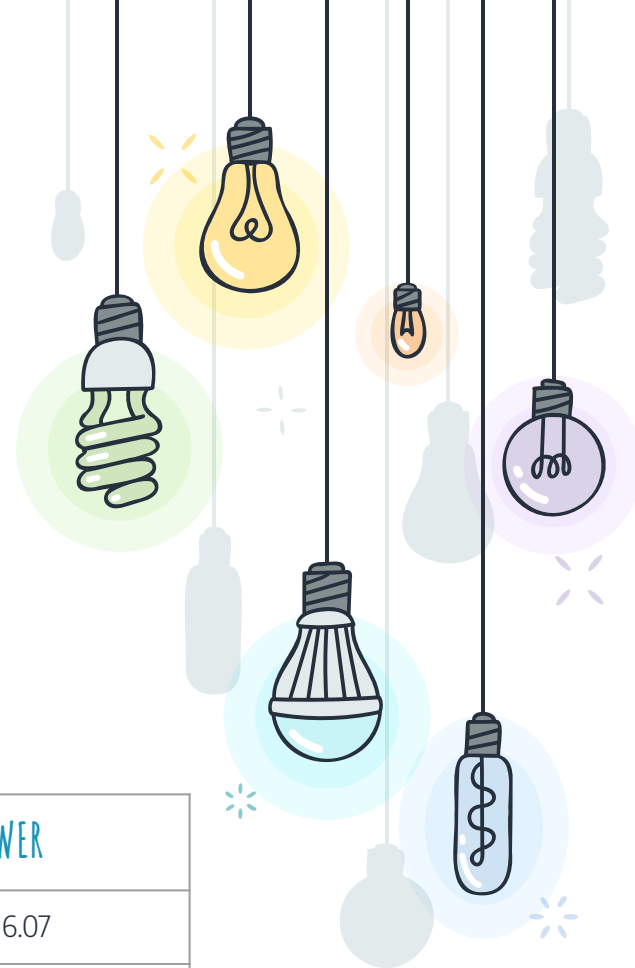
Zestawienie wartości WER  
dla dobrej i złej dykcji lektora (set 2)



