

corresponding sessions. Enquiries about this feature should be addressed to Professor J. Lachlan Mackenzie (VU University Amsterdam, NL and ILTEC, PT): **lachlan_mackenzie@hotmail.com** Further details will appear on the conference website on 30 September 2015 and in the Winter 2015 issue of *The European English Messenger*. The deadline for the submission of applications will be **28 February 2016**.

ACADEMIC PROGRAMME COMMITTEE:

Professor Patrick Lonergan (NUI Galway, Galway, Ireland) (chair)

Professor Lieven Buysse (University of Leuven, Belgium)

Professor Claire Connolly (University College Cork, Ireland)

Professor Anne Fogarty (University College Dublin, Ireland)

Professor Irene Gilsenan Nordin (Dalarna University, Sweden)

Dr Aoife Leahy (Independent Scholar, Ireland)

Professor Biljana Mišić Ilić (University of Niš, Serbia)

Professor Nóra Séllei (University of Debrecen, Hungary)

Associate Professor Slávka Tomaščíková (P. J. Šafárik University in Košice, Slovakia)

GENERAL QUESTIONS ABOUT THE ESSE 2016 CONFERENCE IN NUI GALWAY SHOULD BE SENT BY EMAIL TO info@esse2016.org

LANGUAGE IN USE

Investigating Academic Identity Traits

Maurizio Gotti
Università di Bergamo, Italy

In the last few decades there has been a considerable increase in the globalisation of research practices in Universities all over the world. The wish to be read and quoted as widely as possible at a global level has induced many scholars to publish in English. These globalising trends in the academic world have had great implications for research, resulting in very complex formulations. Indeed, in spite of the homogenizing effects deriving from the process of globalisation, academic discourse does not seem to be at all uniform, but shows great variation based on several factors, some of them clearly reflecting aspects of the local tradition and culture. The analysis presented here originates from a recent research project on identity and culture in academic discourse carried out at the University of Bergamo (Italy), whose data show that the (native or non-native) Anglophone textual realisations are clearly influenced by their authors' cultural allegiance to their linguistic, professional, social, or national reference groups.

1. Globalising trends in the academic field

Academia is a field in which important changes due to globalising trends have emerged, with several cases of language variation linked to the encounter/collision of different cultural frameworks. The internationalisation of English academic discourse has not only been observed in Anglophone countries, but wherever institutional and professional settings have evolved in ways that transcend the linguistic, cultural and conceptual standards of their local communities. The gradual globalisation or hybridisation of discursive practices that first appeared in English-speaking environments now

significantly affects also smaller languages, which are subject to standardising pressures in their semantic, textual, sociopragmatic and even lexicogrammatical construction.

Hegemonic tendencies have been identified, especially in the language policies commonly adopted by major international publications employing English as the world's academic lingua franca. Non-native scholars are expected to have good English literacy skills, so as to be able to present their papers in that language at conferences and publish them in peer-reviewed journals and volumes. This expectation has greatly influenced academics, with the result that the last three decades have seen a massive conversion of journals from other languages to English, thus determining "a real loss in professional registers in many national cultures with long scholarly traditions" (Swales 2000: 67).

This trend has had a number of serious consequences. The first is the concentration of immense power in the hands of a restricted group of academic gatekeepers, located in very few countries in the world. These few countries have attained the right to enforce norms and to certify the academic recognition of research carried out all over the world. Their academic power in certain disciplines is so strong, that it can decide the careers of scholars who need to publish in leading international journals to validate and disseminate their research findings. There is therefore a risk of linguistic monopoly, scholarly chauvinism and cultural imperialism, which may give rise to unintentional - or even intentional - discrimination against non-native speakers on the part of the editors of specialised publications. The demands associated with writing and publishing in English are usually very strict and can be used by academic publications to filter foreign contributions. Moreover, since only the British or American varieties are favoured, a failure to comply with the journal's linguistic standards is usually penalised with rejection.

Difficulties may also be experienced by non-native writers in interpreting editorial feedback after the refereeing process and in discerning what revisions are to be made. In many cases, in order to have their articles or books accepted, non-native writers have to rewrite them several times, modifying the original style typical of their local way of writing, and adopting the rhetorical conventions commonly shared by the Anglo-American scientific community - what Hyland (2000: 13) calls the prevailing 'tribal lore' of a specific discipline. The result of this policy of *linguisticism* - to use the term coined by Phillipson (1992) to denote language-based discrimination - is a marked increase in the intellectual dominance of English-speaking centres. There is a risk that 'periphery' perspectives in the various disciplines may have no influence on the trends developed in intellectual centres located in a small number of monopolising academies. The periphery, instead, may play a healthy role by questioning views prevailing in the centre and providing alternative perspectives. In recent years, there has been a heightened awareness in the academic world of the valuable contribution of non-Anglophone scholars working within dominant research paradigms and agendas. However, this increased awareness has rarely "translated into a recognition that the discipline[s] are] also 'owned' nowadays (to use the new management-speak) by a very large number of people for whom English is neither a first, nor a second language" (Kayman 2003: 52).

In some cases, 'periphery' publications have changed their language or even title to suggest a more international collocation. For example, in 2006 the *Italian Heart Journal* (already published in English) changed its name to the *Journal of Cardiovascular Medicine*. As local journals are considered second-class research tools by the Italian medical community and since medical literature is regarded as being more competitive if published in the UK or the US, the scientific board of the *Italian Heart Journal* decided to conceal the peripheral provenance of the journal by assigning it to an American publisher, while maintaining an Italian editor.

While these hegemonistic developments have caused a reaction of alarm in non-English-speaking countries, there is a growing awareness of the need to free international

English from native-speaker ethnocentricity. More and more non-native writers who agree to adopt an internationally-used language such as English want to retain a personal style in their additional language. In a few cases, editors sympathise with this view and accept variation in discourse style and nativised varieties, the key criterion of acceptability being whether or not the readership would be likely to understand the contribution. On the other hand, some internationally-renowned journals and book series accept papers and manuscripts written in other languages than English, particularly in the fields of linguistics and education and whenever such languages are the object of specialised research.

This reaction is not an isolated case. Indeed, the strong English-language policies frequently adopted by academic authorities in many countries have aroused people's awareness of the risk that the increasing use of English in publishing and higher education might greatly reduce the role of national languages for scholarly purposes. Many countries are becoming aware of the problem of erosion of functionality in their languages and have launched policies aimed at strengthening the role of the local medium at different stages of education and in various domains of communication. This is particularly true of some European nations, which are committed to defending the prestige of the local language. For example, the Academy of the German Language has warned universities against reducing the standards of scholarly German and replacing it with 'bad simple English', and has pointed out the dangers of reducing German to "a system with restricted functional range" (Görlach 2002: 16). Nordic countries have started a research project called *Nordens språk som vetenskapsspråk* ['The Nordic languages as languages of science'] to defend the use of their languages for academic and scientific purposes, as they deem this fundamental for the acquisition of a strong competitive position in culture and science. Stimulated by the results of this project, policies of domain (re)conquest are being promoted in several contexts.

2. The CERLIS research project

Within this context, various members of CERLIS (the research centre on specialised discourse based at the University of Bergamo) have chosen to investigate the relationship between socioculturally-oriented identity factors and textual variation in English academic discourse, focusing in particular on the identification of identity traits typical of different branches of learning. Within such domains, we have investigated to what extent the cultural allegiance of (native or non-native) Anglophone discourse communities to their linguistic, professional, social, or national reference groups is affected by the use of English as a lingua franca of international communication. To a certain extent, the process of internationalisation of English has strengthened its hegemonic tendencies, with the result that local communities are often marginalised, thus preventing an authentic intercultural discourse. This process is most evident in discourses (e.g. academic, technical, scientific and legal communication) where the socialisation/textualisation of knowledge plays a crucial cohesive role.

Early results from research carried out by our group of investigators indicated that the internationalisation which makes English the preferred choice of code is coupled with textual inconsistencies and ambiguities that advise against straightforward, simplified conclusions: the apparent dominance of 'Anglocentric' models in the domains considered revealed specific adaptive attitudes and evidence of cultural resistance in the textual strategies that construct identity-shaping differences. Furthermore, background research by our team highlighted the complex pragmatic functions of the texts concerned, which are mostly constructed by deploying strong culturally-connoted values. Our work has therefore moved in this direction, identifying cases of language variation linked to the encounter/collision of different cultural frameworks within English academic discourse.

The approach has singled out various aspects for textual analysis, such as the macrostructural elements of linguistic variation, the lexico-semantic development of disciplinary discourses, their rhetorical-pragmatic strategies, and specific textual phenomena, such as hedging and verbal modality.

In order to enable an easy comparison of results, the structure of our investigation is similar to that of other research projects on identity traits in academic discourse, especially KIAP and SERAC. The KIAP Project (*Cultural Identity in Academic Prose*, <<http://www.kiap.uib.no/>>), carried out by the University of Bergen, Norway, has adopted a double contrastive perspective, investigating the linguistic characteristics of research articles from different disciplines (Economics, Linguistics, Medicine) and different languages (English, French, Norwegian). Its main results show that authors of research articles tend to write more like their disciplinary colleagues using other languages than their language-community co-members working in other disciplines (Fløttum, Dahl, Kinn 2006). The SERAC (*Spanish/English Research Article Corpus*) Project, conducted at the University of Zaragoza, Spain (<www.interlae.com>), was designed to enable its researchers to carry out an intercultural/rhetorical comparison of the work of Anglo-American scholars writing research articles in English, and of Spanish scholars writing research articles in English and Spanish. It has identified various rhetorical strategies which are used in a different way by native-English scholars and Spanish academics when they write in international or local journals (see, among others, Pérez-Llantada 2012).

Within this context, the aim of our project was to improve the understanding of identity-forming features linked to 'local' or professional cultures, as communicated through contemporary English in various specialised domains by native and non-native speakers. Through an exploration of major social and academic factors, it has evaluated how far international audiences in key intercultural domains trigger textual reconfigurations that simplify, distort or even remove non-congruent institutional and cultural traits, while enhancing the identities of specific socio-professional communities.

3. The CADIS Corpus

As corpora constitute a remarkable tool for the study of discourse, a specific corpus (Corpus of Academic Discourse, or CADIS) was assembled as the core and foundation of this line of research (<www.unibg.it/cerlis/cadis/>). In view of an in-depth analysis of variation in intercultural communication, we have selected a range of texts produced by scholars and academic institutions in various parts of the world. To identify textual variants arising from the use of English as a native language or as the lingua franca of science, we have used a corpus formed by English texts for academic communication. The corpus also comprises some Italian texts for comparative purposes. As regards the analysis of non-native English, we should remember that such texts are usually the result of careful editing carried out by native speakers, either before submission or during the process of revision for publication. This is the reason why we investigated the writing of Italian academics not only in their texts in English, but also in their own native language.

