

Digital Deployment of the Signs of Ireland Corpus in Elearning

Lorraine Leeson, Brian Nolan

Centre for Deaf Studies, School of Linguistic, Speech and Communication Sciences,
Trinity College Dublin, Dublin, Ireland.

Department of Informatics, Institute of Technology, Blanchardstown, Dublin, Ireland.

E-mail: leesonl@tcd.ie brian.nolan@itb.ie

The Signs of Ireland Corpus



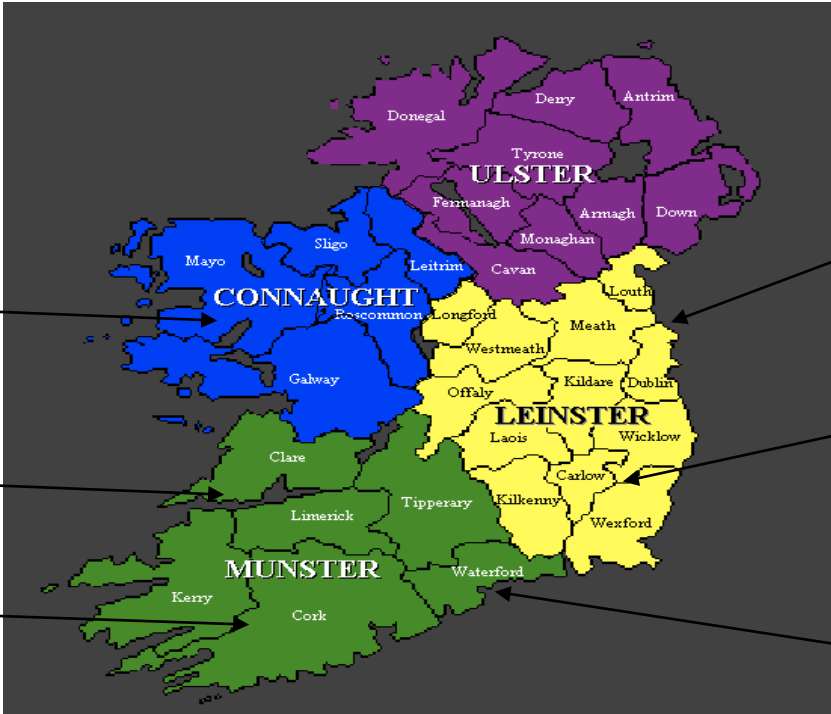
- **Aim:** to develop a multi-functional corpus (research and teaching);
- 40 (16=m, 24=f) Deaf ISL signers from across the republic (2004);
- All educated in the schools for the Deaf in Dublin;
- Male and female, non ISL teachers, range of age groups: aim to capture potential sociolinguistic variation;
- 5 Locations: Dublin, Cork, Galway, Waterford, Wexford (in 2004)
- Limerick data to be collected in 2008*
- Currently the largest annotated corpora of a SL in Europe.
- **Crucially:** permission given to use data for research & teaching, including publication.

