

# REPRESENTING CULTURE IN OPENCOURSEWARE LECTURES

## A corpus-based semantic analysis

BELINDA CRAWFORD CAMICIOTTOLI  
UNIVERSITY OF PISA

**Abstract** – This paper explores the representation of culture in a corpus based on fifteen OpenCourseWare (OCW) lecture transcripts available from the Massachusetts Institute of Technology and Yale University. The corpus is divided into three components of five lectures each from the broad disciplinary areas of humanities, soft sciences, and hard sciences, thus allowing for a contrastive analysis of cultural references across different domains of specialized knowledge. The analytical approach is based on the concept of a “cultureme”, i.e., a unit of analysis for a culture-specific phenomenon and its linguistic expression. The corpus was processed with the semantic annotation tool of Wmatrix that automatically assigns lexical items in a corpus to pre-established semantic domains. The lexical items in domains associated with human cultural experiences (e.g., education, religion, history, food and drinks, sports, the media, entertainment, geographical names, proper names) were then examined to identify culturemes. Extensive follow-up cross-domain analysis was necessary to tease out culture-specific meanings across the corpus. The results indicate that, although not present in high frequencies, a number of culturemes were dispersed throughout the corpus, with most referring to the domain of education, followed by government & politics and entertainment. The paper concludes with some reflections on the pedagogical implications of the findings in the context of helping L2 learners cope with the comprehension challenges of culture-specific meanings in lecture discourse.

**Keywords:** lectures; OpenCourseWare; culture; culturemes; semantic annotation.

## 1. Introduction

The university lecture has often been criticized as an outdated and passive learning approach which is not sufficiently interactive for today’s modern students (see, among others, Limbach, Waugh 2005; DiPiro 2009). Yet it is still the predominant teaching format in higher education institutions worldwide. This is likely due to the fact that the lecture remains the most practical and popular way to teach the large classes of students that are frequent in most universities (Bligh 2000; Parini 2004; Clay, Breslow 2006).

However, in recent years, the traditional university lecture has been transformed by technology which offers new delivery formats and modes that increase learning opportunities by overcoming the physical, temporal, and economic constraints of attending university classrooms. For example, video recorded classroom lectures are widely available through OpenCourseWare (OCW), which was pioneered by the Massachusetts Institute of Technology (MIT) in 2001 to promote an educational vision that values learning over profit (Lerman, Miyagawa 2002). According to the MIT OCW website,<sup>1</sup> the

<sup>1</sup> <https://ocw.mit.edu/about/presidents-message/>

mission of MIT is “to advance knowledge and educate students, and to bring knowledge to bear on the world's great challenges for the betterment of humankind. Open sharing of knowledge is the purest manifestation of this”. This vision thus reflects a new paradigm that strives to provide a high quality educational experience to anyone who wishes to learn at no cost, which is in stark contrast to competitive and market-based models often found in higher education today.

The OCW movement has grown dramatically and now offers a wide array of learning resources, including audio/video recorded lectures, which constitute a core component of OCW (Vladoiu 2011). It is now possible to find literally thousands of video recorded lectures from universities all over the world. An important source of video lectures is the Open Education Consortium portal with a searchable database of courses and materials of all kinds.<sup>2</sup> Large numbers of university lectures are also freely available on iTunes U, while many universities operate their own OCW websites.

What learners can access in OCW learning environments varies among hosting platforms, but they can often find complete courses that contain a range of different types of materials, including syllabi, resource lists, outlines of lecture topics covered in each session and corresponding notes, audio and/or video recorded lectures, transcripts of recorded lectures, links to access free materials or purchase textbooks, and post-course student feedback surveys. Annex 1 is a screenshot of the web page that corresponds to Lecture 1 of the Open Yale University Course History 116 “The American Revolution”.<sup>3</sup> As can be seen, there is a paragraph-length overview of the lecture, the lecture video that can be accessed directly or downloaded in two different bandwidths, the audio file, and the lecture transcript with the content articulated in “lecture chapters” to facilitate navigation.

