Case for support: Translating science for young people

1. Fit to the call and contribution to the theme
The proposed research fits the "Translating Cultures Innovation Grants" call strategic question (2) "How can research into languages and cultures [...] extend analysis of the ways in which translation serves as a form of transmission and circulation of ideas, ideologies and forms of knowledge between geographical locations, historical moments and cultural contexts?"

The primary aim of the project is to investigate the translation and transformation of current scientific knowledge in communications for and with young people aged 11-16 (Key Stages 3 and 4 in English schools). Understanding what happens when scientific knowledge and ideas are translated is important because scientific issues are central to young people's future lives as active citizens (Schreiner et al, 2008), yet texts produced by scientists for their peers are not generally accessible to non-specialists. Translation here is understood as the interpretation and transformation of ideas between and across related genres in the same language, in this case, between experts in science, educators and other writers, and young people. We are also concerned with issues of what is not translated, or is apparently mistranslated, and how young people's talk about their understandings and attitudes towards science may reflect this translation, non-translation and mistranslation.

The secondary aim of the project is to develop understandings of how linguistic metaphorical and metonymical framing devices, and related linguistic mechanisms, are used and translated in a real-life context to explain, simplify and encode information, agency and evaluative stance.

Language choices and use are at the core of translation, whether inter-lingual, or as in our proposed research, intra-lingual, but most research in the communication of science to the public and in science education for young people uses informant interviews and content analysis as its central methods. This project is innovative in that it brings detailed linguistic expertise and techniques to analysing a large body of relevant texts. We have chosen to focus on a specific scientific topic, climate change, to ensure some parity of subject matter across the corpora studied. The topic chosen is a Socioscientific issue (Morris, 2014) that has a considerable body of research literature, is dealt with in science and other curriculum areas in schools, has implications for lifestyle and society more widely, and potentially major consequences during the lifetimes of current school students.

The project will extend the range of corpus linguistic research techniques to this important societal issue. The PI and one Co-I (2013) have established international reputations in both the quantitative and qualitative analysis of linguistic data, including in science communication (e.g. 2013; 2011). Together and separately they have analysed variations in figurative language use according to genre and register, and the metaphors of expert and popular science writing. The project also takes forward the second Co-I (2013) research into the learning of science by secondary school students, and their views of the role of science in their lives. The project builds on both these bodies of research and also represents a new collaborative venture in bridging the two fields.

2. Research questions
We will address the following questions:

Overarching questions:
How is specialist scientific knowledge translated, not translated or apparently mistranslated in texts accessed by young people?
How do young people account for their understanding and attitudes towards scientific knowledge about climate change?

Specific questions:
1. What are the predominant metaphorical and metonymical frames and related linguistic devices used in the language produced by experts for their peers or knowledgeable counterparts in texts dealing with climate change?
   Which aspects of each topic are highlighted by these frames and devices?
   What evaluative meanings are presented by these frames and devices?

2. What are the predominant metaphorical and metonymical frames and related linguistic devices used in popular and educational texts aimed for and used by young people dealing with climate change?
   Which aspects of each topic are highlighted by these frames and devices?
   What evaluative meanings are presented by these frames and devices?
3. What are the predominant metaphorical and metonymical frames and related linguistic devices produced by young people in talk about climate change?
   Which aspects of each topic are highlighted by these frames and devices?
   What evaluative meanings are presented by these frames and devices?
   How does the language young people use suggest their understanding, attitude and perception of agency with respect to the topic?
4. What commonalities and differences exist in the framing of knowledge and world view between texts produced by and for experts, popular and educational texts, and talk produced by young people?
5. How do the language devices listed above vary in the way they frame knowledge, and stance, agency and evaluation towards science depending on the text creator and intended receiver?