Dykcja

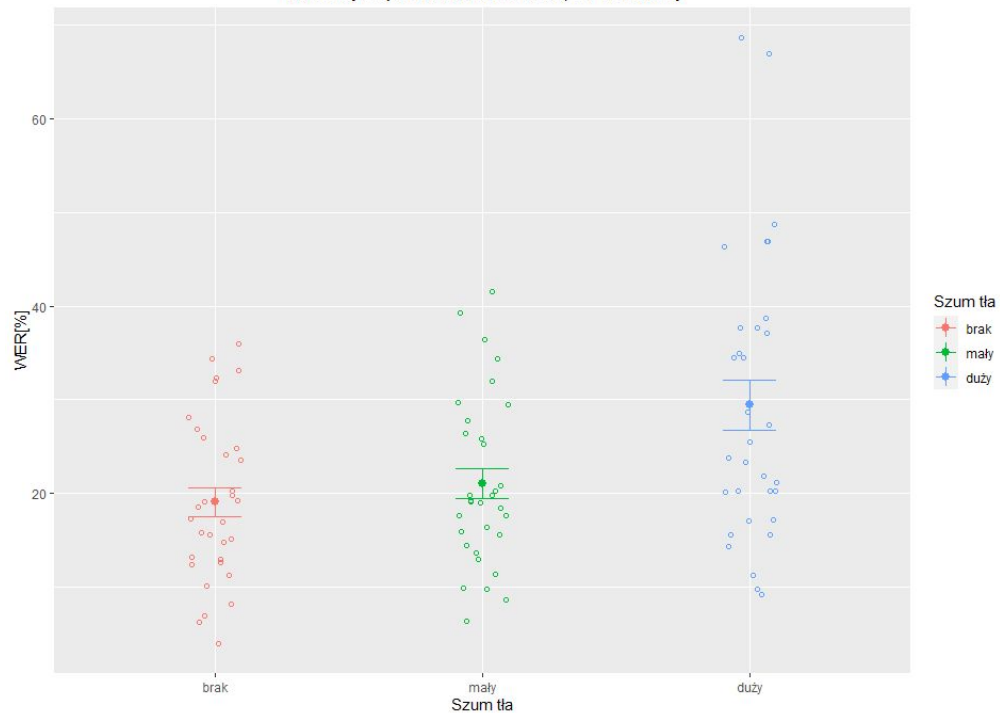
- dobra
- przeciętna

DYKCJA	WER
dobra	16.07
przeciętna	21.59

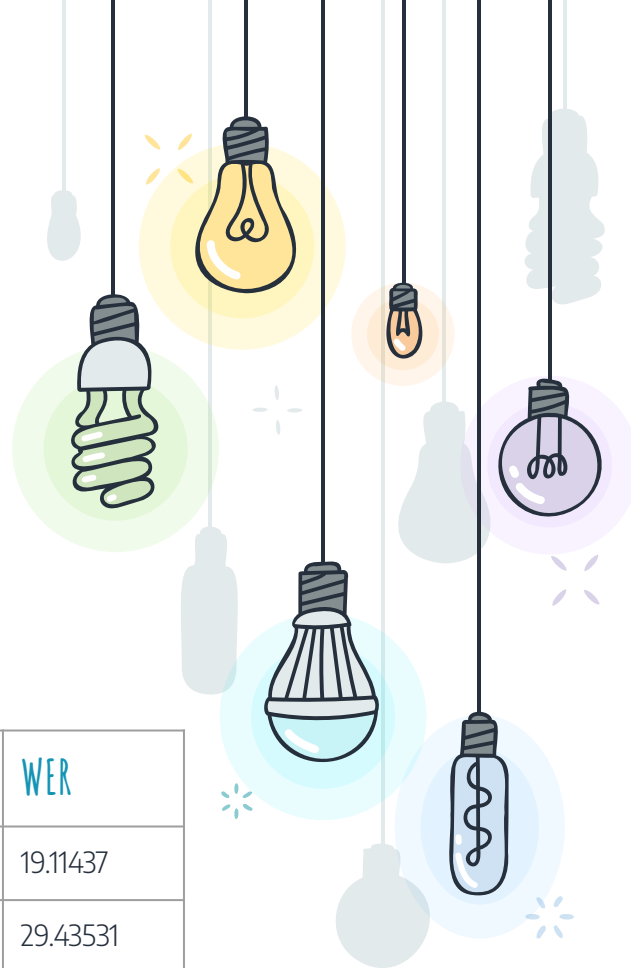


# \* SZUM

Zestawienie wartości WER  
dla różnych poziomów szumu tła (set 1, lektorzy)