Although OCW courses offer a variety of learning resources, because they are non-credit bearing, there is no formal assessment or grading procedure. Within the Open Educational Resources (OER) landscape,<sup>4</sup> OCW can be distinguished from MOOCs (Massive Open Online Courses) on a number of levels. According to Martinez (2014), MOOCs are usually offered by educational companies (e.g. Coursera, Udacity). They are accessible only during the course itself, include assessment, offer credits, and have a collaborative format for participants. OCW is instead usually provided by higher education institutions and can be characterized as static, always accessible, without assessment, without accreditation, and designed for an individual learning experience.

One of the important features of OCW lectures is that their online delivery opens them up to a vast audience of learners, including those with different linguistic and cultural backgrounds. Thanks to these increasingly available digital resources, L2 learners can now reap the benefits of lectures delivered in English by professors in prestigious universities, thus acquiring specialized knowledge across a variety of academic disciplines. However, it is well known that lecture discourse may contain specific cultural references that can be challenging for listeners of different linguistic and cultural backgrounds (Flowerdew 1994; Mulligan, Kirkpatrick, 2000; Zhu, Flaitz 2005; Crawford Camiciottoli 2005). To this regard, Miller (2002, p. 147) articulated four different aspects

<sup>2</sup> <http://www.oeconsortium.org/>

<sup>3</sup> Joanne Freeman, Introduction: Freeman’s Top Five Tips for Studying the Revolution (Yale University: Open Yale Courses), <http://oyc.yale.edu> (Accessed December 14, 2017). License: Creative Commons BY-NC-SA.

<sup>4</sup> The Hewlett Foundation defines OER as “teaching, learning and research materials in any medium – digital or otherwise – that reside in the public domain or have been released under an open license that permits no-cost access, use, adaptation and redistribution by others with no or limited restrictions. (<https://www.hewlett.org/strategy/open-educational-resources/>)

of culture that may emerge during a lecture and create comprehension difficulties for L2 students:

1. ethnic culture: aspects deriving from a mismatch between the cultural backgrounds of the lecturer and the L2 students;
2. local culture: aspects of the lecture linked to the local setting which may be unknown to L2 students;
3. academic culture: different practices in educational institutions with which L2 students have little experience;
4. disciplinary culture: discipline-specific ways of presenting knowledge that L2 students may not know.

Clearly, these cultural aspects can have an impact on whether or not L2 learners are able to successfully understand lectures in English. This is an issue that takes on special importance considering the growing numbers of international students pursuing their higher education in contexts where English is the language of instruction, reflecting a process of internationalization within a globalized academic community (Coleman 2006). Indeed, international student enrollments in universities located in English-speaking countries are constantly on the rise.<sup>5</sup> Moreover, OCW lectures now provide even more opportunities for learners of different language backgrounds to experience high quality lectures delivered in English from anywhere in the world. Yet, regardless of the traditional classroom or ‘virtual’ OCW delivery format of the lecture, L2 learners all face the same challenges in understanding cultural meanings in lectures. For a better understanding of the role of culture-specific references in lecture discourse, I investigate the representation of culture in a corpus of OCW lectures. The analysis has been guided by the following three research questions:

1. Are cultural references present in OCW lectures and, if so, to what extent?
2. Which type of cultural references are the most prominent?
3. Are there disciplinary differences in how cultural references are used by OCW lecturers?

In addition, a fourth methodology-oriented question was formulated with a view to evaluating the efficacy of corpus techniques, specifically semantic annotation, for the automated retrieval of open-class and highly variable linguistic elements such as cultural references. As Bianchi (2012) points out, semantic analysis is a complex undertaking, as various complicating issues may come into play when deciding which semantic domain a particular item belongs to, among which are polysemy, homography, and figurative expressions. Moreover, a given item may fall under more than one semantic domain. Previous research using semantic annotation to analyze some very specific types of discourse has suggested that it may present some limitations. For example, in a corpus of blogs written by teens and young undergraduates, Ooi (2006) found that the semantic