3. Research context
3.1 Science, the public and young people
Scientists recognise the need to communicate with non-scientists, and the way they do this is an object of study in itself. There are two major journals dedicated to this: Science Communication, and Public Understanding of Science; the communication of climate change is of especial interest. Hargreaves et al (2004) and Davies (2004) argue that access to scientific knowledge is important in order for a democracy to function, so that the public can make informed decisions about what is in their interest. In the UK, modern interest can be traced back to the Royal Society’s publication of ‘The Public Understanding of Science’ in 1985 (Hansen, 2009: 106). Since then (ibid), researchers have moved away from a deficit model which stressed the need to provide information, to a concern with public engagement, including the public’s attitudes towards scientific advances, and controversies. The process of sharing scientific information with the public is not straightforward, and divergences between scientists’ beliefs and public opinion have been noted. For instance, in the case of anthropogenic climate change, Farnsworth and Richter (2012) claim that while scientists are reaching consensus on all but the details of the human contribution to change, public perceptions are moving in the opposite direction, with increasing scepticism.
   Within the academic study of Science Education, there is debate about how to develop in young people an understanding of the broad context and contributions of science, and the ability to interpret media information about science critically. McClune and Jamann (2012) write that young people need to understand how science can impact outwards on social questions; they are not just individuals but citizens with the implied duty to make informed and responsible choices. Schreiner et al (2012) relate this specifically to climate change, arguing that young people need to be empowered by scientific knowledge to act for their own future. Morris (2014) reviews the issues in enabling science students to understand the repercussions of science on our own lives, and our responsibilities.

3.2 Linguistic analysis of metaphor, metonymy and related devices across genres
This project will identify and analyse metaphors, metonymy and related linguistic devices in scientific texts. Metaphor is agreed to be central in the development of scientific ideas and their communication (Brown, 2003). For example, the development of scientists’ understanding of the structure of molecules was driven by a succession of metaphorical models and frames that enabled them to visualise and predict molecular structure (ibid). Metaphors also pervade the communication and pedagogy of science: these include protein folding (ibid) the balance of nature (Deignan et al, 2013) and language metaphors for DNA, including script, code, and draft (Seminio, 2008).
   The linguistic analysis of metaphors and metonymies has also been found to be illuminating in showing the frames that underlie world-views and evaluative stance in both general and scientific text (Chartiers-Biack, 2004; Deignan, 2005, 2008). Linguistic metaphors and metonyms can enable the researcher to identify the metaphorical frames through which a topic or an entity is constructed. For example, the existence of the linguistic metaphors flood and waves in a text describing immigrants to the UK could indicate a frame in which immigrants are natural phenomenon without human agency, and one which has catastrophic negative consequences (Seminio, 2008). A balance metaphor for nature implies harmony and thus a negative stance towards change. Metaphorical frames generally have entailments that are not explicitly stated but can be constructed by the reader or hearer, for example, entailments of floods are that they cause damage and should be guarded against.
   Within the linguistics and metaphor literature, there have been a number of studies of how scientific issues such as climate change are presented in non-specialist texts, often but not always
focussing on the quality press. Nerlich and her colleagues have made significant contributions. For instance, Nerlich (2010) analysed media texts around the unauthorized publication of climate scientists’ emails in 2009, showing that the media’s use of religious metaphors tended to frame climate science negatively as “preaching”. She also examined the metaphors in news texts on geengineering as a climate mitigation strategy (2012), identifying metaphors describing geengineering as a thermostat, as sunblock, and as a toolkit, among others. The metaphorical frames oversimplify the very complex science involved, and suggest that the remedy is as easy as, for example, repairing a car. Jaspat al (2013) analysed laypeople’s comments on climate change topics on the Daily Mail website, which used discursive strategies to denigrate scientists. They called for further research into the language of laypeople within the “volatile social context in which science is embedded” (ibid 405), a gap which this project aims to partly meet.

Metonymy is the use of an aspect of a topic to represent the whole topic, and is closely related to metaphor. For instance, in literature on climate change, carbon tends to stand for carbon dioxide, and metonymically for issues associated with climate science, in compounds such as carbon crusade (Koteyko et al, 2010). Like metaphors, they can convey partial or evaluatively-laden perspectives.

Another linguistic devices can also reveal important patterns of meaning. For example, collocation is the study of how words occur together. Deignan et al (2013) analysed a small corpus of texts written by scientists for their peers about the effects of climate change on glaciers and a popular text based on these peer texts. They found that in the popularisation, more than half of all uses of model(s) collocated with show (ending’s), in sentences such as “new data and models show that...”. In scientists’ texts, the collocation of model** with show** did not occur, even though model and the forms modelling, modeled etc were frequently used, and show is a high frequency verb. Close analysis showed that this was because the texts in the two corpora used subtly different understandings of what a model is and can do, which in turn signified different attitudes to evidence and probability. There are a number of other ways in which probability is signaled in language.