Besides including two different languages, CADIS represents four separate disciplinary areas: Law, Economics, Applied Linguistics and Medicine. For each disciplinary area, various textual genres have been considered: abstracts, articles, book reviews, and editorials. These texts reflect the original project, but in the last few years CADIS has been widened to include further genres. The first addition consists of a subcorpus of research letters in the field of Medicine, an emerging genre which has become widely used in medical journals as a brief, timely and useful tool for rapid publication in the field. Another subcorpus which has been added to the original

categories is that of academic posters presented at international conferences. This addition to the corpus is in line with recent developments in the literature on academic discourse, which have turned from the 'open genre network' of academic writing (Swales, Feak 2000) comprising the most visible research genres (e.g. journal articles, abstracts, dissertations and conference proposals) to some of the lesser known 'occluded' genres of academia. Conference poster presentations rank among the latter and have often been treated as 'a poor country cousin' to oral presentations:

The structural complexity of CADIS reflects its contrastive orientation: it is designed to be internally comparable, so its texts can be analysed not only by disciplinary area, genre, language and culture, but also historically. This is possible because the corpus covers a time frame of over thirty years, from 1980 to 2011. Including all language groups - native speakers and non-native speakers of English, and native speakers of Italian - a total of 2,738 texts (from 635 to 739 per disciplinary area) have been inserted in the corpus, which now includes over 12 million words.

Over the last few years the CADIS corpus has been used as a basis for several studies linked to our research programme. The variety of texts chosen for inclusion in the corpus and their grouping into homogeneous categories have enabled our team to analyse a range of macro/microlinguistic variants in terms of academic identity and to interpret our findings in the light of recent scholarship. Specifically, CADIS has proved to be a useful tool for research targeting the following features:

- a) the generic macrostructure and its lexico-grammatical realisations;
- b) speech acts expressing positive/negative evaluation, whether exophoric or metatextual;
- c) the pragmatic, interpersonal plane of discourse (stance, hedging, politeness strategies);
- d) the various functions of verbal and lexical modality;
- e) textual evidence of gender-related phenomena and academic status.

Although its coverage is not exhaustive, CADIS has proved to be a highly useful tool for exploring authentic discourse, as shown by its impact on all the areas of language analysis represented by our research unit.

4. Main results

Some of the main results of the project are reported in Gotti (2012), with chapters by Ulisse Belotti, Larissa D'Angelo, Davide Giannoni, Maurizio Gotti, Stefania Maci and Michele Sala. To facilitate a comparison of the various perspectives taken by their authors, such contributions have been grouped into three sections, each of which highlights different aspects of identity traits in academic texts. These analytical sections are preceded by two introductory chapters by Maurizio Gotti that sum up some of the previous studies on the subject, present the research objectives and activities, and describe the corpus on which our investigation is based.

The first section of the volume deals with identity traits across languages and cultures, investigating how the use of a given language affects the writing of a scholar, especially when it is not his/her native language. This is particularly evident in the case of English, whose recurrent use by non-native speakers consequently requires a degree of adaptation of their thought patterns and expressive habits. Moreover, as language is strictly linked to the setting in which it is used, cultural elements have been found to operate as key contextual constraints, influencing both the level of discursive organisation and its realisations. Indeed, Davide Giannoni's analysis of Anglo-American journals, English-medium Italian journals and standard Italian journals suggests a considerable extent of intradisciplinary variation, both within and across languages/cultures. Larissa D'Angelo - by analysing book reviews written in English and Italian by native and non-native speakers - shows how reviewers of different nationalities, within the disciplines of Applied Linguistics, Economics, Law and Medicine, express positive and negative ap-

praisals of their peers' work. Comparing the argumentative strategies employed in medical research articles written by native speakers of English with those written by Italian non-native speakers of English, Stefania Maci identifies relevant cross-cultural differences in terms of argumentative devices employed by their authors. Michele Sala has found that different rhetorical styles and strategies are employed by native and non-native speakers of English and by experts of Common Law vs. experts of Civil Law systems when they discuss legal subjects.

The second section of the volume comprises investigations of identity features that characterise specific disciplinary communities or that mark a differentiation from other branches of knowledge. Indeed, several studies have described how scholars in different fields represent themselves and their work in very different ways. This theme is analysed in five chapters. Ulisse Belotti examines how economists manifest their identity in research article abstracts written by single authors or multiple authors. Evaluation and popularisation in journal editorials is the subject investigated by Davide Giannoni, whose chapter focuses on the presence of these two closely-related phenomena in a corpus of journals in Medicine and Applied Linguistics. In the following chapter, Davide Giannoni investigates how leading scholars in four academic disciplines (Applied Linguistics, Economics, Law and Medicine) portray themselves through metaphoric realisations. The corpus material considered consists of editorials (a genre whose intensely evaluative orientation tends to generate disciplinary conflicts linked to the validation of new knowledge claims), focusing in particular on the *message editorial* subgenre, which is where editors share experiences with readers or solicit their collaboration. In the next chapter, Michele Sala examines the use of interactive metadiscursive resources, comparing research articles in Applied Linguistics, Economics, Law and Medicine published over a thirty-year period from the early 1980s to the present day. His purpose is to trace recognisable trends in the use of such resources and to relate possible variations to socio-cultural changes in discursive practices. In the following chapter, Michele Sala explores the interpersonal dimension of academic discourse by investigating the use of interrogative forms in research articles from a diachronic perspective. His study analyses a corpus from the domains of Applied Linguistics, Economics, Law and Medicine in the thirty-year period ranging from 1980 to 2010.

The third section of the volume deals with identity aspects emerging from genre and gender variation. The chapters in this section confirm that written academic communication greatly relies on compliance with textual norms governing the construction of its different genres, whose conventions are learned through training and engagement with the disciplinary community. In recent years, however, the availability of new technological tools has expanded the range of texts targeted by analysts. There is an increasing interest in phenomena that lie beyond single genres, including a variety of non-traditional semiotic modes such as visual presentations and online material. This has led to the identification of discrepancies between traditional textual conventions and their actual realisations, through new concepts such as *genre mixing*, *repurposing* and *hybridisation* to account for generic dynamism. One of these new genres is the research letter, a type of publication meant to facilitate fast dissemination of primary research; a similar genre was first identified by Hyland (2000) in Physics, Chemistry and Microbiology. Stefania Maci investigates the use of research letters in medical publications and finds that their scope and purpose are different in the medical field: while Hyland's letters allowed the reporting of new results and ideas to a wider community than specialists and are a means of promoting young scientists, the authors of medical research letters range from young scientists to established scholars. The next two chapters deal with conference posters, a genre which has often been marginalised across disciplinary fields and is considered less prestigious than paper presentations. However, poster presentations play an

important part in scientific conferences and constitute an interesting alternative to papers. By facilitating informal discussions between presenters and their audience, posters provide a more intimate forum for exchange than regular paper presentations. Also, compared with other texts, this multimodal academic genre differs for its lack of prescriptive guidelines, allowance for creativity and individuality, space for narratives and stories, and the goal to both inform and persuade. In her analysis of a corpus of 120 posters from four different disciplines (Applied Linguistics, Medicine, Economics and Law), Larissa D'Angelo has identified a range of features: the majority of the posters collected follow conventions based on best-practice guidelines and advice currently available online from university websites, personal weblogs, conference webpages, university writing centres, poster websites and online poster journals. Moreover, as the poster presentation is a multimodal communicative event, with writing, graphics, colour, speech, and even gesture used to convey meaning, this genre clearly displays 'hybrid' features and cannot merely be considered a visual, abbreviated rendition of the research paper. Indeed, what makes the poster unique is the fact that its visual and textual features work together to successfully convey meaning. Posters are also investigated by Stefania Maci, whose chapter analyses the relationship between text and images, as well as the relevance of images in relation to text. The last chapter in this volume deals with the role of gender in defining authorial identity in academic writing. Using a corpus of book reviews and research articles, Larissa D'Angelo analyses the different persuasive resources chosen by male and female scholars and by novice and expert members of the academic community to argue their views, in order to determine whether, to what extent, and in what terms gender and authority influence their argumentation.

5. Conclusive remarks

The CERLIS research project has confirmed that academic discourse is not at all uniform, but varies according to a host of factors, such as language competence, local culture, disciplinary field, community membership, professional expertise, gender and generic conventions. Moreover, the differences observed in the data analysed offer valuable evidence of the existence of individual proclivities in the choice of linguistic items and the use of particular argumentative strategies. A recognition of the conscious choices made by academic writers means that textual differences cannot merely be interpreted in terms of desirable vs. non-desirable or standard vs. non-standard options, but are clearly a result of the deliberate strategic (and often creative) avenues explored by authors to further their rhetorical objectives.

Taken together, the research results also reflect the considerable challenges and opportunities that confront scholars seeking to achieve a delicate balance between their willingness to adhere to the norms and conventions of their professional community and the desire to express individual values and identity traits. Such factors have been found to interact, producing transversal identities that often betray their dependence on local traits and traditions, thus giving rise to textual realisations characterised by hybridising forms deriving from intercultural clashes.

References

- Fløttum, Kjersti, Trine Dahl, and Torodd Kinn 2006. *Academic Voices: Across Languages and Disciplines*. Amsterdam: Benjamins.
- Görlach, Manfred 2002. *Still More Englishes*. Amsterdam: Benjamins.
- Gotti, Maurizio (ed.) 2012. *Academic Identity Traits: A Corpus-based Investigation*. Bern: Peter Lang.
- Hyland, Ken 2000. *Disciplinary Discourses. Social Interactions in Academic Writing*. London: Longman.
- Kayman, Martin 2003. The Foreignness of English. *The European English Messenger* 12/2, 52-53.

- Pérez-Llantada, Carmen 2012. *Scientific Discourse and the Rhetoric of Globalization: The Impact of Culture and Language*. London: Continuum.
- Phillipson, Robert 1992. *Linguistic Imperialism*. Oxford: Oxford University Press.
- Swales, John 2000. Languages for Specific Purposes. *Annual Review of Applied Linguistics* 20, 59-76.
- Swales, John, Christine Feak 2000. *English in Today's Research World: A Writer's Guide*. Michigan: Michigan University Press.