<sup>5</sup> Some statistical evidence of this trend in the United States has been provided by NAFSA (National Association of Foreign Student Advisers) Association of International Educators. In the academic year 2014-2015, approximately 975,000 international students were enrolled in US colleges and universities, representing a record high. In the academic year 2015-2016, international enrollments in UK higher education institutes grew to over 400,000 students (UK Council for International Student Affairs). In 2016, over 300,000 international students were enrolled in Australian universities (Australian Government Department of Education).

annotation tool of Wmatrix (Rayson 2008) assigned high numbers of items to the so-called unmatched category as it was unable to recognize creative spellings and non-standard lexical items. Similarly, Collins (2015) notes that it was necessary to carry out some manual correction in a corpus of climate change discourse compiled from *The Guardian*, when several items that had been assigned to the unmatched category needed to be re-assigned appropriately. With particular reference to cultural aspects of discourse, Potts and Baker (2012, p. 295) set out to explore the “viability of automated semantic tagging as a tool of cultural analysis” in a comparison of British and American written English. However, they concluded that semantic tagging was best suited to the initial phases of investigation which should then be integrated with follow-up cultural analysis. Thus, the fourth research question of this study is:

4. Are corpus tools useful for identifying cultural references in a corpus of OCW lectures?

## 2. Methodology

### 2.1. The OCW corpus

The corpus compiled for this study is based on the transcripts of fifteen OCW lectures in courses representing the humanities, social sciences, and hard sciences from Yale and MIT, for a total of 99,135 tokens. Yale and MIT are two of the most prestigious universities in the United States and in the world, and are both leaders in the OCW movement (Rhoads *et al.* 2013). The fifteen lecturers can be characterized as highly experienced and often distinguished faculty members. For example, the Yale website describes participating lecturers as “Leading Yale scholars and scientists who teach outstanding courses”.<sup>6</sup> MIT has referred to their participating professors as educators who “believe their purpose is to create and impart knowledge not only to MIT students but to society at large” (D’Oliveira, Lazarus 2016, p. 13). In an effort to improve comparability, the first lecture of each course was selected for inclusion in the corpus. The lectures were delivered and video-recorded during class sessions of regular ongoing courses within a timeframe that spans from 2004 to 2013. All the lectures had a frontal delivery style with limited audience interaction and took place in a large lecture hall, with the video camera positioned mostly on the lecturer, while occasionally panning to supporting visuals.

Table 1 provides an overview of the corpus, which is sub-divided into five lectures in each of the three overarching disciplinary areas mentioned above, the particular subject within the discipline, the title of the course, whether the lecture was available through Yale or MIT, gender of the lecturer, and number of tokens per lecture.

Area	Discipline	Course title	Univ.	Gender	Tokens
Humanities	Art History	Roman Architecture	Yale	female	7182
Humanities	English Studies	The American Novel Since 1945	Yale	female	6634
Humanities	History	The American Revolution	Yale	female	7654
Humanities	Religious Studies	Introduction to the Old Testament	Yale	female	7915

<sup>6</sup> <https://oyc.yale.edu/about>

Humanities	Music & Theater Arts	Musical Improvisation	MIT	male	5592
Social sciences	Political Science	Introduction to Political Philosophy	Yale	male	4238
Social sciences	Environmental Studies	Environmental Politics and Law	Yale	male	6323
Social sciences	Sociology	Foundations of Modern Social Theory	Yale	male	5271
Social sciences	Economics	Principles of Microeconomics	MIT	male	5920
Social sciences	Cognitive Sciences	Introduction to Psychology	MIT	male	8134
Hard sciences	Chemistry	Principles of Chemical Science	MIT	female	4715
Hard sciences	Biology	Introduction to Biology	MIT	male	5691
Hard sciences	Mathematics	Linear Algebra	MIT	male	4996
Hard sciences	Physics	Quantum Physics I	MIT	male	11808
Hard sciences	Biomedical Engineering	Frontiers of Biomedical Engineering	Yale	male	7066
					99,135

Table 1  
The OCW lecture corpus.