Ireland

GALWAY

LIMERICK

CORK



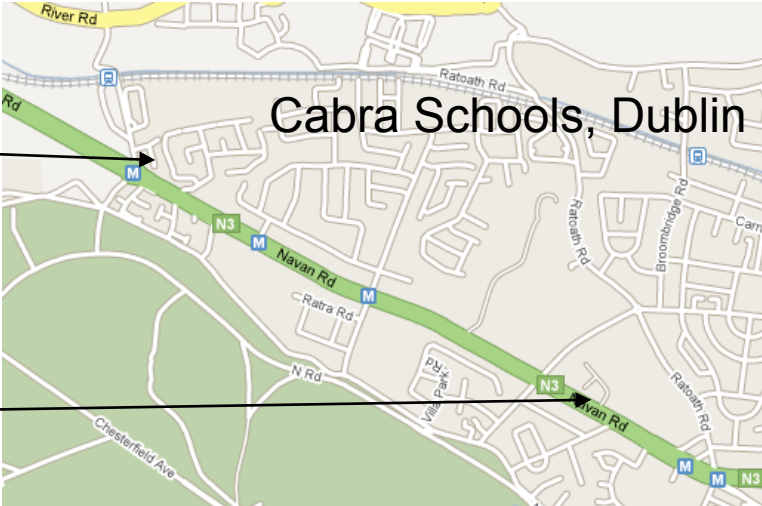
DUBLIN

WEXFORD

WATERFORD

St. Mary's

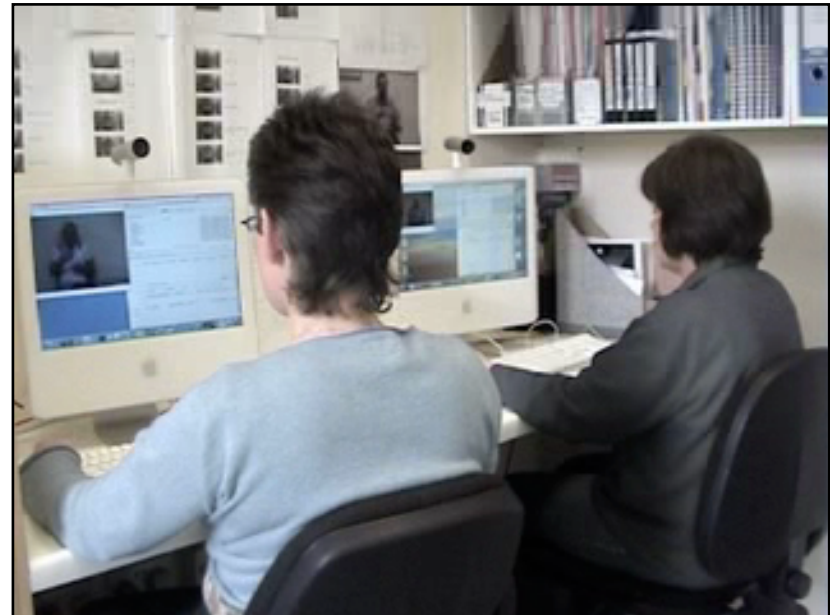
St. Joseph's



Cabra Schools, Dublin

Transcription

- Annotated using ELAN (EUDICO Linguistic Annotator) from 2005-7;
- 40 self-selected narratives + Volterra data task (transitivity test) + ‘Frog Story’ - picture sequence data;
- No writing system for ISL = challenge for codification;
- Used ECHO Project notations as basis. Modified additionally for ISL data.