Checking comprehension in English-medium lectures in technical and scientific fields¹

Alessandra Molino
University of Turin

1. Introduction

This paper reports the results of a study on some interpersonal strategies adopted by Italian lecturers using English as a medium of instruction to engage students in discourse, keep their attention alive and ensure comprehension. The focus is on a range of devices including discourse markers, such as *okay? clear?*, as well as full questions, such as *did you understand up to this point?*, which might be grouped together under the umbrella term of comprehension checks. In the field of Second Language Acquisition, comprehension checks are defined as resources “to anticipate and prevent a breakdown in communication” (Long 1983: 136) and are seen as pre-emptive teacher-initiated strategies that foster language learning. In this study, the label ‘comprehension check’ (henceforth CC) is used to cover a variety of interrogative structures which lecturers employ to ensure successful communication, either by checking that students have understood concepts and ideas clearly or by making sure that the various macro- and micro- phases of the communicative event are processed correctly. Following Walsh (2011: 7), I will argue that not only do lecturers “modify their interactional resources to assist comprehension”, but they also adjust their talk to “help learners ‘navigate the discourse’”. Using Hyland’s (2005) distinction between interactional (i.e. involving the audience) and interactive (i.e. organizing discourse) metadiscourse, I will show that CCs not only reflect the need of checking understanding and opening the floor to student intervention, but they also have interactive properties that help lecturers structure discourse into meaningful chunks, accommodating the students’ processing abilities.

2. Materials and methods

Seven EMI lectures in engineering-related disciplines delivered at an Italian Polytechnic University were investigated. The lectures were selected in such a way as to include a variety of subjects, ranging from Mathematical Analysis to Chemistry, as well as different lecture types, such as classes presenting introductory information about the course, lectures proving theorems and deriving formulae and lessons in which theoretical models were presented. All the lectures were aimed at large classes of undergraduates (i.e. equal to or more than 50 students). Each lecture was video recorded and transcribed orthographically following the conventions adopted for the *Michigan Corpus of Spoken Academic English* (MICASE) (Simpson *et al.* 2002). Table 1 provides information about the length of lectures and the number of tokens in the transcriptions. Approximately nine

¹ This paper is a product of the research project: “English in Italy: Linguistic, Educational and Professional Challenges” co-financed by the Compagnia di S. Paolo, Progetti di Ateneo 2012, University of Turin, and coordinated by Virginia Pulcini.

hours were analysed, corresponding to circa 56 thousand words. The average speech rate of the seven lectures is 107 words per minute (wpm).²

	Length	Tokens	wpm
Lecture 1	1:19:13	11,567	146
Lecture 2	1:44:23	11,089	107
Lecture 3	1:06:49	5,868	89
Lecture 4	1:17:50	9,307	121
Lecture 5	1:07:46	6,206	93
Lecture 6	1:10:49	6,537	93
Lecture 7	0:59:12	5,911	100
Total/average wpm	8:46:02	56,485	107

Table 1. Lectures' length, tokens and speech rate (words/minute)

In the transcription, question marks were used to signal utterances that functioned pragmatically as questions. Therefore, thanks to the search option of word processing software, all questions were identified and a list of types of comprehension checks was compiled. The list served as a basis to retrieve tokens using a concordance programme, so as to quantify occurrences. Nevertheless, the functional classification of instances was done manually while listening to the recordings, as prosodic features were important for understanding the use of CCs in discourse.

3. Results

3.1 Lecture type, lecturing style, student participation and CCs

As illustrated in Table 2, the total number of CCs is 958, meaning that there is one instance every 59 words, i.e. slightly less than two per minute (average ratio: 1,82 per minute). This suggests that, on the whole, lecturers are rather proactive in addressing the comprehension needs of the audience. It should be noted, however, that the frequency with which lecturers ask for the students' feedback may vary, depending on factors such as the type of lecture and the lecturing style. The highest figures are found when lecturers prove theorems and derive formulae, indicating that they constantly interrupt the flow of discourse to ensure that students understand what is being said. This seems to be the case even if the lecturing style is not exactly "conversational" (Dudley-Evans and Johns 1981).³ While lectures 3 and 4 are not delivered as if the lecturer were reading a paper, the style is not informal and the lecturers explain theorems while simultaneously writing formulae on the electronic whiteboard, so the impression is that they are almost reading what they are writing. On the other hand, the lowest figures are found in lectures in which theoretical models are presented, as suggested by the data obtained for Lectures 5 and 6. Finally, individual proclivities seem to play a role, as also noted by Crawford Camiciottoli (2007: 108): while Lecture 6 is built around the description of a model, the lecturer's performance is almost "rhetorical" (Dudley-Evans and Johns 1981), characterised by some asides and jokes.

² According to Tauroza and Allison (1990), the average speed of lectures in native British English is between 125 and 160 wpm. Table 2 indicates that there is some variation, from slow (89 wpm) to average speed (146 wpm), but on the whole the lectures analysed here show a reduced speech rate as compared to the native speaker average.

³ Dudley-Evans and Johns (1981) identified three types of lecturing style: the *reading* style, in which lectures read aloud or deliver the lecture as if they were reading; the *conversational* style, in which lecturers use notes, opt for a relatively informal style and encourage student involvement; and the *rhetorical* style, in which lectures promote a "stage-like atmosphere" (DeCarrico and Nattinger 1988: 93) with digressions, anecdotes and jokes.

	Lecture type	Lecturing style	Student participation	No. of CCs	CCs/min
Lecture 1	Course introduction	Conversational	2 + show of hands	153	1,94
Lecture 2	Course introduction	Conversational	2 + show of hands	65	0,84
Lecture 3	Theorems and formulae	Reading / Conversational	3	217	3,24
Lecture 4	Theorems and formulae	Reading / Conversational	4	420	6,00
Lecture 5	Models	Reading / Conversational	1	4	0,07
Lecture 6	Models	Conversational / Rhetorical	19	98	0,94
Lecture 7	Models	Reading / Conversational	4	1	0,02
TOTAL				958	1,82

Table 2. Lecture type and style, student participation, total number of CCs and CCs/minute

Finally, it may be interesting to consider the extent of student participation. In Table 2 participation is indicated as the number of interruptions to teacher talk during the whole lesson, following Nesi (2001). In all cases, including Lecture 6 which shows a higher number of interventions, audience involvement is limited to very short turns in which students answer questions or ask for clarifications. Therefore, there does not seem to be a correlation between the frequency of CCs and the degree of student participation, suggesting that CCs, while opening up a space for interaction, do not function as elicitors proper seeking a response from students.

The question therefore arises as to what role CCs have in lecture discourse. In a study dealing with the sociolinguistic variation of the use of discourse markers in English lectures and seminars from the MICASE corpus, Schlee (2005) identifies the following functions for the items *okay?* and *right?*: transition markers, i.e. signalling information stage and indicating discourse structure, and progression checks, verifying whether the audience is following.⁴ Similar functions were noted by Othman (2010) analysing the same items in native speaker lectures, but she also recognises instances functioning as response elicitors and seeks of assurance.

In the following sections, bearing these functions in mind, I will try to determine whether the same uses are found in EMI lectures by Italian native speakers, whether specific ones can be related to the adoption of English as a Lingua Franca and whether any recurrent form-function association pattern is noticeable, given that the present analysis includes a wider range of devices than discourse markers. To this aim, I will analyse CCs in terms of their form (e.g. tag or full question), their position within the utterance and their function.

3.2 The form of comprehension checks

Four types of CCs were identified: tags, nominal constructions, verbal constructions and interrogative clauses (Table 3). Tags include the discourse markers *okay? eh? ah? mm?*

⁴ It should be noted that Schlee (2005) investigates academic speech by both instructors and students. In his classification scheme, therefore, Schlee also includes modal question tags, asking for confirmation or information and backchannel signals, i.e. feedback expressions that indicate that the listener is engaged in conversation while not willing to take the floor. These uses were not found in the present sample of lectures, very likely because of the limited interaction between lecturers and students.

right? clear? fine? and the expressions *okay or not?* and *yes or not?* Tags represent the most frequent category with 890 occurrences. The most widely used tag is *okay?* (589 hits), followed by *mm?* (179 hits) and *eh?* (94 hits). Nominal constructions comprise expressions such as *any questions?* and other verbless structures which reflect the non-native status of lecturers (e.g. *that clear? that fine?*). Verbal constructions, on the other hand, include elliptical expressions such as *got it?* and, again, non-standard instances such as *you all agree?* Interrogative clauses are proper questions formulated with subject-verb inversion or the use of auxiliary verbs. Examples are: *did you get the point? do you agree with this? is everything clear?*

Type of comprehension check	Occurrences
Tags E.g.: <i>okay?/ eh?/ mm?/ right?</i>	890
Nominal constructions E.g.: <i>any questions?</i>	20
Verbal constructions E.g.: <i>got it? /see?</i>	23
Interrogative clauses E.g.: <i>did you understand up to this point?</i>	25
Total	958

Table 3. Type of comprehension check

3.3. The position of CCs within the utterance

The data analysed indicate that CCs may be found in three main positions. They occur at the end of clauses concluding a discourse phase (e.g. discussion of theoretical aspects, the illustration of examples, the summary of what has been discussed, the evaluation of terms or theories) (see Young (1994) for a description of the macro-structure of university lectures) or concluding a topic (see Hansen (1994) for the identification of topics, subtopics and minor points). CCs also occur between clauses that deal with the same issue, thus not precluding to a topic shift. Finally, CCs may also occur immediately after a phrase (any type, but mostly noun phrases and verb phrases) in the middle of clauses. Example 1) illustrates these three cases.⁵

(1) S1: [...] our object is this one **clear?** a thermal machine which exchanges **eh?** some heat **eh?** together with S1 S2 Sn **okay?** this machine exchanges heat and exchanging eat_ heat <SELF-CORRECTION> produces work **mm?** and the total word_ work <SELF-CORRECTION> made by this_ produced by this is W **okay?** of course this is a cyclic machine we are looking at a cycle one cycle and inside this cycle **eh?** we have this exchanges (*sic.*) of heat H1 is the heat exchanged in its thermostatic source T1 i eh draw this arrow and this this arrow indicates that the heat is flowing eh excuse me is absorbed by the thermal machine so in this case H1 is positive **is clear? mm?** in the second one for example ok? the heat is released by the machine to the thermal source number two ok? [...] in this case i indicate by H1 H2 H3 the exchanged heat that can be positive or negative depending on the direction of the flow of heat eh? from thermal source to a machine is positive from machine to thermal source becomes negative **ok? is clear this this assumption? ok?** [Lecture 4]

3.4 The function of comprehension checks

The first aspect that emerges analysing CCs in discourse is that they are multifunctional devices: a large number of occurrences conveys interactive and interactional

⁵ In all the examples, the items on which attention should be focused are in bold type.

metadiscourse meanings concurrently. For ease of discussion, the various uses in discourse will be presented individually, but in fact in many cases it is difficult to identify a prevailing function.