## 2.2. The analysis

The theoretical point of departure for the analysis of the corpus was the concept of “cultureme”, defined by Chesterman (2004, p. 5) as “a cultural phenomenon that is present in culture X but not present (in the same way) in culture Y”. This term was first introduced by Poyatos (1976) in the context of cultural studies to provide a method for the systematic study of a culture by identifying particular cultural activities within it that reflect symbolic values. The cultureme has also become an important notion in the field of translation studies, where it is useful for interpreting culture-specific phenomena that exist in only one of two lingua-cultures being compared in texts to be translated. Indeed, according to Nord (2014, p. 34), “translating means comparing cultures”.

Culturemes are expressed linguistically (Jaskot, Ganoshenko 2015), and have been associated with some particular domains of human cultural experience, including education, religion, history, food and drinks, sports, the media, and entertainment (Pamies 2017; Nicolae 2017). In order to identify culturemes in the OCW lectures, I processed the corpus with the semantic tagger of Wmatrix, which automatically assigns lexical items in a corpus to pre-established semantic domains. The Wmatrix tagset covers over 200 semantic domains, among which are several that could potentially reflect culture-specific phenomena: F1/Food, F2/Drinks, K1/Entertainment generally, K5.1/Sports, P1/Education in general, Q4/The Media, and S9/Religion & the supernatural. Then, I carefully examined all of lexical items assigned to each of these domains to identify possible culturemes. During this process, I focused on items that were strongly associated with American culture (the cultural background of all the lecturers). In addition, I included those that might be unfamiliar to L2 students, while excluding those that are globally entrenched phenomena and thus not likely to be problematic for L2 listeners (e.g. McDonald’s that emerged in F1/Food). On the basis of this process, I then performed follow-up searches, also expanding the analysis to examine a number of other domains, beyond those originally hypothesized, in an attempt to identify additional culturemes.

For further insights, I qualitatively examined the culturemes that emerged from the semantic analysis within their context of usage in an effort to better understand how they

may be used by lecturers to develop and explain content or to interact with the students. Finally, I looked for patterns of alignment among the culturemes across the three broad disciplinary areas (i.e., humanities, social sciences and hard sciences) to identify possible trends in usage.

### 3. Results and discussion

#### 3.1. Identifying culturemes

The initial analysis of the seven semantic domains hypothesized as likely sources of culturemes revealed surprisingly few items. No culturemes were identified in the domains associated with food, religion, and entertainment, in contrast with previous research (Pamies 2017; Nicolae 2017). Overall, very few items (N=5) were identified within the expected domains. One item was found in K5.1/Sports (*softball*) and four items were found in P1/Education in general (*graduate, undergraduate, freshman, college*). However, the presence of *freshman* prompted me to perform a simple corpus query for the three other terms used to distinguish the four years of university and high school education in the United States: *sophomore, junior, and senior*. Among these, only *sophomore* was present in the corpus, but it had been assigned by the semantic tagger to the Z99/Unmatched category, indicating that the software did not recognize this lexical item as belonging to any of its pre-established semantic fields.

This episode suggested that other culturemes could be identified, but that it would be necessary to expand upon the originally hypothesized semantic domains. Thus, I examined the contents of many other domains, such as G1/Government, politics & elections, I3/Work and employment, Y1/Science and technology in general, H1/Architecture, kinds of houses & buildings, as plausible sources of culturemes, as well as the more generic categories, Z1/Personal names, Z2/Geographical names, Z3/Other proper names, and Z99/Unmatched. Table 2 lists the domains in which culturemes emerged during this search process, and the corresponding items that were found within them. Words in brackets following some of the items have been added to provide some interpretive context. As can be seen, with the exception of those found in G1/Government, politics & elections, all the other culturemes were found in the generic Z categories, with Z99/Unmatched having by far the largest number of items.