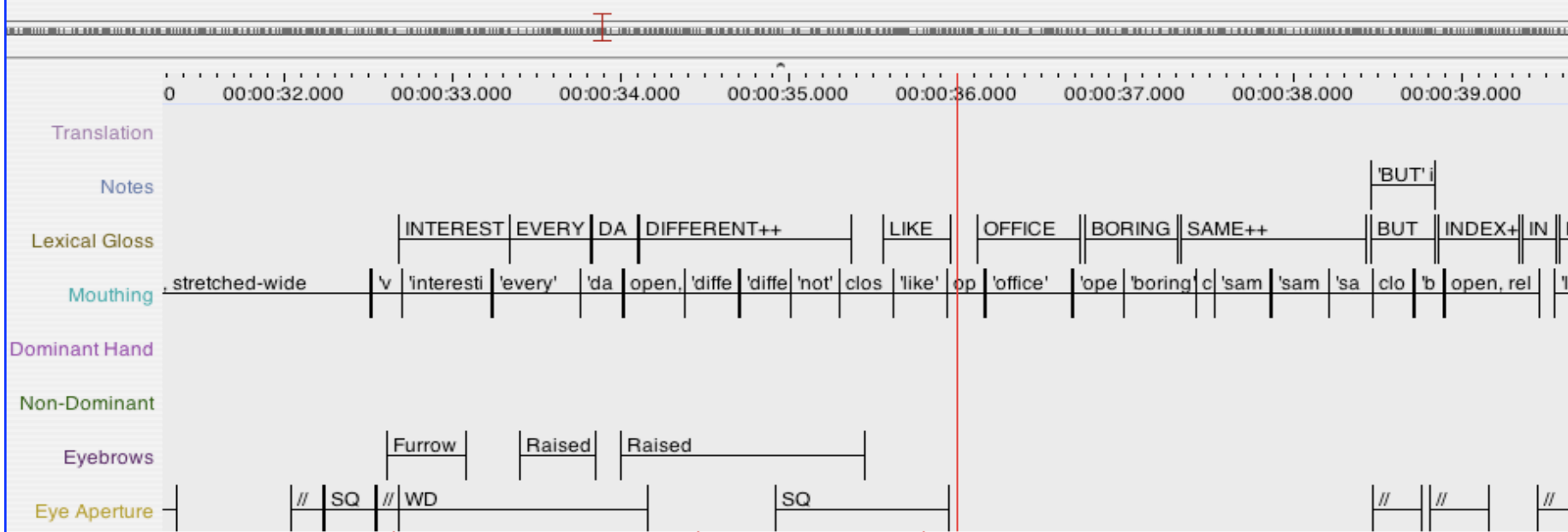
Grid Text Subtitles Controls

Lexical Gloss

Nr	Annotation	Begin Time	End Time	Duration
45	AND	00:00:29.980	00:00:30.840	00:00:00.860
46	INTEREST	00:00:32.680	00:00:33.340	00:00:00.660
47	EVERY	00:00:33.340	00:00:33.820	00:00:00.480
48	DAY	00:00:33.830	00:00:34.100	00:00:00.270
49	DIFFERENT++	00:00:34.110	00:00:35.370	00:00:01.260
50	LIKE	00:00:35.560	00:00:35.960	00:00:00.400
51	OFFICE	00:00:36.127	00:00:36.737	00:00:00.610
52	BORING	00:00:36.767	00:00:37.317	00:00:00.550
53	SAME++	00:00:37.337	00:00:38.437	00:00:01.100
54	BUT	00:00:38.467	00:00:38.847	00:00:00.380
55	INDEX+fl	00:00:38.867	00:00:39.347	00:00:00.480

00:00:36.011 Selection: 00:00:00.000 - 00:00:00.000 0

Selection Mode Loop Mode



Codification & Transcription

- All data verified by Deaf research assistant
- Problems of glossing:
“... glosses impose upon the data a wealth of unwarranted and highly variable lexical and grammatical information (depending upon the spoken/written language used for glossing).” (Pizzutto and Pietrandrea 2001: 42).

- (1) Sean: HEARING
- (2) Fergus M: HEARING - both are glossed the same so that searches show variants



You Get Out What You Put In!



- Search function is restrained by the consistency and accuracy of the annotations inputted
- Variants often glossed as the same thing (e.g. WHAT (1), (2)).
- Tagging re: grammatical function not yet attempted- though we intend to begin to tag certain items (e.g. verbs, mouthings) in 2008.

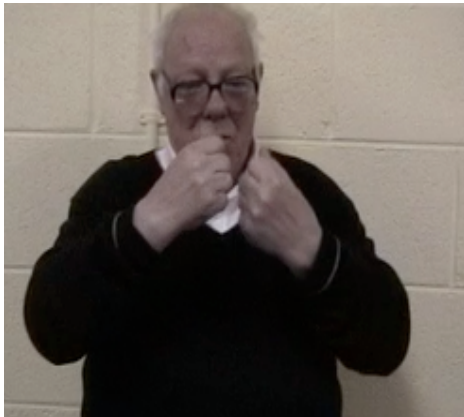
Annotating the SOI Texts

- Decisions had to be made re:
 - non-manual features as dependent tiers, relative to the manual signs that they co-occur with, or as independent tiers containing information that may be supra-segmental in nature.
 - what constitutes a word in ISL (word vs. gesture vs. complex predicate)
 - Borrowings, neologisms, gendered signs
 - Mouth patterns
 - The multi-functionality of the corpus guided our decisions.