According to Hyland (2005: 49), interactive metadiscourse reflects the text producer's awareness of the audience and organises discourse guiding the receivers through the text.⁶ In the lectures investigated there are various examples in which CCs seem to contribute to the structuring of discourse, as was also noted by Schleef (2005) and Othman (2010). CCs may work as topic management resources operating as "closing brackets for old topics" (Hansen 1994: 135) or precluding to a different discourse phase. This was illustrated in example 1 in the previous section and can also be noted in example 2 below:

(2) S1: [...] ok just some extremes ehh and ok and then smaller groups of *elettotecnica* *elettrica* electrical engineer one and two, those people are fighting hard here <LAUGHS> eh <LAUGHS> and other people from ok this one *meccanica* ok he's already repented eh *telecomunicazioni* you're there and this one i don't know, *elettronica* is_ (xx) and Z might be the *magistrale* [S2: <INAUDIBLE>] eh probably, ehh and one *gestionale* and one energetics mmh? you are energetics and *gestionale*? it's out gaining money eh ah ok, ok so just so that you start also knowing each other the only information i'm missing is form the others from this group mmh? i won't to show the names up here in the recorded lecture ok? so we can do that offline. **okay?** so, so much for the organization **any question so far?** [Lecture 1]

Rather recurrent patterns were noted in CCs with a discourse structuring function. These instances appear to have a stronger elicitive force than those occurring in the middle of an explanation or after a phrase. In the case of discourse markers, the clause to which tags are attached present a falling intonation indicating completeness, and a pause often precedes the discourse marker. This suggests that the tag does not refer to the last clause, but to the entire reasoning. In the case of more elaborate constructions, particularly interrogative clauses, the CCs function as elicitors proper. As observed in section 3.1, very rarely do students take the floor, but when they do, it is precisely after elicitors of this kind, as illustrated in example 3.

(3) S1: [...] we have a machine this machine absorbs some heat then releases another heat then absorbs others and so on continuous and during this process okay? moves some mechanical frames some mechanical system produces heat excuse me produces work some work is negative and then at the end of the story we have the total work W okay? mm? so we have to keep in mind mm? <DRAWING ON THE WHITEBOARD> this eh eh statement **okay? mm? any ob-** yes
S2: <INAUDIBLE>
S1: sorry?
S2: <INAUDIBLE>
S1: yes is the total work W [Lecture 4]

Another use that has to do with interactive metadiscourse is when CCs are employed at the end of passages introducing a new point. The clause or phrase to which the CC is attached normally works as a preview of what follows. Thus the CC is employed to ensure that students understand that a new topic or point will be explained. This is illustrated in examples 4 and 5. It should be noticed that in example 5, in particular, the CC is the only interpersonal item signalling the preview, because despite the interrogative form of the clause, the tone is falling, as in statements.

⁶ Hyland's (2005) model of metadiscourse was developed with special reference to writing. However, the distinction between interactive and interactional metadiscourse also seems useful to highlight interpersonal aspects lecture discourse.

(4) S1: [...] let me look at the value of the total work produced by this thermal machine **mm? ok?** the total work produced by this thermal machine is the following **eh?** <DRAWING ON THE WHITEBOARD> total work WT is equal to is equal to ok W [Lecture 4]

(5) S1: [...] what happens in the case of H2 second source **okay?** i build also for the second source another machine reversible machine R2 operating between S2 and T prime again same T prime [...] [Lecture 4]

In Hyland's model, interactional metadiscourse includes devices that engage the audience, giving them the opportunity "to respond to the unfolding text" (2005: 49), in this case the spoken text. Interactional metadiscourse also enables speakers to anticipate objections and set up "an imagined dialogue with others" (2005: 50). It seems paradoxical to talk about 'an imagined dialogue' in spoken interaction, but as the degree of student participation (Table 2) indicates, this happens in lecture discourse too, and not only in writing. The main interactional function of comprehension checks, as the very label indicates, is to ask students whether they have understood what has been said or understand what is being said. However, we may try to investigate why lecturers feel the need to intrude with a question checking comprehension. Such investigation leads to the identification of different sub functions for CCs.

A first use is to make students focus on a specific concept or idea (examples 6 and 7), and sometimes on both the idea and the words chosen to convey it (example 8). It is as if lecturers were asking students whether they understood the meaning and significance of that particular idea as well as its formulation.

(6) S1: [...] and then this will show_ this shows that eh this idea of limit is an idea which is a general idea **okay?** [Lecture 3]

(7) S1: [...] now we are studying we are proving proofing the Clausius theorem <WRITING ON THE WHITEBOARD> ok? eh? which is another thing is a a new theorem **ok?** [Lecture 4]

(8) S1: [...] eh we will also try to learn to reuse as much as possible, about what is already available eh there is the trend a a a temptation **okay?** to say ok i want to, i need to have one kind of a sensor or something that controls the light [Lecture 1]

A related use is when CCs are employed to make sure that correct information is recorded in the students' mind after repairing errors or after uncertainties. This use is illustrated in example 9.

(9) S1: [...] i made a mistake <WRITING ON THE WHITEBOARD> thank you so much <WRITING ON THE WHITEBOARD> for being so patient <WRITING ON THE WHITEBOARD> you can use an eraser <WRITING ON THE WHITEBOARD> what does we share is the control <WRITING ON THE WHITEBOARD> **okay?** not the the input, it's- <WRITING ON THE WHITEBOARD> **okay?** [Lecture 6]

Numerous instances behave as "progression checks" (Schleef 2005), that is, they are employed to ask if students understand what the lecturer is saying and whether the explanation can continue. This is illustrated in example 10.

(10) S1: [...] in this case we have three different possibilities one is left flip flop feeding me the other one is right flip flop feeding me and the third one is parallel loading feeding me so this multiplexer becomes at three inputs one output, **did you get it?** <P: 04> and the three inputs (xx) bit are from left flip flop, from right flip flop and from parallel load, **got it?** so how many control signals do we need for three-to-one multiplexer? we need, this <POINTING AT DRAWING> is not i need to go to the restroom but this is exactly the answer, two control signals, **got it?** [Lecture 6]

CCs may also occur in contexts in which lecturers ask students to activate their background knowledge to understand a given passage or piece of information, thus creating a “shared discourse environment” (Hansen 1994: 135). This use is also related to instances in which lecturers switch to Italian to ensure the correct comprehension of technical concepts and terms, as illustrated in examples 11, 12 and 13.

- (11) S1: [...] carbon monoxide oh sorry carbon dioxide i beg you pardon mm? *anidride carbonica* **mm**? [Lecture 2]
(12) S1: [...] sulphonamides *sulfamidici* in Italian **mm**? sulphonamides [Lecture 2]
(13) S1: [...] some of these devices are already existing you can use them and some of them are say this home automation technology smart phones mm in some languages we use the term domotics (*sic.*) or *domotica* **mm**? i i won't be able to use it because is not the_ in the (*sic.*) English the word does_ doesn't exist [Lecture 1]

There are some passages in which lecturers check whether the information provided about the organisation of classes, exams or the entire course is clear. In these contexts the check for comprehension seems combined with a “seek for agreement” (Othman 2010), probably because it refers to actions directly involving the students, even though there would be no concrete possibility to question the lecturer's decisions, and indeed this is never done. An example is in 14.

- (14) S1: [...] the oral is rather fast i will ask you a couple of questions here and there in the entire programme generally these questions are something that i cannot put in the written examination something that_ where you can out of some work of your mind explain in more, broad details, but it will last if you are rapid maybe ten minutes it can be as small as this because i have some experience and i understand when you know perfectly things you know so and so or you know ok <LAUGHS> go ahead eh or you really need to spend more time and try once again, **okay**? this is how the exams are held [Lecture 6]

4. Conclusion

This study presented an investigation of the way lecturers ensure understanding not only in terms of discipline-specific content, but also in terms of the lecture as a communicative event. The results confirm that CCs have interactional and interactive metadiscourse functions, engaging students in discourse, opening up a space for intervention, accounting for the audience's processing needs, signalling discourse phases and breaking up the discussion into ‘digestible’ bits. One of the concerns of this study was to verify whether distinctive uses could be noticed in relation to the adoption of English as a Lingua Franca in EMI classes. No significant divergences were observed in terms of the functions of CCs, as the uses identified in studies dealing with native speaker lecture discourse (Schleef 2005; Othman 2010) were also noticed in EMI lectures. CCs are mainly employed as discourse structuring devices, progression checks, elicitors and seeks for agreement. Probably the most notable feature related to the non-native use of English is the form of CCs, which include structures such as *that clear? that fine? you agree?* and *yes or not?* As for the frequency of CCs, it seems that the lecture type and individual proclivities play an important role. However, due to the high incidence of use of CCs in most lectures, i.e. on average almost two per minute, it is also possible to hypothesise that they reflect greater uncertainty on the part of non-native speakers about their effective use of English. Clearly, this issue requires further investigation involving comparisons that take into account the proficiency level of lecturers. It would also be interesting to compare these results to similar ones in other European academic contexts to determine whether the lingua-cultural background of lecturers leaves a mark on their preferred ways of ensuring understanding.

References

- Crawford Camiciottoli, Belinda 2007. *The Language of Business Studies Lectures. A Corpus-assisted Analysis*. Amsterdam: Benjamins.
- DeCarrico, Jeanette and James R. Nattinger, J. 1988. Lexical phrases for the comprehension of academic lectures. In *English for Specific Purposes* 7(2), 91-102.
- Dudley-Evans, Tony and Timothy F. Johns 1981. A team teaching approach to lecture comprehension for overseas students. *The Teaching of Listening Comprehension* (ELT Documents Special). London: The British Council, 30-46.
- Hansen, Christa 1994. Topic identification in lecture discourse. In Flowerdew, John (ed.), *Academic Listening: Research Perspectives*. Cambridge: Cambridge University Press, 131-145.
- Hyland, Ken (2005) *Metadiscourse*. London: Continuum.
- Long, Michael H. 1983. Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics* 4(2), 126-141.
- Nesi, Hilary 2001. A corpus based analysis of academic lectures across disciplines. In Cotterill, Janet and Anne E. Ife (eds), *Language Across Boundaries*. London: Continuum Press, 201-218.
- Othman, Zarina 2010. The use of *okay*, *right* and *yeah* in academic lectures by native speaker lecturers: Their 'anticipated' and 'real' meanings. *Discourse studies* 12(5), 665-681.
- Schleef, Erik 2005. Gender, Power, Discipline and Context: On the Sociolinguistic Variations of *Okay*, *Right*, *Like* and *You Know* in English Academic Discourse. *Proceedings of the Twelfth Annual Symposium about Language and Society Forum* 48, 177-86.
- Simpson, Rita, Sarah Briggs, Janine Ovens, and John Swales 2002. *The Michigan Corpus of Academic Spoken English*. Ann Arbor, MI: The Regents of the University of Michigan.
- Tauroza, Steve and Desmons Allison 1990. Speech rates in British English. *Applied Linguistics* 11(1), 90-105.
- Walsh, Steve 2011. *Exploring Classroom Discourse: Language in Action*. London: Routledge.
- Young, Lynne 1994. University lectures – macro-structure and micro-features. In Flowerdew, John (ed.), *Academic Listening: Research Perspectives*. Cambridge: Cambridge University Press, 159-176.