Tag/domain	Items
G1/Government, politics & elections	Declaration of Independence, Republican Party, Democratic Party, Homeland Security
Z1/Personal names	Mickey Mantel, Wendy's, Lisa Kudrow
Z2/Geographical names	New York Yankees
Z3/Other proper names	Mickey Mouse, Congress, Sunday school, Ivy League, Boston Red Sox, Friends [TV show], Yale Bowl
Z99/Unmatched	Mexican-American War, Vassar [College], upperclassmen, sophomore, valedictory [speech] StubHub, Ticketmaster, HBO, Repo movie, Oprah [Winfrey], pre-med, midterm, McCain/Palin, Founding Fathers, Founder-like, Founder-type [guys], Fido [name of robot]

Table 2  
Expanded domain search results.

A careful examination of all the items in Table 2 led to additional insights. More specifically, some of the items actually do encode cultural meanings related to some of the originally hypothesized domains (i.e., food, entertainment, and religion), but they had been assigned to other domains. For example, *Wendy's*, which denotes an American fast-food restaurant was found in Z1/Personal names. Several items that encode cultural meanings related to entertainment were found in various other domains. For instance, *Lisa Kudrow* was found in Z1/Personal names, *Friends* and *Mickey Mouse* were found in Z3/Other proper names), while *StubHub*, *Ticket Master*, *HBO*, *Repo movie*, and *Oprah* were all found in Z99/Unmatched. Various items with cultural meanings associated with education (*Vassar*, *upperclassman*, *sophomore*, *valedictory*) were also assigned to Z99/Unmatched. The American baseball legend *Mickey Mantel* was found in Z1/Personal names, rather than K5.1/Sports. Similarly, the famous baseball teams *New York Yankees* and *Boston Red Sox* appeared in Z2/Geographical names and Z3/Other proper names, respectively, and rather inexplicably, in two different domains. The typically Protestant religious activity of *Sunday school* was found in Z3/Other proper names, rather than S9/Religion & the supernatural. It was only by using this type of cross-domain analysis that I was able to accurately identify many of the culturemes present in the corpus that would have otherwise escaped detection. Thus, from the results described above, it appears that semantic tagging as a tool to identify culturemes within a corpus is not a particularly straightforward process, as the vast majority of the items were found either in semantically unrelated domains or in generic domains without a clear semantic alignment.

Table 3 presents all the culturemes identified through the process described above, now reassigned and regrouped according to the cultural domain to which they actually correspond. In terms of frequency, due to their highly specific and variable nature, culturemes will obviously not occur in large quantities in a small specialized corpus such as the one investigated in this study. A total of 37 culturemes were identified. They were distributed across 13 of the 15 lectures represented in the corpus. No culturemes were found in the lectures on linear algebra and quantum physics. I will return to this finding in the context of the cross-disciplinary analysis in section 3.3.

The highest number of culturemes occurred in the domain of education (N=12), reflecting experiences and entities of American university life that are relevant to the lecture participants. Culturemes related to government and politics were also relatively frequent (N=9), likely at least partly due to some content overlap with two lectures (Introduction to Political Philosophy and Environmental Politics and Law). However, discussions of politics are very much entrenched in American culture, particularly among highly educated people, which could also explain their prominence. Culturemes associated with the world of entertainment (N=8) may reflect an effort on the part of lecturers to relate the content to what students are typically interested in outside of academics, such as films and television shows. Culturemes from the domains of sports, food and religion were relatively infrequent, and only two were classified according to the generic category of proper names.

Cultureme	Cultural domain	N
Graduate, undergraduate, freshman, college, Vassar [College], upperclassmen, sophomore, valedictory [speech], pre-med, midterm, Ivy League, Yale Bowl	Education	12
Declaration of Independence, Republican Party, Democratic Party, Homeland	Government, politics & elections	9

Security, Congress, McCain/Palin, Founding Fathers, Founder-like, Founder-type [guys]		
Lisa Kudrow, Mickey Mouse, Friends [TV show], StubHub, Ticketmaster, HBO, Repo movie, Oprah [Winfrey]	Entertainment	8
softball, Mickey Mantel, New York Yankees, Boston Red Sox	Sports	4
Wendy's	Food	1
Sunday school	Religion	1
Mexican-American War, Fido [name of robot]	Proper names	2
		37

Table 3  
Frequencies of culturemes.