Next steps with the corpus

1



2



- Use in classroom
- Deployment in elearning
- Analysis using ELAN
 - Phonology
 - e.g. symmetry constraint as in (1)
 - Morphology
 - e.g. role of mouthing, NMFs (e.g. squint, as in (2))
 - Syntax / Semantics /
 - Discourse structure
 - Humour
 - Developing ontologies for machine translation/
linguistic modelling (with ITB)

Using the Corpus

- Multi-functional: In teaching ISL, interpreting/ translation, sign linguistics
- For linguistic research purposes:
 - A corpus led description of ISL (Leeson & Saeed, in prep.)
 - PhD on mouthing (Fitzgerald, in prep.),
 - Possible machine translation projects (with ITB)
- For sociolinguistic description:
 - Identifying the extent and form of variation
 - For capturing the contemporary usage of ISL among ‘naïve’ signers aged 18-65.
 - For exploring the possibility of regional variation in ISL
 - For considering collocational patterns in ISL
- Cross-referencing of co-occurrence of data on several layers possible
 - (e.g. mouthings and manual signs, eyeblink and offset of clause, NMFs and classifier constructions, etc.)

What we need most:

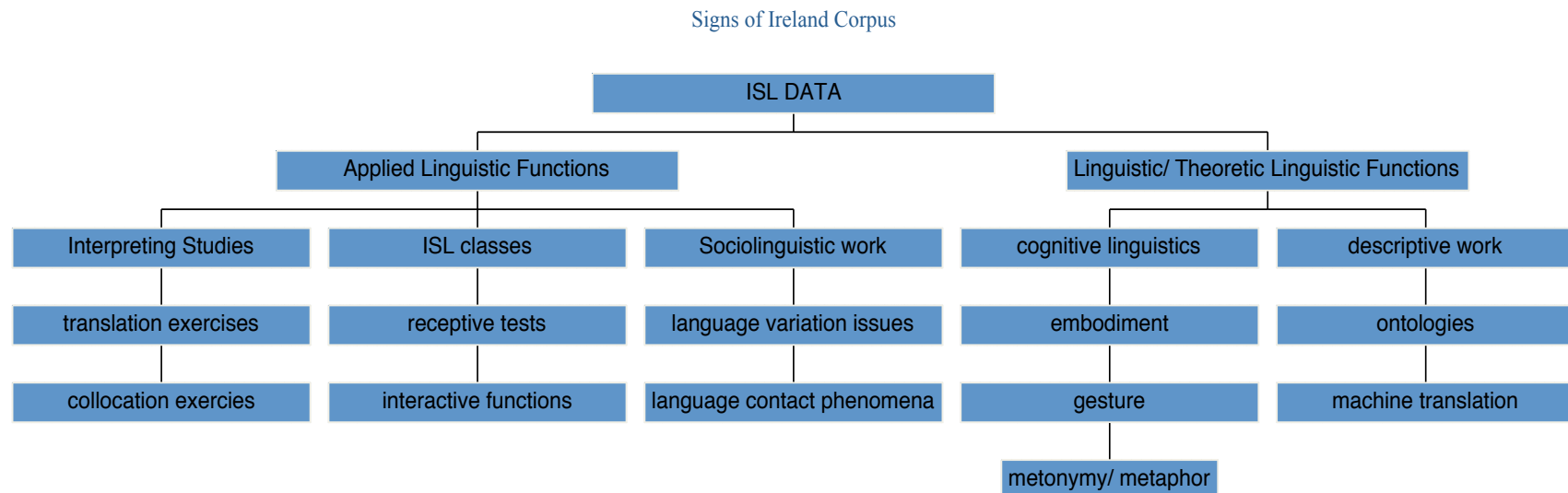


Fortunately...along came SIF II!

