

Index of the Ph.D Dissertation:
Confidence Measures for Automatic and Interactive Speech Recognition

by Isaias Sanchez Cortina,
advisors: Dr. Alfons Juan i Ciscar Dr. J. Alberto Sanchis Navarro

~

1 Introduction	13
1.1 Motivation and Thesis Outline	13
1.2 Document Structure	15
1.3 Scientific and Technologic Goals	16
2 Background	17
2.1 Automatic Speech Recognition	17
2.1.1 Feature Extraction: Audio Signal Analysis	18
2.1.2 Statistical approach to Automatic Speech Recognition	20
2.2 Confidence Measures	31
2.2.1 Combination of predictors	35
2.2.2 Utterance Verification as Hypothesis testing	36
2.2.3 Comparison of the different CM approaches	38
2.3 Interactive Speech Transcription	40
3 Interactive Speech Transcription	43
3.1 Introduction	43
3.2 Interactive speech transcription balancing error and supervision effort	45
3.2.1 Interactive Speech Transcription Procedure	45
3.2.2 Estimation of the Transcription Error	47
3.3 Experimentation	49
3.3.1 Experimental Setup	49
3.3.2 Results	52
3.4 Conclusions	56
3.4.1 General remarks on the behaviour of the IST method	56
3.4.2 Remarks on the large task WSJ1+0	57
4 IST prototype	59
4.1 IST prototype balancing error and user effort	59
4.2 User interface of the prototype	61
4.3 Demonstration	63
4.4 Usability	64
4.5 Conclusions	65

5 transLectures: Transcription of massive repositories of lectures	67
5.1 Introduction	67
5.2 Deployed architecture for cost-effective transcription of online repositories	68
5.2.1 the transLectures Player	71
5.2.2 ASR & SMT Systems	72
5.3 Overview of the videoLectures.net and poliMedia repositories:	73
5.4 First results on poliMedia	75
5.5 Conclusions	78
6 Intelligent interaction with limited user effort	79
6.1 Introduction	79
6.2 IILUE approach	81
6.3 Constrained search for speech recognition	83
6.4 Evaluation	85
6.4.1 Experimental setup	85
6.4.2 Results	88
6.5 Conclusions	92
7 Improved and Speaker-adapted confidence measures	95
7.1 Introduction	96
7.2 Speaker-adapted naïve Bayes classifier	98
7.2.1 Estimation of the smoothing parameters	101
7.3 Speaker-adapted logistic regression confidence estimator	103
7.3.1 Discriminative learning	106
7.4 Experiments	108
7.4.1 Experimental setup	108
7.4.2 Evaluation of CMs	110
7.4.3 Results	112
7.5 Interactive Speech Transcription Application	117
7.6 Conclusions	121
8 Conclusions	123
8.1 Summary	123
8.2 Publications	125
8.3 Future work	126