### 3.2. Culturemes in context

Follow-up analysis of the culturemes in their context of usage revealed some particularly interesting insights into how the lecturers used them in various ways. For example, culturemes emerged during episodes in which the lecture is engaging with the audience on an interpersonal level in relation to experiences at university. In example (1), the chemistry lecturer is recounting her own experiences as a student and utilizes the term *pre-med*, the so-called pre-medical curriculum track that American students must follow in the first four years of university in order to be admitted to medical school. In example (2), the term *upperclassmen* refers to students in their third and fourth years of a four-year undergraduate degree program (i.e., *juniors* and *seniors*). In example (3), the lecturer uses *Mickey Mouse*, not as the proper noun to denote the Disney character, but rather as an adjective to describe as something that is too easy or banal to be taken seriously. Indeed, in both North American and British academia, a “Mickey Mouse” course is commonly known to be an easy course that students enroll in to receive high marks with little effort. However, this alternative meaning might be lost on L2 students.

- (1) So, I actually also started *pre-med*. Is anyone else *pre-med* here? Okay, so a pretty good showing. So maybe you can relate to some of the reasons I wanted to be *pre-med*. Part of it was the interest in the science and the biology. Also, I wanted to help people. It seemed like a really clear way that I could have a career that was challenging and involved in science, but also helping others. (MIT Chemistry lecture)
- (2) This is important because your final paper assignment will be an exercise in exegesis, an interpretation. The skills that you will need for that paper I am fairly certain are not things that you would've acquired in high school and, if we have some *upperclassmen* — I don't know, but maybe not even some *upperclassmen* will have acquired here yet. (Yale Religious Studies lecture)
- (3) I don't want this to be a “*Mickey Mouse*” course. I want this to be serious. Right? I want you to be challenged, I want you to think, I want you to read, and I want you to remember what you learned in this course. (Yale Sociology lecture)

In example (4), the American History lecturer opts to refer to important historical figures with a colloquial expression *Founder-type guys*, typical of an informal register, also seeming to take a slightly humorous slant that highlights stereotypical impressions of the so-called Founding Fathers of the United States as relatively old white males.



- (4) You may have encountered The Federalist Essays as the grand source of authority on the Constitution. Right? How could it not be that when you have *Founder-type guys* talking about the Constitution and they were the guys who were at the convention? (Yale American History lecture)

Culturemes were also used to reinforce particular points or explain concepts by drawing comparisons to culturally familiar entities, or what Miller (2002) refers to as aspects of local culture. In example (5), the lecturer refers to two well-known figures from American politics to illustrate different lecturing styles, while in example (6) the lecturer explains the concept of partisanship by means of analogy with the famous rivalry between two American baseball teams. In example (7), the lecturer illustrates the concept of price mechanism with an example that relates directly to the modern ticket purchasing experiences of the students. In example (8), the publishing experience of an American novelist is described by drawing a comparison with a well-known American celebrity who successfully promotes contemporary literature.

- (5) Now, in the past, sometimes students have found this whole thing a little frustrating, that they just get used to one lecture style, and then all of a sudden there's another lecture style, and that can be true. I mean sometimes the styles of the two professors couldn't be more different - think *McCain/Palin*, odd couples. (MIT Chemistry lecture)
- (6) Regimes are necessarily partisan, that is to say they instill certain loyalties and passions in the same way that one may feel partisanship to the *New York Yankees* or the *Boston Red Sox*, or to Yale over all rival colleges and institutions, right? (Yale Political Science lecture)
- (7) So how many people have waited on line to get a concert ticket? That's amazing. So if I asked this question 30 years ago, 90% of the hands would have gone up. OK? So what that means is the price mechanism has started to be used. It has replaced the line mechanism as a way to allocate those tickets. And we see prices working. That wasn't true 30 years ago. There wasn't *StubHub*. There weren't these secondary ticket sellers. You had to wait on line to get the tickets. (MIT Microeconomics lecture)
- (8) So you write the story of your life. It's nearly 400 pages long. It gets a really nice reception at a very good publisher. It's in page proofs. Everything's going great. You're thrilled. And then someone says to you, "You know...". Imagine this is *Oprah*. She's thinking about putting it on her book club and, if any of you know anything about contemporary literature, getting on *Oprah's* Book Club makes your sales for the next 20 years. It's huge. She says, "This is great, but you know what? I think that last hundred pages you should get rid of it." Well, this is what happened to Richard Wright. (Yale American Novel lecture)