What I mean is- *what is it doing in conversational interaction?*

Karin Aijmer
University of Gothenburg

1. Introduction

Traditionally there has been a bias towards written language and 'written grammar' in linguistics (see Linell 1982). The balance is now starting to change and linguistic structures are studied both in speech and writing. If we consider how language is used in spoken interaction, we ask other research questions and explain grammatical features from a different perspective than in writing.

In order to describe linguistic structures in spoken language we must consider both 'temporality' and the interactional dimensions of spoken interaction. Spoken language is produced 'from left to right' and speakers are constrained to plan what to say on-line. As a result, they have 'to buy extra time' by using linguistic structures for pausing or word-search. In an interactional approach to grammar, we are interested in highlighting functions or activities that syntactic structures can have in the sequential context rather than in their compositional structure. Elements 'outside the clause' or 'inserts' are a special challenge, since they depend on the preceding discourse and raise expectation about upcoming talk.

According to Biber et al (1999: 1082), inserts 'comprise a class of words that is peripheral, both in the grammar and in the lexicon of the language'. An important group consists of discourse markers. They tend to occur at the beginning of a turn or an utterance where they 'signal a transition in the evolving progress of the conversation' or 'an interactive relationship between speaker, hearer and message' (Biber et al. 1999:1086). Discourse markers can be regarded as an open or emergent category to

which new elements are recruited over time. One such category of novel or non-prototypical kind of discourse markers is represented by *the fact is*, *the thing is*, which ‘look like clauses’ (see Aijmer 2007, Günthner 2011).

There are also structures which need to be analysed differently, depending on whether they are found in spoken or written language. The *wh*-cleft construction has been extensively studied in written language, both formally and functionally. However it has different formal and functional properties in spoken language (Günthner and Hopper 2010 (German and English), Günthner 2011 (German), Pekarek Doehler 2011 (French)).

My aim in this article is to study the construction *what I mean is*, which can be used as or ‘become’ a discourse marker in conversation. The examples I am interested in can be exemplified by (1):

- (1) Sorry, I'm not saying this right. What I mean is, you are absolutely lovely. I can see why Eric fell for you. (COCA)

Biber et al (1999: 1075) consider such expressions as overtures (‘longer expressions from a stock of ready-made utterance openers’). Overtures ‘provide a more explicit way of signaling a new direction in the conversation’ (than discourse markers). However we need a more detailed description of the form and function of such markers.

What I mean is is not a unique phrase. There are similar phrases with other verbs (*what I'm thinking of*, *what I would like to say*) illustrating the same construction. According to Traugott and Trousdale (2013: 144), *do*, *happen* and *say* are the most frequently used verbs in *wh*-clefts. ‘Other verbs, e.g. *want*, *mean* occur with very low frequency’ (Traugott and Trousdale, *ibidem*). I could have dealt with other uses of the *wh*-cleft. However I wanted to focus on a ‘construction’ with a fairly fixed or idiomatic form associated with particular discourse functions.

The COCA Corpus (Corpus of Contemporary American English, <http://corpus.byu.edu/coca/>) was chosen for the empirical analysis because of its large size (about 450 million words). There were 63 examples of *what I mean is* in the spoken part of the corpus. It was also frequent in fiction (153 examples). There is no reason to believe that the phrase is characteristic of American English only. It was also found in the British National Corpus, although there were fewer examples (only 14 examples in the spoken part).

What I mean is in spoken language can best be understood in an interactional or dialogical perspective. It is not used as a main clause, but as ‘a sort of discourse marker’. I will first discuss the structures in which *what I mean is* appears (Section 2). *What I mean is* is described as a collocational frame in Section 3. Section 4 discusses *what I mean is* in the perspective of emergent grammar. In Section 5, I will take an interactional approach to *what I mean is*, looking at how it is used for projective ends. The concluding section 6 will be a summary of the uses of *what I mean is* in interaction and a more general discussion of syntax and ‘spoken grammar’.

2. The use of *what I mean is* in different structures

In the canonical *wh*-cleft, *what I mean is* cannot stand alone, but requires a *that*-clause. In example (2), *what I mean is that* is the first part of a *wh*-cleft construction:

- (2) BOLLING: What do you mean? BECKEL: What I mean is that there are people who make a lot of money when she's out. (COCA)

The *wh*-cleft construction consists of two parts: a matrix clause introduced by *what*, the copula *is* followed by a subordinate clause introduced by *that*. According to Carter and McCarthy (2006: 787), the *wh*-cleft construction has the function of ‘highlighting a whole clause or a longer stretch of discourse instead of focusing on one clause element’.

However, Carter and McCarthy do not discuss the ways in which the *wh*-cleft can be used in new ways formally and functionally in conversation.

Example (3) illustrates the discourse usage of *what I mean is* (cf example 1). *What I mean is* is followed by a sentence (or a discourse segment) which is not integrated in the preceding structure. Similar examples may contain a comma (marking a pause) after *what I mean is*:

- (3) SPRINGER: Well, what I mean is, you meet these guys on -- o -- obviously, they're in jail, they're kind of lonely. Somebody says, Hey, talk to this woman friend.' You're very nice. You're very nice on the phone. They say nice things about you and they -- they may very well mean it at the time they're saying these nice things about you, because, boy, yeah, you're sitting looking at four walls and -- and -- and bars, you would love to be with. (COCA)

The speaker begins with 'well, what I mean is' as a stepping-stone to upcoming talk. The continuation reflects the moment-by-moment unfolding of the speaker's contribution to the conversation.

The disappearance of *that* is a general phenomenon in conversation. This makes it difficult to say if the stretch of talk introduced by *what I mean is* consists of a main clause or if the connection between the clauses is achieved by the sequential placement of the second part only. In (4) there is however no overt sign of subordination following *what I mean is*.

- (4) Ms-ROBEL: I -- I think you're -- I -- what I mean is you're growing in -- in your -- your process, in your journey of... WINFREY: Oh, I know -- I know what -- we know what you mean, yeah. (COCA)

The first part is relatively fixed. The second part has a variable form. In the following sentence, the second part is an interrogative structure which cannot be preceded by *that*. *What I mean is* unambiguously a discourse marker pointing forwards to a question which is relevant against the background of the preceding context:

- (5) "No, no, no, what I mean is, how did you erase it?" # "What?" It seemed I was challenging her technical proficiency. (COCA)

In (6) (an example from fiction), *what I mean is* is followed by an utterance in direct speech ('So the middle child gets screwed again'). Such examples are additional evidence that in conversational interaction speakers avoid complex structures where one clause is subordinate to a main clause. The structural organisation can be understood by considering the cognitive constraints imposed by the necessity to plan and produce spoken language on-line. Speakers confront the problems they encounter in conversational interaction by ignoring the complexity of the embedded structure and producing 'one clause after the other'.

- (6) "So we're only going for one night?" I ask. "And we're camping in the woods?" What I mean is, So the middle child gets screwed again? So number two son is ignored one more time, in a lifetime of getting ignored? So good old Uncle Stevie takes another one for the fucking team?" (COCA)

An element can be inserted between *what I mean* and *is*:

- (7) KOPPEL: And we are back once again. Dr. Hale, I would assume there is a difference between the way you talk to an adult about AIDS and the way you talk to a child about AIDS. What I mean specifically is, do you talk to the children about the fact that they are going to die? (COCA)

By that was especially frequent as an insertion after the verb:

- (8) Mr-BROOKS: Yeah, and specifically what I mean by that is, Is he like me? Does he get what I get? And that's a question of resilience. Why doesn't he talk about fighting? Has he suffered, has he struggled the way I've suffered? This guy seems to have led a charmed life. Why do I think he understands me? And I think that's one of Hillary Clinton's strengths. (COCA)

What I mean is in conversation (unlike the 'ordinary' *wh-cleft*) co-occurs with hesitation phenomena. *What I mean is* is followed by a pause and is preceded by another discourse marker (*listen*):

- (9) Listen, I'm mean - what I mean is- I mean, certainly 1 percent of the country was liberal or something. They would have said Joe McCarthy or McCarthyism- in other words, what I'm saying is- Mr. Winkler, maybe it was a big problem out there at a dinner party. There might have been hysteria there, but among the American people these was the late Truman era and the Eisenhower era and there really was no hysteria in America. (COCA)

Further evidence that *what I mean is* can be exploited as a resource in conversation controlling the speaker's activities is the 'double copula'. In (10) 'double *is*' is found before a *that*-clause:

- (10) When I say save, what I mean is, is that they're gon na know they're gonna be okay. (COCA)

According to Andersen (2002), the double copula occurs systematically (but infrequently) in *wh-cleft* sentences (as well as in clauses such as *the thing is is*). He makes the following comment (Andersen 2002: 51): 'In terms of pragmatic function, these structures [what I mean, what I think/what I'm saying/what I'm suggesting is is (that)] allow the speaker to buy processing time and may thus help in the planning of the utterance'. The double copula is not a performance phenomenon (marking hesitation, stuttering), but supporting evidence for the existence of a discourse marker with functions which are best explained by the temporal constraints imposed by the interaction.

The copula *is* is an obligatory element of the *wh-cleft* construction. However *is* can be missing. Günthner and Hopper (2010:13, 15) discuss such structures at some length. According to them, the *wh-cleft* (with or without the copula) 'contextualizes' a continuation of the first part which unfolds in time and can therefore be of different length. An example of this use of *what I mean is*:

- (11) Mr. LIKHOTAL: Well, I suppose that first of all we should treat it that way. There is a lot of speculation and I believe that in the United States there is an expression that something of like- self-fulfilling prophecy. And what I mean, when you speak about forces being- probability of forces being used in the Lithuania, you add to the anxiety and to the nervousness of different people, and they could be different mistakes in this respect. (COCA)

Even without the copula, *what I mean* seems to function as a discourse marker drawing attention to the following argument (that) you add to the anxiety and nervousness of different people when you speak about forces being used in Lithuania. The absence of the copula does not indicate that the speaker has planning difficulties, but can be regarded as a variant of *what I mean is*.