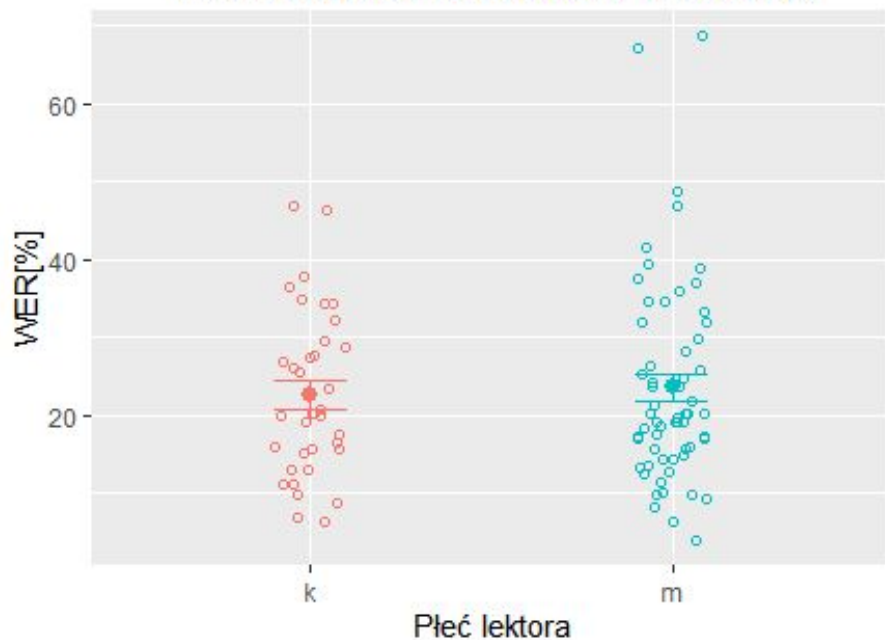


SZUM	WER
brak	19.11437
duży	29.43531
mały	21.09156



# \* PŁEĆ

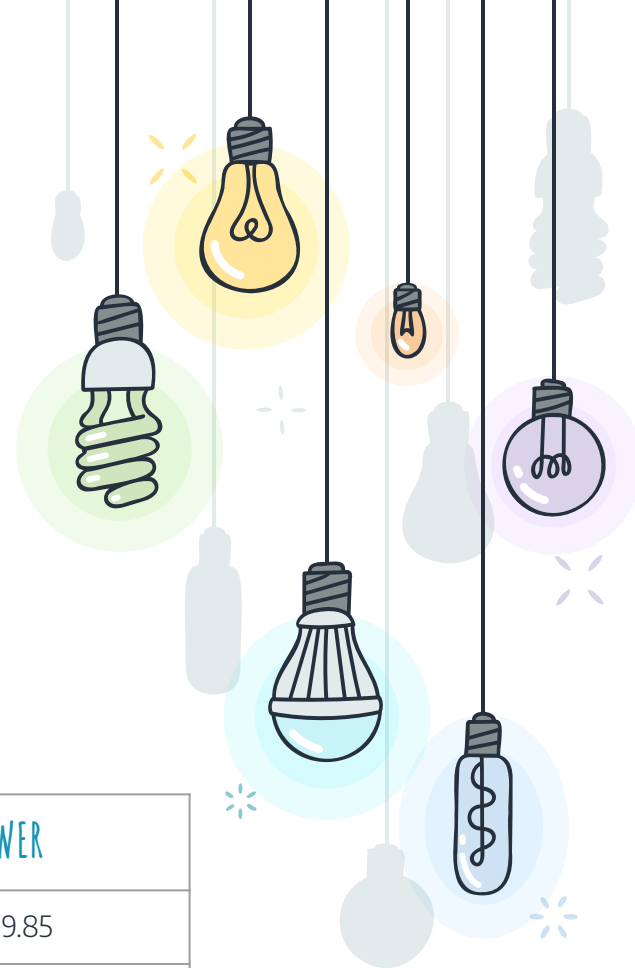
Zestawienie wartości WER  
dla różnej płci lektorów (set 1, lektorzy)