The function, or purpose of these language items can also be significant. Deignan et a’s study (2013, discussed above) found that metaphors and metonyms served different functions in the two datasets analysed: in texts written by scientists for their peers, metaphors and metonyms tended to be used as technical terms and to convey research ontology, while in the popular texts they were used to convey a world view, dramatizing and simplifying the message of the peer texts.

4. Research methods
The setting for the research is young people learning about science and related topics in KS3 and KS4, and access to the participants will be through schools. Leeds works with a number of schools through its initial teacher-training arm and has a successful tradition of research partnerships with a number of them. Aspects of the study will inevitably be concerned with what has been studied in school, and therefore, texts used in school, but this is not primarily a study of curriculum and materials. The focus is on language experienced in a broader sense, through popular print and digital media as well as educationally oriented texts.

The two central and complementary analytical tools that this project will use are corpus linguistic analysis and text analysis. Corpus linguistic analysis refers to the automated linguistic analysis, using specialised software, of collections of texts (corpora) that are normally too large to process manually. It is generally acknowledged that for the purposes of identifying frequent patterns of form and meaning in naturally-occurring language this is superior to using unaided intuition; even expert speakers are generally unable to predict language use accurately (Deignan, 2005). Text analysis is the detailed, qualitative study of patterns of form and meaning in texts. Text analysis complements corpus analysis, offering contextual detail and the analysis of fine-grained meaning that can be missed with corpus techniques.

The project will build three corpora, as follows:

(1) A corpus of research and policy texts in the various subfields of climate sciences produced within the last five years.

(2) A corpus of popular and educational materials in fields aligned as closely as possible to the first corpus, consisting of texts aimed at and accessed by young people, including curriculum materials, educational websites such as ‘BBC Bite-size’, popular science texts, internet forums,
Twitter feeds and other texts used by young people (selection informed by interviews with them and with teachers).

3. A corpus of transcribed interviews with young people aged between 11 and 16, in which they discuss their understanding of the issue.

The corpus described in (1) will be approximately 500,000 words. The texts in corpus (2) will be, on average, considerably shorter; for example, a typical article in ‘Focus’, a science magazine for young people, is around 1500 words. Corpus (2) will therefore include approximately 300,000 words. The corpus described in (3) will consist of around 200,000 words of transcribed interview data, made up of approximately 40 interviews of around 30 minutes each (20 hours). All personal identities in Corpus (3) will be anonymised. The texts will be organised so that they can be analysed separately or in batches if required; for instance, in (2), it will be possible for texts from different sources to be analysed separately. At a number of points in the study, we will also talk to science teachers about the process of communicating and translating scientific ideas for young people. This will inform the selection of texts for our corpora, and our analysis of them.

Copyright permission will be sought where required. This will apply to some texts in the first two corpora. This is generally unproblematic for corpus studies, because the texts will not at any point be made available to be read beyond the project team: the corpora will not be made publicly available. Any extracts that are published in project outputs will not be longer than 400 words, and will generally be considerably shorter, and are thus covered by fair-use provisions.

We will analyse the ways in which metaphorical and metonymical frames communicate information, world-view (in particular evaluation and agency) and relationships between ideas, such as compatibility, contrast, cause-effect and instrumentality. We will also analyse associated linguistic features of the texts, such as collocational patterns, modality, function and use of technical language. We will compare the analyses of the three datasets and identify commonalities and divergences in what information and stance are communicated, through which linguistic mechanisms, and, using the third dataset, consider how scientific information and stance are understood and reframed by young people. We will use the analysis to identify general trends in the translation of scientific knowledge and stance into popular and educational genres.

The three corpora will be analysed in various ways as follows. In the first stage, preliminary analysis will involve automatic and manual techniques, in order to identify lines of analysis for full searches, in procedures discussed in the literature on using corpora to investigate metaphor and related phenomena (for example, Cameron & Deignan, 2003; Charteris-Black, 2004). These automatic techniques will consist of identifying the most frequent word forms in each sub-corpus, using corpus software (SketchEngine, https://www.sketchengine.co.uk; WordSmith Tools v.6, http://www.lexically.net/wordsmith), to give an indication of the most frequent semantic themes. We will also identify key semantic areas using WMatrix 3 (http://ucrel.lancs.ac.uk/wmatrix/), using techniques developed by the one of the Co-Is (Jaye) and her colleagues to identify frequent metaphorical patterns (Koller et al 2008). The Key Words function from WordSmith Tools will be used to compare sub-corpora against each other. The Key Words tool compares word frequencies across any two corpora and identifies word forms that are used significantly more frequently in one of the corpora compared to the other. This gives a snapshot of the differences in language choices across corpora. Manual text analysis of samples of each corpus will also be undertaken; this also gives an indication of semantic and formal patterns to be investigated using automatic techniques. Each component of the preliminary analysis will indicate language items to be studied across the full corpora.