I have regarded *what I mean* without the copula (or with a double copula) as variant forms of *what I mean is* and not as a performance phenomenon. In (12) *what I mean* without *is* (and together with the attention-getter *look*) points forward to a clause which is 'incrementally produced' in the interaction. *What I mean* is not followed by a

subordinate clause, but the connection depends on the context and the prospective properties of *what I mean*:

- (12) MR-LEHRER: Forget what he says. What did you mean when you said that? Did you mean that Jews -- MR-BUCHANAN: Of course not. MR-LEHRER: All right. MR-BUCHANAN: What I mean, look, if you take the armed forces of the United States, I think the idea that Irish and Hispanic and blacks represent the majority of the ground troops is accurate. It's a good phrase. There's nothing wrong with it. MR-LEHRER: But he's wrong when he claims that was anti-semitic, anti-Jewish? MR-BUCHANAN: Yes, he is. (COCA)

What I mean emphasises that 'it's a good phrase' to say that the Irish, Hispanic and blacks represent the ground troops. (On the other hand referring to people as McCallister or Murphy may be taken as anti-semitic).

In (13) *what I mean* is separated from the discourse it draws attention to by an inserted clause ('it says lower to the floor') reflecting the moment-by-moment unfolding of conversation:

- (13) (Mr-RILENCE: What I mean -- it says lower to the floor. Is there some -- some reason why you need to be closer to the floor because it is a... (Footage-of-Mike-Ri) Mr-RILENCE: The problem is is acclimatizing. That's the reason... AARON: (Voiceover) And at his old high school, it's the same' can do' message. (COCA)

3. *What I mean as as a collocational frame*

From a formal point of view, *what I mean is* is not a main clause, but can be described as a collocational frame (Aijmer 1996:27) or construction (Traugott and Trousdale 2013: 11) (see Figure 1.) It looks like a *wh*-cleft construction, but is employed in new and different ways. It is introduced by *what*; it contains the copula *is* (not *was*) and the subject is the first person. The predicate is often *mean*, although other stance verbs or verbs of saying are also possible. It can be associated with modifications, insertions, reduction.

(No, well, but) what I {mean, want to say...}(by this, specifically) (is) + 'continuation'
--

Figure 1: The collocational frame of *what I mean is*

The continuation consists of an independent sentence (declarative or interrogative). It can also be a lengthy stretch of discourse. The collocational framework captures the flexible nature of *what I mean is*. Some elements are optional and the class of verbs is fairly open. Besides the simple *mean*, we find insertions such as *mean by that* or *mean specifically*.

Language users also have knowledge about the kinds of pragmatic functions associated with the frame or the schema. The functions of *what I mean is* rely on the projective force of the construction. The importance of projection is further discussed in Section 5.

To sum up, in this section I have questioned the analysis of *what I mean is* in spoken language as a matrix clause which is followed by a complement clause introduced by *that*. In spoken language, *what I mean is* can be regarded as a discourse marker and the following clause is no longer integrated with the main clause.

4. *What I mean is and emergent grammar*

In an interactional approach, structures are not fixed but flexible or 'emergent' in the communication situation (on emergent structures, see Hopper 1987). As we have seen, the clause following *what I mean is* is not treated by speakers as a subordinate clause, but as an independent sentence (or, more generally, as a discourse segment). *What I mean is*,

on the other hand, can be regarded as a discourse marker whose subsidiary status is indicated by its position outside the sentence or clause.

The emergence of new uses of the *wh*-cleft in the conversation and their ‘routinization’ is also compatible with grammaticalization or ‘constructionalization’ (‘the creation of a form_{new}-meaning_{new} pairing; Traugott and Trousdale 2013). *What I mean is* as a discourse marker can be regarded as the end-point of a process by means of which the main clause *what I mean is* is ripped out of the sentence and reanalysed as a discourse marker in an outside-the-clause position. It is used as a linguistic resource for performing pragmatic functions associated with the structuring of context (‘the pragmatics of indexing upcoming discourse’ Traugott and Trousdale 2013: 145).

5. *What I mean is* and projective force

The interactional approach implies that the explanation for how *what I mean is* is used should be in terms of social action and cognition. ‘[W]e should always prefer explanations in terms of cognitive processing and/or social action and interaction (e.g. constraints on turn design conditioned by sequential contexts), and let purely structural (language-internal) accounts only stand proxy for such explanations’ (Linell 2008: 97-98).

What I mean is and similar expressions or constructions can be explained by the sequential organization of talk. We expect constructions to be designed to either respond to what has been said or to set up conditions on subsequent sequences of talk. According to Günthner (2011: 19), ‘in everyday talk, speakers make use of various types of *projective constructions*. *What I mean is* seems to be a good example of a marker which has been designed to function in the interaction as a ‘projector phrase’ raising expectations about a continuation’ (Hopper and Thompson 2008, Günthner and Hopper 2010, Günthner 2011)¹. The use of a projector phrase (and its continuation) fulfils important roles in the communication. ‘[Projector phrases] may provide a cognitive breathing space for formulating the next utterance in a maximally efficient way’ (Günthner 2011: 19). *What I mean is* emphasizes a subjective point of view (the speaker’s orientation to the discourse sequencing) and functions as a turn-holder. The organization into a projecting part and a projected part also has interactional advantages. Projector phrases are convenient for ‘pragmatic framing’ (the projective phrase or the frame is separated from the framed part; Günthner and Hopper 2010: 20). After *what I mean is*, we can, for example, expect an explanation or clarification of different length and complexity.

What I mean is can have scope over a lengthy discourse segment. This is illustrated in (14), where the speaker argues that Shuster’s report may ‘have gotten a little bit ahead of the facts’. The speaker presents his points over several utterances reflecting the temporal unfolding of talk in the talk:

- (14) David Shuster’s reports have not been inaccurate. He’s been on the cutting edge.
MCMAHON: I’m not disparaging David Shuster at all. But what I mean is, sometimes these reports have gotten a little bit ahead of the facts. If David Shuster’s report turned out to be accurate, I would agree with you that the president has a very serious problem. Twelve counts of obstruction, or something in that neighbourhood...
HANNITY: Perjury. (COCA)

¹ Compare also Biber et al’s term ‘overture’ and Stubbs’ (1983) preface. Stubbs (1983: 183) defines ‘preface’ as follows: ‘[s]peakers use such items to indicate to hearers that they should not attempt to use placement in sequence, in order to analyse the point or illocutionary force of the following utterance.’

But what I mean is foreshadows some opposition from the hearer and has the function to keep the hearer's attention until the long and complex argument has been completed (cf Günthner 2011: 14, for similar examples with N be *that* – constructions in German conversation).

6. Conclusion

In order to understand linguistic structures, we need to consider how they are used in conversation by real people. Things 'happen to' linguistic structures when they are used in conversation. This can be explained by the constraints imposed by turn-taking or the needs to plan and produce talk on-line. The form and functions of the *wh*-cleft, for example, can only be understood on the basis of the sequential organization of talk. In this paper I have focused on a particular 'fixed phrase', *what I mean is*, and its use in spoken language.

What I mean is (that) 'looks like grammar' since it has syntactic structure. However it is no longer integrated in a sentence, but occupies a slot outside the sentence or utterance, which is relevant for spoken language and for interaction. *What I mean is* is not a main clause, but can be regarded as a sort of discourse marker with discourse-structuring functions.

What I mean is has been shown to have the following properties:

- *what I mean is* is an 'insert' (an unattached element) with projective force
- It is a flexible structure which can be exemplified in different forms (other verbs than *mean* can be used, the copula can be doubled or omitted, things can be inserted between the verb and the continuation)
- It has developed out of a complex sentence on the basis of the speaker's recurrent use of the structure in different contexts 'an emergent product of interactional practices' (Günthner 2011: 29).
- It can be described as a collocational frame (grammatical construction or schema)
- It has projective properties, that is it raises expectations about a continuation. As indicated by the collocational frame, the continuation is a part of the construction
- *what I mean is* is used as a pragmatic frame at the boundary of a conversational exchange, marking upcoming talk as a specification or clarification. It can also have a floor-holding function, making it possible for the speaker to plan ahead what to say.

References

- Aijmer, Karin 1996. *Conversational routines in English. Convention and creativity*. London: Longman.
- Aijmer, Karin 2007. The interface between discourse and grammar. In Celle, Agnès and Ruth Huart (eds). *Connectives as discourse landmarks*. Amsterdam and Philadelphia: John Benjamins, 31-46.
- Andersen, Gisle 2002. Corpora and the double copula. In Breivik, Leiv Egil and Angela Hasselgren (eds.). *From the COLT's mouth ... and others'*. *Language corpora studies in honour of Anna-Brita Stenström*. Amsterdam and New York: Rodopi, 43-58.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, and Edward Finegan 1999. *The Longman Grammar of Spoken and Written English*. London: Longman.
- Carter, Ronald and Michael McCarthy 2006. *Cambridge Grammar of English. A Comprehensive Guide. Spoken and Written English. Grammar and Usage*. Cambridge: Cambridge University Press.
- Günthner, Susanne 2011. N be *that*-constructions in everyday German conversation: A reanalysis of 'die Sache is/das Ding ist' ('the thing is')-clauses as *projector phrases*. In Laury, Ritva and Ryoko Suzuki (eds.), *Subordination in conversation. A cross-linguistic perspective*. Amsterdam and Philadelphia: John Benjamins, 11-36.
- Günthner, Susanne and Paul J. Hopper 2010. Zeitlichkeit & sprachliche Strukturen: Pseudoclefts im Englischen und Deutschen. *Gesprächsforschung - Online Zeitschrift zur verbalen Interaktion*. Ausgabe, 11. 1-28. (www.gespraechsforschung-ozs.de)
- Hopper, Paul J. 1987. Emergent grammar. In Aske, Jon, Natasha Berry, Laura Michaelis, and Hana Filip (eds.), *Berkeley Linguistics Society 13: General Session and Parasession on Grammar and Cognition*. Berkeley, CA: Berkeley Linguistics Society, 139-157.

- Hopper, Paul J. and Sandra A. Thompson 2008. Projectability and clause combining in interaction. In Laury, Ritva (ed.), *Cross-linguistic studies of clause combining. The multifunctionality of conjunctions*. Amsterdam and Philadelphia: John Benjamins, 99-123.
- Linell, Per 1982. *The written language bias in linguistics*. Linköping: Department of Communication Studies.
- Linell, Per 2008. Grammatical constructions in dialogue. In Bergs, Alexander and Gabriele Diewald (eds.), *Context and Constructions*. Amsterdam: John Benjamins, 97-110.
- Pekarek Doehler, Simona 2011. Clause-combining and the sequencing actions: Projector constructions in French talk-in-interaction. In Laury, Ritva and Ryoko Suzuki (eds.), *Subordination in conversation. A cross-linguistic perspective*. Amsterdam and Philadelphia: John Benjamins, 103-148.
- Stubbs, Michael 1983. *Discourse analysis. The sociolinguistic analysis of natural language*. Chicago: University of Chicago Press.
- Traugott, Ekizabeh C. and Graeme Trousdale 2013. *Constructionalization and constructional changes*. Oxford: Oxford University Press.

