

UNIVERSITAT POLITÈCNICA DE VALÈNCIA
DEPARTAMENT DE SISTEMES INFORMÀTICS I COMPUTACIÓ



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Different Contributions to Cost-Effective Transcription and Translation of Video Lectures

Thesis

presented by Joan Albert Silvestre Cerdà

supervised by Dr. Alfons Juan Císcar and Dr. Jorge Civera Saiz

November 18, 2015

CONTENTS

Agraïments	vii
Abstract	ix
Resumen	xi
Resum	xiii
Contents	xv
1 Introduction	1
1.1 Motivation	2
1.2 Scientific and Technological Goals	4
1.3 Document Structure	5
Bibliography	7
2 Preliminaries	9
2.1 Automatic Speech Recognition	10
2.2 Audio Segmentation	10
2.3 Machine Translation	11
2.4 Language Modelling	12
2.5 Recommender Systems	13
2.6 Evaluation Metrics	14
2.6.1 Segmentation Error Rate	14
2.6.2 Word Error Rate	14
2.6.3 Bilingual Evaluation Understudy	14
2.6.4 Translation Error Rate	15
2.6.5 Real Time Factor	15
Bibliography	17

3	Explicit Length Modelling for Statistical Machine Translation	19
3.1	Introduction	20
3.2	Implicit Length Modelling	21
3.3	Explicit length modelling	21
3.3.1	Standard length models	22
3.3.2	Specific length models	23
3.3.3	Estimation of phrase-length models	24
3.4	Experimental results	26
3.5	Conclusions	34
	Bibliography	37
4	Efficient Audio Segmentation for Speech Detection	39
4.1	Introduction	40
4.2	Corpora	40
4.3	System description	41
4.4	Experimental results	42
4.5	Albayzin 2012 Evaluation	42
4.6	Conclusions	43
	Bibliography	45
5	The transLectures-UPV Platform	47
5.1	Introduction	48
5.2	poliMedia	49
5.3	transLectures	51
5.4	System Architecture	53
5.4.1	Web Service	54
5.4.2	Player	55
5.4.3	Database	56
5.4.4	ASR & SMT Systems	57
5.5	System Evaluation	61
5.5.1	Automatic Evaluations	61
5.5.2	User Evaluations	65
5.6	Conclusions	70
	Bibliography	71
6	Recommender Systems for Online Learning Platforms	75
6.1	Introduction	76
6.2	Recommendation system overview	76
6.3	System Updates and Optimisation	79
6.4	Integration into VideoLectures.NET	80
6.4.1	The LaVie project	80
6.4.2	The VideoLectures.NET Repository	81
6.4.3	VideoLectures.NET user-lecture interaction analysis	81
6.4.4	Topic and User Modelling	86
6.4.5	Learning Recommendation Feature Weights	87

6.4.6	Evaluation	87
6.5	Conclusions	89
	Bibliography	91
7	Language Model Adaptation Using External Resources for Speech Recognition	93
7.1	Introduction	94
7.2	Document Retrieval	95
7.3	Language Model Adaptation	95
7.4	Experiments	96
7.4.1	Corpora	97
7.4.2	Acoustic Models	97
7.4.3	Language Models	98
7.4.4	Evaluation	99
7.5	Conclusions	99
	Bibliography	101
8	Transcription and Translation Platform	103
8.1	Introduction	104
8.2	The transLectures-UPV Platform	104
8.2.1	Use Cases	106
8.2.2	Database	110
8.2.3	Web Service	111
8.2.4	Player	112
8.2.5	Ingest Service	113
8.3	Integration with poliMedia	117
8.3.1	ASR and SMT Systems	117
8.3.2	Automatic Evaluations	119
8.3.3	User Evaluations	126
8.4	MLLP's Transcription and Translation Platform	128
8.5	Conclusions	130
	Bibliography	133
9	Conclusions	135
9.1	Summary	136
9.2	Publications	137
9.3	Future Work	139
	Bibliography	141
	List of Figures	143
	List of Tables	147