

Automatic pronunciation grading for Dutch

Catia Cucchiarini & Helmer Strik

A²RT, Dept. of Language and Speech
University of Nijmegen

Contents:

- assessment of pronunciation
read speech (4 scales)
- assessment of fluency
read speech
extemporaneous speech

assessment of pronunciation

Goal

To determine whether expert pronunciation ratings can be predicted on the basis of automatically calculated measures

Method

80 speakers: 60 NNS, 16 NS & 4 SDS

2 sets of 5 phonetically rich sentences
read speech
orthographically transcribed

CSR: 38 monophones & lexicon
Viterbi alignment of speech signals &
orthographic transcriptions

segmentation on phone level

various automatic measures:

tdur1, tdur2,
art, ros, ptr, mlr, #p, tdp, alp

Human ratings

3 groups of 3 experts:

1. Phon : Phoneticians
2. ST1 : Speech Therapists 1
3. ST2 : Speech Therapists 2

scored the 10 sentences of all 80 subjects on 4 pronunciation scales:

- no specific instructions
- 80 speakers were divided over the 3 raters in a group

range

session 1

1 OP : overall pronunciation [1, 10]

session 2

2 SQ : segmental quality [1, 10]

3 FL : fluency [1, 10]

4 SR : speech rate [-5, 5]

Results

Reliability:

intrarater: 0.76 - 0.98

interrater: 0.76 - 0.97

Agreement:

means and standard deviations varied

- between the raters in a group
- between the raters in different groups

→ normalization:

standard scores / Z scores

Correlations between
autom. measures and
raw human scores

		OP	SQ	FL	SR
tdur2	Phon	-.73	-.68	-.90	-.82
	ST1	-.78	-.77	-.97	-.86
	ST2	-.72	-.65	-.86	-.85
ptr	Phon	.69	.64	.83	.75
	ST1	.76	.74	.92	.75
	ST2	.70	.68	.85	.78
ros	Phon	.76	.72	.92	.83
	ST1	.80	.79	.93	.87
	ST2	.75	.70	.85	.85

Correlations between
autom. measures and
normalized human scores

		OP	SQ	FL	SR
tdur2	Phon	-.79	-.75	-.91	-.90
	ST1	-.81	-.77	-.94	-.88
	ST2	-.73	-.70	-.91	-.88
ptr	Phon	.76	.73	.86	.86
	ST1	.78	.74	.88	.78
	ST2	.72	.72	.89	.80
ros	Phon	.82	.79	.93	.92
	ST1	.83	.79	.91	.89
	ST2	.77	.76	.90	.89

Correlations between the 4 scales

		OP	SQ	FL	SR
OP	Phon		.97	.87	.73
	ST1		.96	.87	.60
	ST2		.91	.77	.64
SQ	Phon			.86	.69
	ST1			.91	.61
	ST2			.76	.62
FL	Phon				.87
	ST1				.83
	ST2				.83

Means and standard deviations
for read and spontaneous speech

	read speech				spont. speech	
	NS		NNS		LP-HP	
	_	sd	_	sd	_	sd
ros	12.74	1.35	9.68	1.94	5.65	1.12
ptr	93.17	2.79	82.66	8.57	47.09	9.32
art	13.65	1.19	11.61	1.37	12.04	1.06
#p	1.42	1.23	7.20	5.47	65.91	34.40
tdp	0.45	0.42	3.10	2.76	63.50	36.91
alp	0.20	0.13	0.38	0.13	0.97	0.25
mlr	34.26	5.85	21.52	8.77	9.41	2.23

Correlations between the raw human scores

	OP	SQ	FI	SR
ph-st1	.92	.90	.94	.90
ph-st2	.80	.57	.82	.88
st1-st2	.90	.69	.83	.81

Correlations between the normalized human scores

	OP	SQ	FI	SR
ph-st1	.96	.91	.94	.93
ph-st2	.90	.87	.90	.86
st1-st2	.94	.84	.90	.89

Correlations between 7 quantitative variables

	ros	ptr	art	#p	tdp	alp	mlr
ros		.91	.96	-.87	-.86	-.71	.88
ptr			.75	-.97	-.96	-.73	.94
art				-.72	-.71	-.61	.74
# p					.97	.63	-.91
tdp						.67	-.86
alp							-.76