- In late February 2008, the Minister for Education and Science announced funding decisions regarding the Strategic Innovation Fund (SIF II).
- CDS, in partnership with the Institute of Technology Blanchardstown , were awarded c.2.8 million euro.
 - collaboration partners in deployment include University College Cork, Limerick Institute of Technology and the Tipperary Institute.
- Also working on a Leonardo da Vinci Project, SIGNALL II, promoted by Interresource Group (Ire.) Ltd. to deliver a “Perspectives on Deafness” course with content from Polish, Czech, UK and Finnish partners.
- **Aims** - to develop blended learning delivery of Diploma programmes nationally and to develop a 4 year degree in Deaf Studies

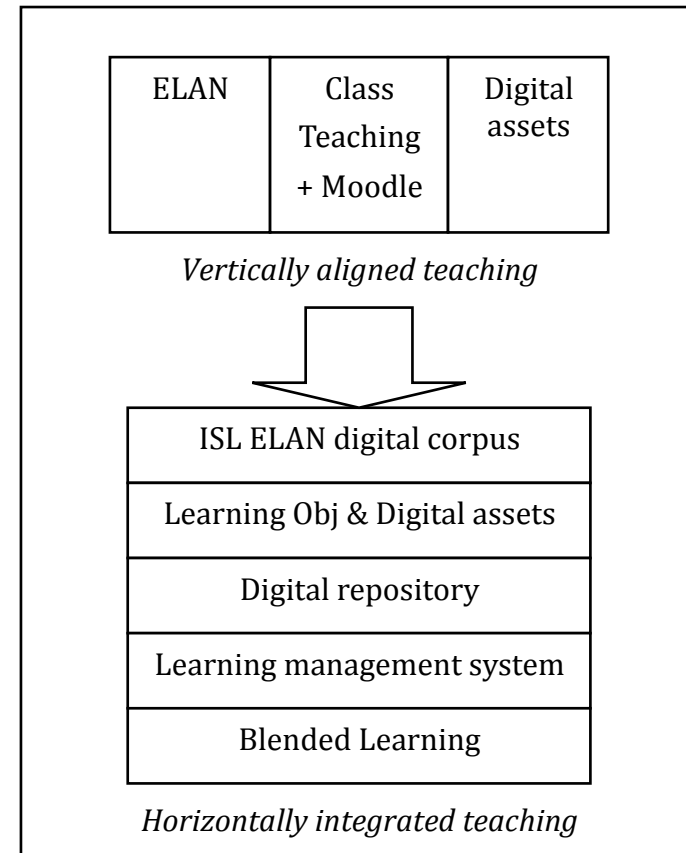
How does this link to the corpus?

- Our project entails the development of digital learning objects as a pedagogical tool
- The Signs of Ireland corpus clearly links with this objective
- The enlargement of the corpus to include the “gapped” regional data as well as elderly signer and child signer data can be managed within this framework
- Annotation may not be as detailed moving forward, but will be required.
 - Some of this work may be covered by postgraduate students at CDS / ITB.



Project requirements

- Leverage the ISL digital corpus within a blended learning environment with digital learning objects
 - At the moment we have Moodle populated with a wide variety of modules within the programme.
 - We have the ISL digital corpus tagged in ELAN.
 - Currently, we have traditional classroom delivery.
 - The programme architecture is very vertical in orientation.
- Achieve horizontal integration through the use of information technology, the Internet and a blended learning approach.
- Getting the programme complete and online



Some key questions

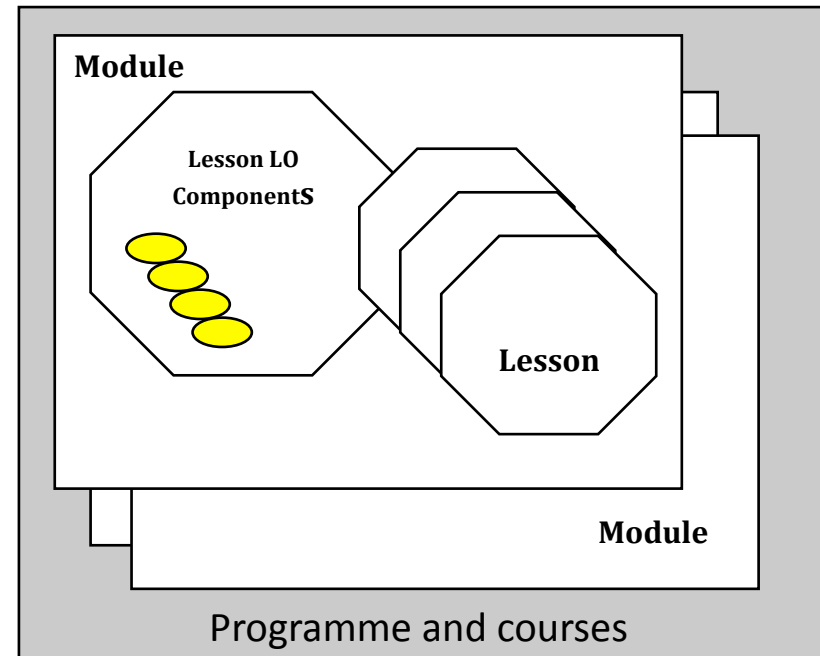
- What is an appropriate **architecture for a digital repository** to support sign language learning
- What is an appropriate assessment model for our online blended learning environment
-
- what **types of learning objects** do we need for
 - each of the programme modules
 - each lecture,
 - and how many?