### 3.3. Cross-disciplinary analysis

Table 4 illustrates the disciplinary areas in which culturemes were extracted from corresponding lectures. As can be seen, humanities lectures contained culturemes associated with five domains: Education, Government & politics, Food, Religion, and Proper names. Social science and hard science lectures contained culturemes associated with four domains: Education, Government & politics, Entertainment, and Sports; and Education, Government & politics, Entertainment, and Proper names, respectively. Thus,

there were no clear alignments in terms of how culturemes are distributed across the three broad disciplinary areas.

Only the two most prominent domains in the corpus (see Table 3), i.e., Education and Government & politics, were found across all three disciplinary areas, indicating that these domains are common in lecture discourse, regardless of the disciplinary content. Among the less prominent domains, there appears to be no evident trend. Sports culturemes were found only in social sciences lectures, but with so few and with two of them used by the same lecturer, no real conclusions can be drawn. Moreover, the presence of *Sunday school* in the lecture on religion is clearly linked to the topic of the lecture. Perhaps what is most interesting is the fact that no culturemes were identified in two lectures representing the hard sciences. In the linear algebra lecture, the lecture spent the entire time working through algebraic formulas on the blackboard. Similarly, the quantum physics lecture was dedicated to the description of experimental activities with extensive use of the blackboard. Although more lecture data would be necessary to determine possible disciplinary alignments, the lack of culturemes in these two lectures suggests that they may be less likely to emerge in the empirically-oriented content of hard science lectures.

Cultureme	Cultural domain	Disciplinary areas where present
Graduate, undergraduate, freshman, college, Vassar [College], upperclassmen, sophomore, valedictory [speech], pre-med, midterm, Ivy League, Yale Bowl	Education	Humanities Social Sciences Hard Sciences
Declaration of Independence, Republican Party, Democratic Party, Homeland Security, Congress, McCain/Palin, Founding Fathers, Founder-like, Founder-type [guys]	Government, politics & elections	Humanities Social Sciences Hard Sciences
Lisa Kudrow, Mickey Mouse, Friends [TV show], StubHub, Ticketmaster, HBO, Repo movie, Oprah [Winfrey]	Entertainment	Social Sciences Hard Sciences
softball, Mickey Mantel, New York Yankees, Boston Red Sox	Sports	Social Sciences
Wendy's	Food	Humanities
Sunday school	Religion	Humanities
Mexican-American War, Fido [name of robot]	Proper names	Humanities Hard Sciences

Table 4  
Culturemes across disciplinary areas.

#### 4. Concluding remarks

In this study, I have attempted to shed some light on how culture is represented in a corpus of OCW lectures drawn from three broad disciplinary areas (i.e., humanities, social sciences and hard sciences), using the cultureme as the unit of analysis. This was accomplished by implementing a two-pronged methodology that integrated quantitative corpus methods and qualitative contextual analysis. Returning to the research questions posited out the outset, with reference to the first question, culturemes were present in all but two of the fifteen lectures that comprised the corpus, although not in high quantities.

This is in line with their highly specific and variable nature, as well as the fact that they were typically limited to episodes of exemplification or interpersonal commentary, and were not themselves the topics of the lectures. Moreover, many culturemes appeared only once, displaying a primarily one-off nature in this communicative context. Yet this low frequency does not correspond to irrelevance. In fact, in the context of L2 lecture comprehension, the presence of even one unfamiliar term/concept can be problematic for listeners, potentially causing them to lose concentration and to feel isolated from the L1 counterparts who are able to interpret culturemes without difficulty. A similar finding emerged in Crawford Camiciottoli's (2007) analysis of idioms in a corpus of business studies lectures. Idioms were also not particularly frequent and largely used only once, but they were nonetheless features that could present significant challenges for L2 listeners.