The second stage of the analysis will use corpus linguistic techniques, principally concordance and collocation analysis, to analyse in detail the linguistic features identified during the first stage. In techniques pioneered by the PI (2005), metaphorical and metonymical uses will be identified. Metaphor and metonymy analysis will first identify the most frequent and salient groups of metaphors and metonyms used within each corpus, and then identify metaphorical and metonymical frames and their entailments. We will then analyse the most frequent collocations in each corpus, and linguistic markers of evidentiality and probability. We will analyse words that the software identifies as frequent in each corpus for evaluative meaning and function. We will use the British National Corpus, available through SketchEngine, and the Oxford English Corpus, as reference corpora, that is, a guide to non-specialised use as experienced by a typical language user, in order to inform our intuitions about typical meanings and uses of words and structures.
In the third stage of the analysis, we will consider how the world-views constructed through the devices identified compare across the three corpora, and how knowledge, evaluation and stance has been constructed, reconstructed and translated across the different groups of language users.

We will interview science teachers of Key Stages 3 and 4 at a number of points in the project (see timetable below) to inform our choice of texts in the sections of corpus 2, and to establish: (a) their views on communicating, translating and reformulating scientific knowledge to young people, and (b) their opinions of our analysis of samples of the three corpora and our interpretation of them.

5. Management and co-ordination

The project crosses the disciplines of Applied Linguistics and Education. As applied linguists, the PI (0.2 fTE) and one Co-I (0.05 fTE) are well-known scholars of figurative language and the construction of world views. is a former teacher, and an expert in corpus linguistics, including corpus studies of figurative language, and the corpus analysis of language across different genres and registers in science (e.g. et al, 2013). brings expertise in the detailed examination of discourse, especially scientific texts, and especially the use of metaphor theory (2008, 2011). She also brings an established track record of managing funded projects and associated impact activities. The other Co-I (0.1 fTE) is an educational researcher. She is a former secondary science teacher and currently works with science teachers and student teachers in local secondary schools. She has experience of funded research projects with secondary school pupils and teachers (e.g. et al, 2010; & , 2011).

As PI, will oversee all aspects of the management, planning and implementation of the project. will oversee the corpus collation, building, and analysis. She will also oversee the engagement, impact and dissemination activities. will be responsible for the detailed text analysis, will assist with the comparative analysis, and will carry out impact activities with groups other than educators. will oversee all aspects of data gathering from young people and teachers, and will advise on the linguistic analysis from a science and education perspective. She will also carry out impact activities with education professionals. An RA (0.6fTE) will be appointed. He/she will carry out data collection and corpus building. The RA will also transcribe the interviews that make up corpus (3).

The PI and Co-Is will train the RA in the different aspects of data gathering, corpus building and analysis, and will conduct analysis on samples of the data with the RA. The RA will carry out a substantial proportion of the analysis, having been trained and under the supervision of the PI and Co-Is. The RA will be based at Leeds, and will meet at least fortnightly with and will visit Leeds at least five times during the project for review and planning meetings, and will also maintain contact with the Leeds Investigators and RA by Skype and email on a less formal basis.

The project will be benefit from an Advisory Group, which will meet face-to-face in Leeds to review and advise on progress at three points in the project, and which can be consulted less formally at other points. The following have agreed to join the group: University of Leeds, PI on a number of science education projects; and , Lancaster University, author of three relevant books: (Wisconsin, 1990), (Continuum, 2010). We will also be recruiting local science teachers, representatives from science education publishers and examination boards, and curriculum advisors and designers.

Professor , Director of the ESRC-funded , has indicated that the project can be affiliated to . Members of the project team already have links with , who is on the ‘Challenge Panel’ of outside experts advising on ’s work, and leads a project associated with . This affiliation will give access to additional computational linguistic expertise held at . In technical areas such as automatically harvesting Twitter and online forum data (for the second corpus), and other specialist techniques that may be needed as the project proceeds.