Quality Measurements of Error Annotation - Ensuring Validity through Reliability

Nikola Dobrić

Alpen-Adria-Universität Klagenfurt

Major obstacles to achieving high levels of reliability (and by extension validity) of error annotation of learner corpora range from defining errors in general, the lack of an error taxonomy sufficiently applicable in corpus annotation, insufficiencies of any set linguistic norm as background for tagging, to the lack of well-defined measurements of quality of annotation. The paper first looks at the theoretical issues behind the definition of an error. It expands the discussion by focusing on a more practically applicable account of errors aimed at error annotation. It goes on to offer a more robust error taxonomy which could help address issues of interpretability inherent in linguistic categorization and could ensure more consistency. In the end, the paper suggests an alternative definition of an error applicable for corpus annotation, based on inter-annotator agreement and aimed at being the primary indicator of validity.

1. Introduction and background

Error annotation of learner corpora¹ is problematic when it comes to defining errors in practical terms, in terms of the error classification chosen as the background for a tagset, in terms of the annotators' training to use a set linguistic norm, and in terms of assessment of the quality (or 'correctness') of annotation. Assessing the quality of annotation is essential for gauging the validity of the linguistic information we wish to extract from the corpus. Since validity cannot be assessed directly, due to the lack of 'ground truth' in human (linguistic) judgment, the only thing we can assess is the reliability of annotation as indicating validity of the tags assigned (Plaban, Pabitra and Anupam 2000). The reliability is mirrored in consistency of annotation (Brants 2000), done both by each individual annotator (intra-annotator agreement) and by more annotators when compared to each other (inter-annotator agreement). High levels of consistency signify high quality of data processing. It is hence important to ensure high levels of consistency of tagging, which would lead to high reliability of annotation and hence signify valid information in the corpus.

¹ For more on the construction and uses of learner corpora see Granger (1998), Tono (2003) or Callies and Zaytseva (2013).

Starting with the issue of accounting for errors, there is a general consensus that errors constitute failures in language competence (Corder 1971; Lennon 1991; Lengo 1995). From the point of view of a practical application in any form of error analysis, as in corpus annotation, such a take on errors is not really useful. Firstly, following such a theoretical premise in practice would lead us to a conclusion (reached already by other authors, such as for example James (1998:79)) that errors cannot be produced by 'native' speakers. Only mistakes could be possible.² Secondly, language competence is in broad terms accessible only indirectly through language performance and needs to be observed accordingly. In terms of error annotation (and applied linguistics in general), it is clear we need a more practice-oriented definition of an error. In other words, we need something more tangible that annotators can actually hold on to while tagging.

A common approach to solving this problem is to define errors using a set linguistic norm as a background. The opaque (and often unjustly negatively connoted) term 'error' is then (much more accurately) referred to as a 'non-norm adequate form'. In practice, the linguistic norm can be tied to a standard variety of a language, its grammatical description, dictionaries, and in case of learner corpus annotation, to the training of annotators. This is a more applicable take on errors – there is a common linguistic background set for all annotators. It can clearly indicate to annotators what the acceptable, norm-adequate performance is, at least up to a certain hierarchical level of language description. It makes their determining what is not adequate simpler and ideally produces more agreement between their categorizations.

2. Error annotation in practice

The next stumbling block on the road towards higher consistency of learner corpus annotation is the choice of a taxonomy of non-norm adequate features to be used as a tagset. This task is, of course, also riddled with difficulty. If we look at what has been done so far in terms of enumerative lists of non-norm adequate features, we can see several different approaches:

(1) classification according to level of linguistic description: this most commonly applied taxonomy of errors employs the various levels of linguistic analysis (phonology, morphology, syntax, semantics, etc.) as the basis for defining error types (Ellis and Barkhuizen 2005; George 1972; Havranek 2002). Such taxonomies identify errors like 'passive voice', 'temporal conjunction', 'transitive verb' or 'wrong word';

(2) classification according to alterations in ideal performance: this less popular and more abstract type of taxonomy describes errors in terms of what has been altered on the 'surface' level of a hypothetical ideal performance. This includes *omissions* (some element demanded by the norm is left out), *additions* (some element barred by the norm is added), *misinformation* (some element is expressed by a form barred by the norm), and *misordering* (some element demanded by the norm has been misordered) (Dulay, Burt and Krashen 1982: 150);

(3) classification combining levels of linguistic description and alterations in ideal performance: this approach describes each error both with regard to level of linguistic analysis and in terms of alteration in the hypothetical ideal performance. It yields error categories such as *tense/omission* or *modal verbs/misordering* (Pibal 2012:10).

² There is a general consensus that errors differ significantly from mistakes. The difference lies in terms of corrigibility (James 1998): the claim is that a speaker would be able to self-correct a mistake (some authors hence terming them slips (Edge 1989), and attribute mistakes to problems of cognitive activation or sheer carelessness. Errors, on the other hand, would be beyond self-correcting and are to be seen as failures in language competence.

(4) classification in terms of presumptive cause of error: this approach attempts to describe errors with regard to the possible source of the error. This may be the learner's L1 or another foreign language or universal cognitive constraints. Error categories in such taxonomies are *interlingual errors* (attributable to interference), *developmental errors* (due to universal cognitive constraints), *ambiguous errors* (attributable to more than one possible source), and *unique errors* (a residue category for unclassifiable errors) (Dulay, Burt and Krashen 1982: 163);

(5) classification according to the degree of message impairment: this approach describes errors in terms of the degree to which they disturb the message in information theory terms. Errors are here characterized with regard to their effect on the listeners or readers. A distinction is often made between *global errors* and *local errors*. *Global errors* involve large amounts of noise and seriously impair comprehensibility. An example would be violations of major syntactic rules. *Local errors* are said to cause noise to a lesser degree and involve a narrower focus. Examples are errors in article use or verb inflections (Dulay, Burt and Krashen 1982: 172).

However, attempting to apply these types of taxonomies (and their various practical incarnations) to authentic learner language data has shown that they sometimes leave room for a considerable amount of subjective judgment. This can render the training of future corpus annotators difficult and have serious impact on consistency and reliability. Also, reports on the applications of any of the listed types of taxonomies in learner corpus annotation are hard to find. Such a state of affairs has prompted a development of an error classification which could leave less room for subjective judgment, prove less demanding in annotator training, and produce high levels of inter-annotator and intra-annotator agreement:

(6) the classification according to *substance* and *scope* (the *SD Error Taxonomy*): this approach to defining errors revolves around two questions: how can you tell something is an error when you look at a given language performance in terms of the context involved; and what is it exactly that creates an error in terms of deviation from the norm? If we term the textual and extra-textual context necessary to detect an error SCOPE and the non-norm adequate form to be amended an error SUBSTANCE, we can get 14 classes of errors (Dobrić and Sigott 2014):

- SCOPE phrase SUBSTANCE phrase;
- SCOPE phrase SUBSTANCE punctuation;
- SCOPE clause SUBSTANCE phrase;
- SCOPE clause SUBSTANCE clause;
- SCOPE clause SUBSTANCE punctuation;
- SCOPE sentence SUBSTANCE phrase;
- SCOPE sentence SUBSTANCE clause;
- SCOPE sentence SUBSTANCE sentence;
- SCOPE sentence SUBSTANCE punctuation;
- SCOPE text SUBSTANCE phrase;
- SCOPE text SUBSTANCE clause;
- SCOPE text SUBSTANCE sentence;
- SCOPE text SUBSTANCE text; and
- SCOPE text SUBSTANCE punctuation.

We can first see that the scope of an error may be *phrase*, *clause*, *sentence* or *text*.³ The substance of an error covers the same denominations, with the addition of *punctuation*. To illustrate the process we can examine two examples.

[Example 1] At evening I and my best friends always speak at our problems and other things.

[Example 2] Our teacher expects us to know a lot about the historical background of the covered literary epoch.

If we look at the first example above, for instance, a fully proficient speaker of English would most probably recognize the error of ‘At evening’ immediately. However, if we ask ourselves how we realize this is an error, it becomes apparent that we can only do so when we look beyond the individual words themselves. While ‘at’ and ‘evening’ are both perfectly acceptable as lexical items of English, the combination ‘At evening’ violates the norm. Hence, the error only becomes manifest when one takes into consideration the context beyond the individual word. In cases like these, we will say that the scope of the error is the phrase, as the phrase is the necessary context we need to recognize the error in. Such an explanation can be constructed and exemplified for all of the mentioned hierarchically higher levels of scope (namely *clause*, *sentence*, and *text*) as well. When we look at the second example, a brief scrutiny of the noun phrase ‘the covered literary epoch’ is enough for us to become aware of an error, which consists in ‘covered’ being used in premodifying rather than in postmodifying position. The phrase structure in the writing performance will need to be changed to a structure in which the head is followed by a postmodification rather than preceded by a premodification. So a change at phrase level is required in order to remove the error. Here both scope and substance of the error will therefore be said to be *phrase*, because it is the noun phrase (or more precisely its structure) that needs to be altered. Similar examples can be drawn for other types of error substance (*clause*, *sentence*, and *text*).

The next problem we encounter in our efforts to ensure consistency of error annotation is that, even when using a coarsely-grained taxonomy specifically geared towards learner corpus annotation, there are general linguistic issues that cannot be resolved easily through any kind of training or classification attempt. The problem that is always encountered in any form of linguistic categorization (and is a longstanding one in applied linguistics) is that there is no ultimate correct version of language performance at most levels of language. Standardizing a linguistic norm through training to serve as a background for subsequent identification of non-norm adequate features only takes us as far as any grammatical description of language can. In a clear-cut and well-defined manner, this means up to certain hierarchical levels of language description, such as morphology and perhaps syntax. We can, for instance, unambiguously recognize something as a non-norm adequate spelling or a non-norm adequate tense form. Any grammatical account of language viewed as a norm has little to offer in semantic and pragmatic (discourse) terms. Instead, at this level of language use, we have theoretically an infinite possibility in terms of acceptable and adequate performance by which to express a particular message. Relaying, for example, the message of ‘opening’ a window can pragmatically take multiple norm-adequate forms, such as ‘Please open the window.’, ‘Open the window, please.’, ‘Could you open the window, please?’, ‘Oh, it is really hot in here!’, and more. Each of these adequate forms of relaying the same message would potentially highlight different parts of the same performance as being non-norm adequate.