Płeć lektora

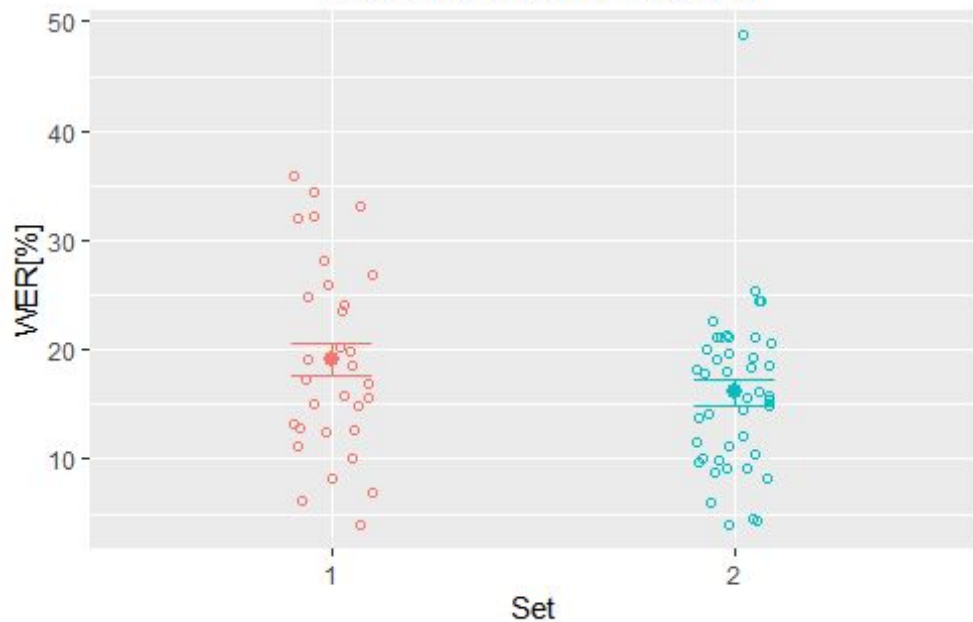


PŁEĆ	WER
kobieta	19.85
mężczyzna	20.25



# \* OGÓLNE PORÓWNANIE SETÓW

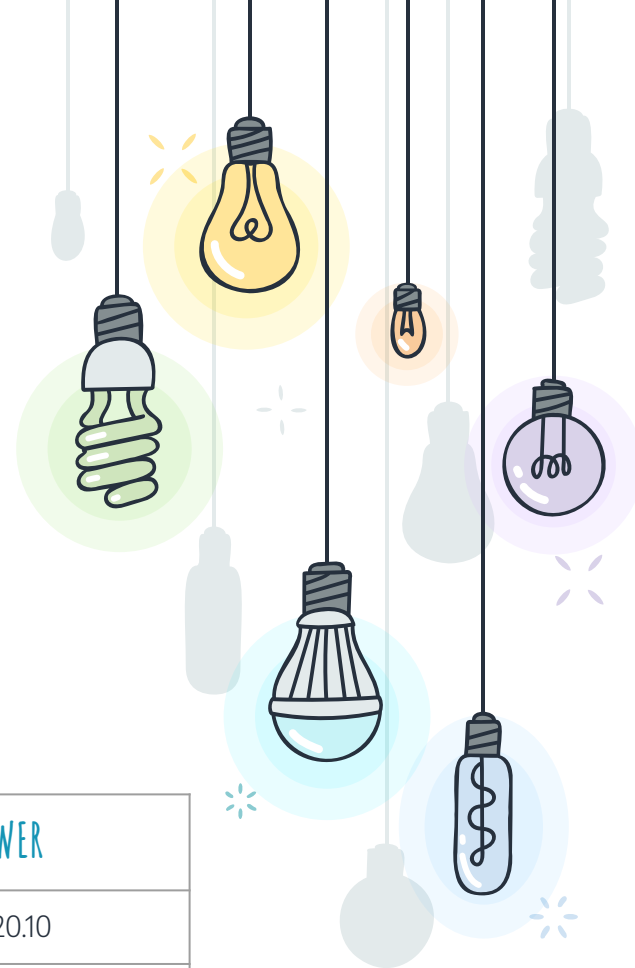
Zestawienie wartości WER  
dla różnych typów nagrań