6. Timetable

The project will encompass the following tasks:

1. Collect texts for Corpus 1: code, upload.
2. Interview secondary school pupils
3. Transcribe interviews to create Corpus 3
4. Collect texts for Corpus 2: code, upload
5. Corpus and text analysis of Corpus 1
6. Corpus and text analysis of Corpus 3
7. Corpus and text analysis of Corpus 2
8. Interview teachers
9. Comparative analyses of Corpora
10. Outputs, dissemination and impact activities

Timing of the tasks is as follows, starting from November 1st 2014:

<table>
<thead>
<tr>
<th>2014</th>
<th>2015</th>
<th>2016</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>J</td>
<td>E</td>
</tr>
<tr>
<td>N</td>
<td>D</td>
<td>J</td>
</tr>
<tr>
<td>O</td>
<td>E</td>
<td>M</td>
</tr>
<tr>
<td>E</td>
<td>A</td>
<td>M</td>
</tr>
<tr>
<td>A</td>
<td>e</td>
<td>A</td>
</tr>
<tr>
<td>e</td>
<td>a</td>
<td>p</td>
</tr>
<tr>
<td>c</td>
<td>p</td>
<td>a</td>
</tr>
<tr>
<td>o</td>
<td>u</td>
<td>u</td>
</tr>
<tr>
<td>e</td>
<td>e</td>
<td>c</td>
</tr>
<tr>
<td>l</td>
<td>o</td>
<td>e</td>
</tr>
<tr>
<td>e</td>
<td>a</td>
<td>e</td>
</tr>
<tr>
<td>a</td>
<td>a</td>
<td>p</td>
</tr>
<tr>
<td>p</td>
<td>a</td>
<td>u</td>
</tr>
<tr>
<td>a</td>
<td>u</td>
<td>u</td>
</tr>
<tr>
<td>l</td>
<td>o</td>
<td>e</td>
</tr>
<tr>
<td>e</td>
<td>a</td>
<td>e</td>
</tr>
<tr>
<td>a</td>
<td>a</td>
<td>p</td>
</tr>
<tr>
<td>p</td>
<td>a</td>
<td>u</td>
</tr>
</tbody>
</table>

7. Outputs and dissemination

Through our science teacher education work at Leeds, we have close contacts with teachers and other education professionals, with whom we have discussed the project. We have also made contact with other user organisations such as the Royal Society and the UK Youth Climate Coalition.

We will present outputs at the Association for Science Education conference (UK, early 2016), ESERA (European Science Education Research Association, Helsinki, 2015); and at specialised linguistic conferences such as: Researching and Applying Metaphor (Leiden, 2015); International Conference of Corpus Linguistics (Lancaster, 2015).

We will write four journal articles, targeted at science communication journals such as Science Communication and Public Understanding of Science, at journals of science education such as Studies in Science Education and International Journal of Science Education, and at the specialised linguistic journals International Journal of Corpus Linguistics and Metaphor in the Social World. We will also submit articles discussing our findings for the non-specialised educated public to quality newspapers such as the Times Educational Supplement and the Guardian.

We will organise engagement and dissemination events for education professionals in the Yorkshire region. We will also hold a national event for curriculum developers, science journalists and policy-makers in London, at the Royal Society, following the model of the successful event for a similar group held by a previous funded project by Leeds staff (EISER) (see Pathways to Impact). We will seek advice from the AHRC Theme Fellow on extending the engagement and dissemination activities, and developing further research and impact projects, in collaboration with other projects funded under the theme where possible.

In terms of legacy, the project will create collaborative links between corpus work at Leeds, the internationally-known expertise in Science Education, also at Leeds, and work at Lancaster, known for its applied linguistic and corpus excellence. The project aims to start a tradition of bringing language analytical techniques to the investigation of texts for young people, and to the consideration of intralingual translation of research to popular and educational genres.

8. Technical summary
8.1 Digital outputs

The project will produce a website, which will be linked to the investigators' webpages and departmental websites, and to the website at Lancaster University.
The project will produce the three corpora described above, which will be stored securely at Leeds. At present, we do not plan to make corpora (1) and (2) available beyond the project staff because of copyright issues. These corpora are not therefore considered to be an output of the project. A transcribed version of Corpus (3) will be made available to other researchers (see Technical Plan).

8.2 Digital technologies

We will use three specialised pieces of software for processing the language data: WordSmith Tools v. 6; SketchEngine and WMatrix3. WordSmith Tools and SketchEngine are commercially available while WMatrix was developed and is owned by Lancaster University (see Technical Plan).

References