Planning ahead

We need to plan for the capture and creation of the respective digital rich media that we intend to deploy.

- Link the programme learning outcomes:
 - to **module** learning outcomes
 - to **lecture** learning outcomes
 - & **identify learning objects** to support all these learning outcomes



Analysis phase of project

Module	Lecture	Themes	Learning Outcomes	Learning Objects	Assessment
Sign Linguistics & Sociolinguistics	1	Introduction to Sign Linguistics			
	2	Are Sign Languages Real Languages?	<p>The student will be able to:</p> <ol style="list-style-type: none"> 1. explain - what is a language 2. explain what are the characteristic features of language 3. argue whether these hold for Signed Languages 4. provide evidence for Signed Languages as real languages. 	<p>LO1.1 Lecture notes LO1.2 PPT/PDF LO1.3 online pop quiz LO1.4 Video data (c. 10mins interview)</p> <p>LO2.1 LO2.2 LO2.3 LO2.4</p> <p>LO3.1 LO3.2 LO3.3 LO3.4</p> <p>LO4.1 LO4.2 LO4.3 LO4.4</p>	<p>End of year essay question on this theme.</p> <p>Pop quiz</p>

Tagging of the digital assets

1. Topic
2. Description
3. Sections
4. Media
 - a. Source
 - b. Options for reuse
 - c. Context - 'where used now'
 - d. Proof of availability
 - e. Ownership
 - f. Licensing
 - g. Cost
 - h. Payment Method
 - i. Optimum speed of access and use
 - j. Ability to apply style guide
 - k. Types supported
5. Handle tags: Specific topics covered
6. Context
 - a. Modality for delivery
 - b. Format
7. Conversion speed
8. Assessment of topics
 - a. Assessment of specific areas
 - b. Depth of assessment
 - c. Level of adaptability
 - d. Feedback
9. Author
10. Version number
11. Date Created

People

- We will recruit
 - a project manager who is bilingual in ISL/English.
 - academic staff for local delivery of ISL in the regions,
 - interpreting lecturers and general Deaf Studies academics
 - elearning/ digital repository/ digital media specialist.
 - full-time ISL interpreter and administrative support to the project.
- To contribute to the research of the programme, we will engage a number of people on Ph.D level to investigate over the three years in the following research areas:
 1. Assessment models appropriate to ISL in an elearning and blended learning context;
 2. Developing and maturing the ISL ELAN digital corpus, including meta-tagging and enriching the data;
 3. Signed languages interpreting;
 4. Design and build of rich digital media for Irish Sign Language

Go raibh maith agaibh! ‘thank you!’

Any questions?

Brian Nolan & Lorraine Leeson



Thanks to:

- All SOI participants
- Deirdre Byrne-Dunne, Alison Macduff and Cormac Leonard
- Profs. David Little & Jeffrey Kallen, SLSCS
- Arts & Social Sciences Benefaction Fund (2003-7, 2006-7, 2007-8);
- Trinity College Start Up Research Funds (2001-6)
- University of Dublin Emergency Research Funds (2003/4)
- Strategic Innovation Funds, Cycle II (HEA, 2008-11)
- SIGNALL II (Leonardo da Vinci, European Commission, 2007-9)