³ In difference from the full *SD Error Taxonomy* which includes *word* as both possible substance and scope of an error (as seen in more general terms in Dobrić and Sigott (2014)), the version presented here has been shortened to make it more coarsely-grained and hence more applicable as a tagset. In essence, the category of *word* was, in accordance with common linguistic description, conflated with the category of *phrase*.

Each of these adequate forms would also potentially appear as the ‘correct’ one to different annotators, again increasing the probability of a different error classification in the corpus. Hence, what happens in corpus annotation is that when faced with, for example, a sentence in a writing performance, the annotators can compare it (as a whole and as individual constituents) directly to a predefined norm only in terms of morphology and/or syntax. The semantic and pragmatic content of a performance can only be compared to the hypothetical internalized version(s) of how the given sentence should be formed. This ideal version stems from an idea of the performance a proficient (‘native’) speaker of the language at hand would produce, which is in practical terms a rather illusive concept. It is hard to tie this level of language to anything tangible we could use in training annotators. This realization means that even if we define an error as a non-norm adequate form, pre-define and train the said norm, and then constrain the classification of errors within a specially defined list of possible error types, we are still not close to referring to an error in an unambiguous manner. This is because any given linguistic norm is for the most part intrinsically ambiguous. This state of affairs has a significant impact on the quality of annotation (reliability) we can expect.

3. Error as agreement

With a view to gauging the levels of consistency and the quality (correctness) of each of the errors tagged in the corpus (particularly when annotated by multiple annotators), there are two possible ways of observing an error:

- setting up one ‘expert’ version of the annotation as the ‘correct’ one; or
- seeing a ‘correctly’ tagged error through annotator agreement.

The former option would entail ascertaining the quality of the tagging by comparing it (most commonly) to a version that corpus creators would put forward as the ‘right’ one. This is useful for conducting inter-annotator agreement measurements. It would also mean, however, presuming that the ‘expert’ version does not suffer from the same problems of ambiguity and subjectivity as any other possible classification done by any of the annotators. We would in essence be saying that such a version is more ‘correct’ than a number of other equally acceptable versions annotators may come up with. For example, a sentence such as ‘At home he saw his dog.’ could be marked by annotators as having a *SCOPE Sentence SUBSTANCE Sentence* error (‘seeing ‘At home’ as unjustifiably being the point of emphasis) and not marked at all by the ‘expert’ annotation. Should it really be considered as ‘incorrect’ annotation and signify a lack of quality of processing?

The latter and more acceptable option entails defining an error as a point of mutual agreement between trained annotators working on the same corpus project. This in actuality means that, if most annotators processing the same corpus data feel that a certain part of a language performance is not adequate in terms of the linguistic norm they have been trained to follow and if they further agree in which terms (referring here to the location and the classification of an error they have also been trained in applying), we can assume that the given part is marked ‘correctly’. Hence, high agreement in annotation signifies good quality of the classification. In simple terms, the correct version is, for all intents and purposes, defined as agreement between annotators in recognizing parts of language performance as being errors or not.

Viewing ‘correctly’ tagged errors in corpus annotation against the background of high levels of inter-annotator agreement has an impact on the quality control measurements in (learner) corpus annotation. This in turn has an effect on the validity of the information which is to be extracted from the corpus subsequently. By addressing the

lack of a fully discriminate linguistic norm through redefining it as agreement among annotators, we have provided a solid, criteria-based way of ascertaining levels of quality or 'correctness' of the tags assigned. Being able to clearly state the levels of quality of the manual error tagging in a learner corpus, via for example a *Kappa* agreement coefficient (or some more advanced statistical measurement), would help researchers understand the limitations of the corpus data at hand and interpret their results accordingly.

Further work on ensuring more validity of (learner) corpus annotation should focus first on improving general statistical measurements of agreement in linguistic categorization. One problem is that many agreement measurement models (such as the ones described in e.g. Carletta (1996); Di Eugenio and Glass (2004); or Ragheb and Dickinson (2013)) measure agreement with reference to a final correct classification, a model we have discarded as linguistically unacceptable. Most agreement measurements also focus only on pairwise comparison of annotators, whereas large corpora are often manually annotated by multiple judges. Another problem to be solved is which of the existing statistical measurements to use in order to measure inter-annotator agreement. The most commonly used agreement measurement is the *Kappa* coefficient of agreement. The *Kappa* statistic is a well-established measurement focusing on the observed agreement among the coders (for a detailed description see Cohen (1968); Fleiss (1971); Fleiss (1981); or Di Eugenio (2000)). However there are numerous other, more or less advanced, measurements of agreement including raw agreement, then different kinds of *Kappa* (*Cohen's*, *Scott's/Fleiss'*, *Conger's*, *Light's*), then *marginal homogeneity*, *tetrachoric correlation*, *Krippendorff's Alpha*, *McNemar Test*, *Gwet's ACI*, and more. There is much work in progress attempting to streamline the statistical approaches to agreement for (error) annotation of learner corpora (see for instance Sigott, Cesnik and Dobrić (forthcoming)).

References

- Brants, Thorsten 2000. Inter-Annotator Agreement for a German Newspaper Corpus. In Gavrilidou, Maria, George Carayannis, Stella Markantonatou, Stelios Piperidis & Gregory Stainhauer (eds.), *LREC2000: Proceedings of the Second International Conference on Language Resources and Evaluation*, Athens, 165-172.
- Callies, Marcus, Ekaterina Zaytseva 2013. The Corpus of Academic Learner English (CALE) - A new resource for the assessment of writing proficiency in the academic register. *Dutch Journal of Applied Linguistics*, 2/1, 126-132.
- Carletta, Jean 1996. Assessing Agreement on Classification Tasks: The Kappa statistic. *Computational Linguistics*, 22/2, 249-254.
- Cohen, Jacob 1968. Weighted Kappa: Nominal Scale Agreement with Provision for Scaled Disagreement or Partial Credit. *Psychological Bulletin*, 70, 213-20.
- Corder, Pit 1971. Idiosyncratic Dialects and Error Analysis. *International Review of Applied Linguistics*, 9, 149-59.
- Di Eugenio, Barbara 2000. On the Usage of Kappa to Evaluate Agreement on Coding Tasks. In Gavrilidou Maria et al. (eds.) *LREC2000: Proceedings of the Second International Conference on Language Resources and Evaluation*, Athens, 441-444.
- Di Eugenio, Barbara, Michael Glass 2004. The Kappa statistic: a second look. *Computational Linguistics*, 30/1, 95-101.
- Dobrić, Nikola, Günther Sigott 2014. Towards an Error Taxonomy for Student Writing. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 19/2, 111-118.
- Dulay, Heidi, Marina Burt, Stephen Krashen 1982. *Language Two*. New York: Oxford University Press.
- Edge, Julian 1989. *Mistakes and Correction*. London: Longman.
- Ellis, Rod, Gary Barkhuizen 2005. *Analyzing Learner Language*. Oxford: Oxford University Press.
- Fleiss, Joseph 1971. Measuring Nominal Scale Agreement among Many Raters. *Psychological Bulletin*, 76/5, 378-382.
- Fleiss, Joseph 1981. *Statistical Methods for Rates and Proportions*. New York: John Wiley & Sons, Inc.
- George, Vernon 1972. *Common Errors in Language Learning: Insights from English*. Rowley, MA: Newbury House.
- Granger, Sylviane 1998. Computer-aided Error Analysis. *System*, 26, 163-174.

- Havranek, Gertraud 2002. *Die Rolle der Korrektur beim Fremdsprachenlernen*. Frankfurt/Main: Lang.
- James, Carl 1998. *Errors in Language Learning and Use: Exploring Error Analysis*. London: Longman.
- Lengo, Nsakala 1995. What is an error? *English Teaching Forum*, 33/3, 20-24.
- Lennon, Paul 1991. Error: Some Problems of Definition, Identification and Distinction. *Applied Linguistics*, 12, 180-96.
- Pibal, Florian 2012. *Identifying Errors in the Written Manifestations of Austrian English Learner Language at 8th-Grade Secondary Level and their Influence on Human Ratings*. MA Thesis, Alpen-Adria-Universität Klagenfurt.
- Plaban, Bhowmick, Mitra Pabitra, Basu Anupam 2008. An agreement measure for determining inter-annotator reliability of human judgements on affective text. In Artstein, Ron, Gemma Boleda, Frank Keller and Sabine Schulte im Walde (eds.), *Proceedings of the HumanJudge '08 – the Workshop on Human Judgements in Computational Linguistics*, Stroudsburg: Association for Computational Linguistics, 58-65.
- Ragheb, Marwa, Markus Dickinson 2013. Inter-annotator Agreement for Dependency Annotation of Learner Language. In Tetreault, Joel, Jill Burstein and Claudia Leacock (eds.), *Proceedings of the Eighth Workshop on Innovative Use of NLP for Building Educational Applications*, Atlanta: The Association for Computational Linguistics, 11-12.
- Sigott, Günther, Hermann Cesnik, Nikola Dobrić, forthcoming. Measurements of Consistency in Annotation of Learner Corpora - SD Error Taxonomy and Agreement.
- Tono, Yukio 2003. Learner corpora: Design, development and applications. In Archer, Dawn, Paul Rayson, Andrew Wilson and Tony McEnery (eds.), *Proceedings of Corpus Linguistics 2003*, Lancaster University: UCREL Technical Papers, 800-809.

REVISITING LITERATURE

From Theories of Deviation to Theories of Fictionality: The Definition of Literature

Rolf Breuer

University of Paderborn

This essay focuses on the various ways in which literature has been differentiated from non-literature. The criteria of differentiation show themselves to be quite heterogeneous, even incommensurable. Older – essentialist – theories, based on epic and lyric poetry, distinguished between poetic and non-poetic forms of language. Later – relational – theories, often based on the novel, have argued that it is the reference of language to reality that distinguishes fiction from non-fiction. Still more recent theories, accompanied by new forms of literature, see the difference in the eye of the beholder or, rather, reader – and this is a pragmatic criterion of differentiation. Since each perspective yields valuable insights, the question is how the three criteria – essentialist, relational and pragmatic – relate to one another.

1. Language and form as the distinguishing criterion of poetry

Until the 18th century, epic poetry and certain types of poems - ode, elegy, nature poetry - were the key genres of literature. A poet striving for honour and glory had to excel in these genres, in which Homer and Vergil were considered to be the greatest models, with Dante, Camões, Milton and others as the respective national examples.

These epic poems, odes, sonnets, and epistles differed from discursive texts – historiography, homilies, philosophical treatises, laws, or everyday speech, etc. – in their use of language, namely in such deviations from everyday speech as verse, metre, rhyme, poetic diction with liberties in vocabulary (archaisms, for instance) and syntax (a freer order of words).