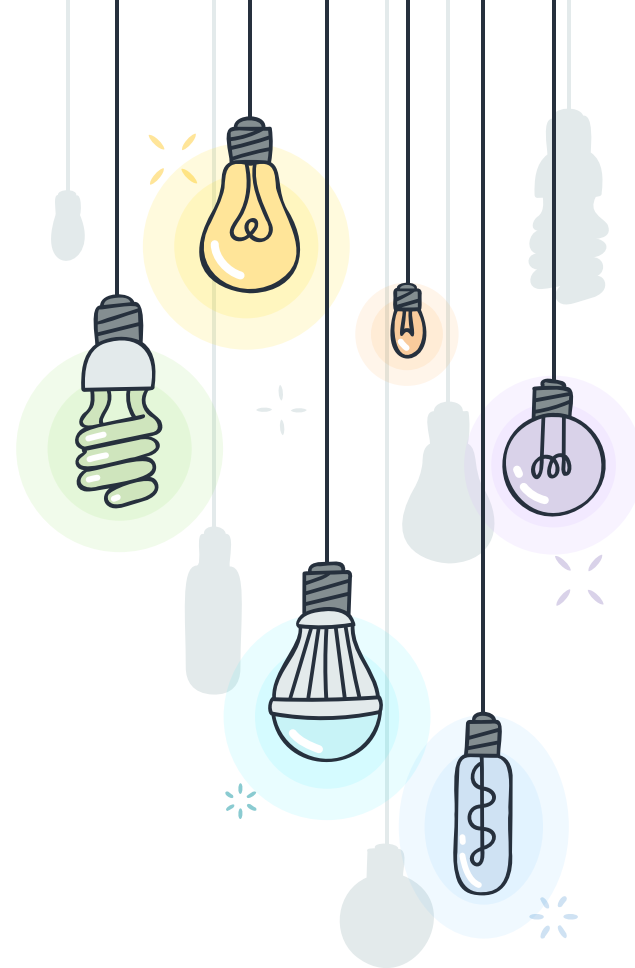
Set

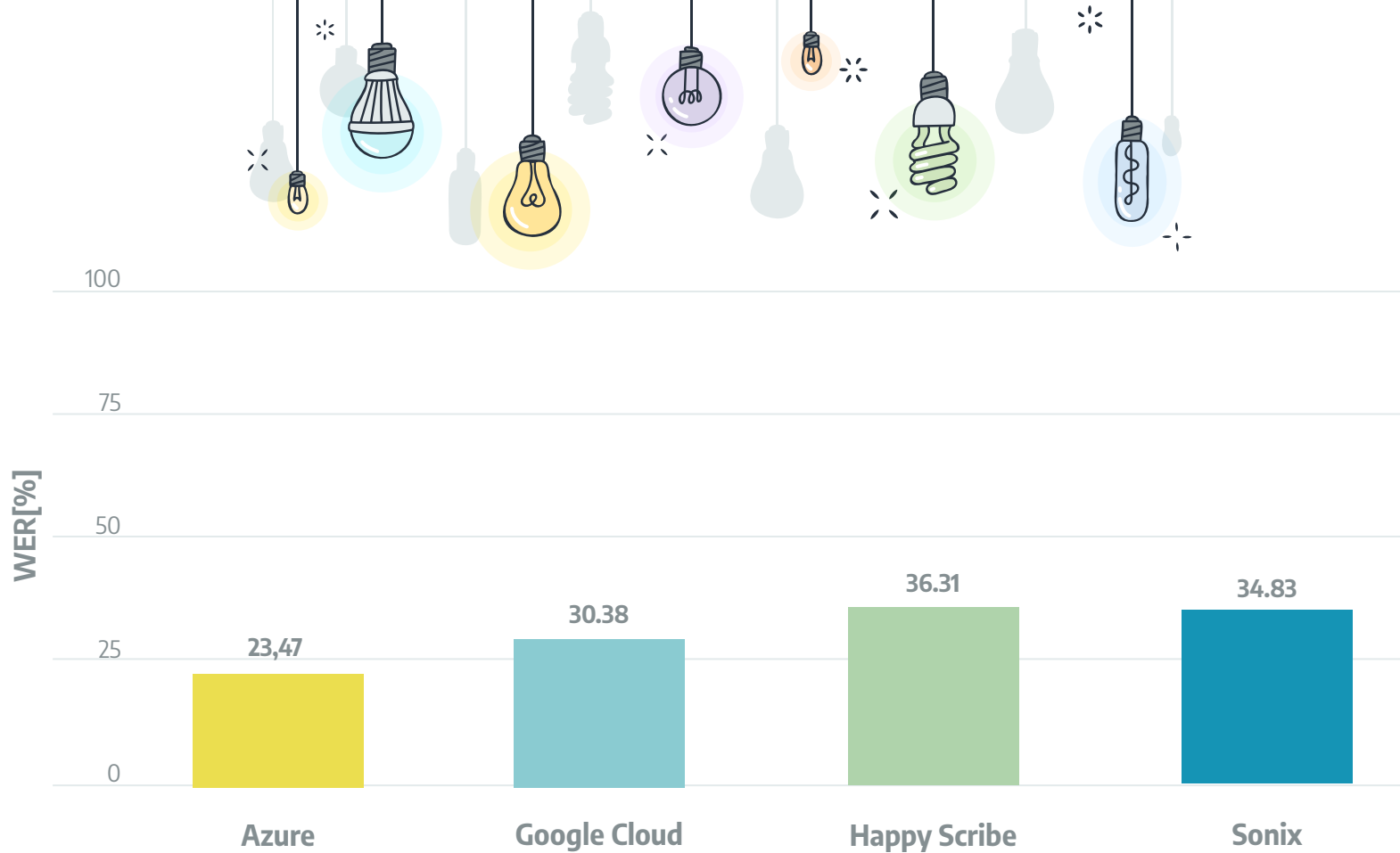


Nr SETU	WER
Set 1	20.10
Set 2	17.88



# \* WPŁYW WYBRANYCH CZYNNIKÓW





# 6

## PROPOZYCJE PRZYSZŁYCH BADAŃ

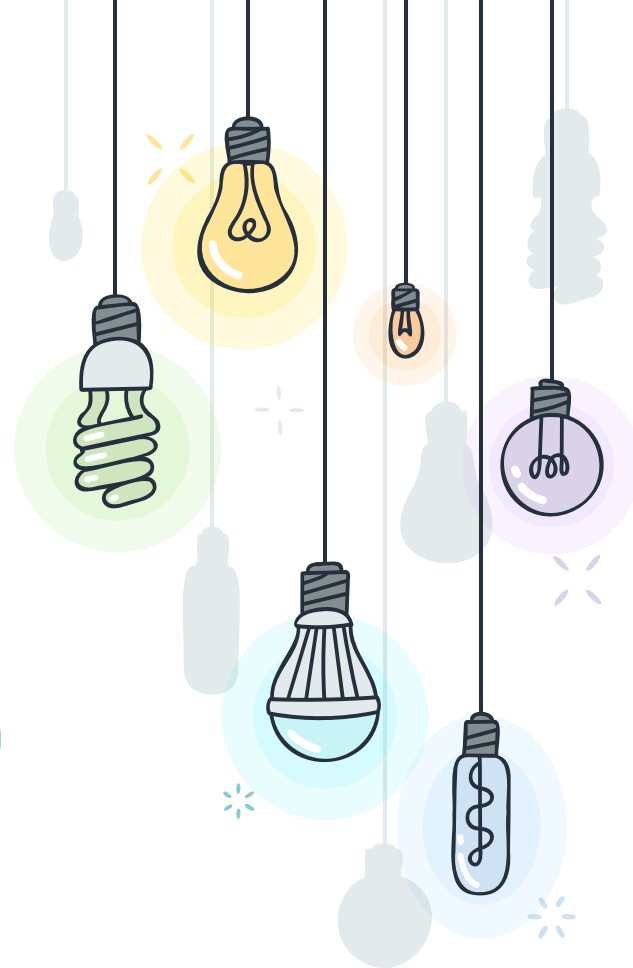
Czyli, jak wykorzystać nasze testy do dalszego rozwoju





## \* POMYSŁY

- + Chmura słów
- + Wpływ szybkości mowy
- + Zestawienie słów specjalistycznych, staropolskich etc.
- + Powiększenie data setu





# DZIĘKUJEMY ZA UWAGĘ

Karolina Głowińska, Julia Kułacz, Martyna Pawlus