From a semantic perspective (research question 2), not surprisingly, culturemes that referred to the field of education were the most frequent, reflecting the strong influence of academic culture (Miller, 2002) across the OCW corpus. This finding could also have been linked to my decision to select the first lecture of each course for inclusion in the corpus (see section 2.1). Because lecturers often address topics related to general organization during the first lecture of the course, they might have had more opportunities to use education-related culturemes. It would be necessary to expand the corpus to include other lectures throughout the course to gain a better understanding of the role of education-related culturemes in lecture discourse. Culturemes in the semantic domains of government/politics and entertainment were also relatively frequent, likely due to the high level of interest in politics in American educated society and to the lecturers' efforts to engage students by relating the lecture to the local culture (Miller 2002).

Concerning possible disciplinary differences in how cultural references were used by the lecturers across the three areas represented in the corpus (research question 3), no clear alignments or patterns emerged, other than the lack of culturemes in the two hard science lectures discussed in section 3. Again, more insights could be gained by expanding the corpus to include more lectures in each of the three disciplinary areas. This would also help to counterbalance potential idiosyncracies of the lecturers themselves (e.g., more or less interactional style that could influence the use of culturemes), which cannot be ruled out when analyzing the speech of a limited number of individuals.

The fourth research question focused on methodology, i.e., whether corpus tools are useful for analyzing cultural references in a corpus of OCW lectures. The answer here appears to be a "mixed bag". On the one hand, because culturemes were often not found in their clearly corresponding domains, but were only identified through extensive cross-domain analysis, the use of automated semantic annotation software was not as efficient as I had hoped, thus corroborating the previous research experiences of others who have used Wmatrix in semantic analyses of corpora of highly specialized language (Ooi 2006; Collins 2015; Potts, Baker 2012). Indeed, the fact that the Unmatched/Z99 tag turned out to contain the highest number of culturemes shows the limitations of this methodological approach for analyzing the representation of culture. On the other hand, the semantic tagging did allow for at least some degree of systematicity by producing lists of items to be searched for culturemes, thus avoiding the need for an in-depth reading of full texts in paragraph format in order to distinguish culturemes manually. Thus, to quote Wilson and Thomas (1997, p. 55), "there is no such thing as an 'ideal' semantic annotation system".

To conclude, I offer some thoughts on the pedagogical applications of the findings of the study. Because culturemes are used by content lecturers, they need to be aware of the potential difficulties that they can create for L2 learners during lectures. Miller

suggests that content lecturers could benefit from some form of training to prepare them to lecture to L2 learners:

Lecturers need to become aware of how they, themselves, use language to talk about their discipline, and then integrate a high level of meta-language into their lectures, that is, they need to talk not only about the content but also explain how they are talking about the content". (Miller 2002, p. 157)

This self-awareness should include a conscious reflection on aspects of culture that may find their way into the content of the lecture. Indeed, lecturers may wrongly assume that all students in the class are familiar culture-specific concepts and their corresponding linguistic expression.

The highly sporadic and content-specific nature of many of the culturemes identified in this analysis suggests that they are not teachable items per se. However, the higher frequencies of both academic and political culturemes suggests that they could be incorporated into orientation activities for L2 learners who either plan to access lectures in English online or participate in a study-abroad program in an English-medium university. For example, such activities could include overviews of the higher education system, terminology, and student life in the chosen country, as well as its governmental structure and political system. This type of preparation would be useful as these are the cultural aspects that appear to characterize the shared educational experience between lecturers and students in university settings and that transcend disciplinary focus.

**Bionote:** Belinda Crawford Camiciottoli is Associate Professor of English Language and Linguistics at the University of Pisa in the Department of Philology, Literature and Linguistics. Her research focuses on corpus-assisted and multimodal analysis of discourse in academic, professional, and digital settings. She has published in leading journals, including *Journal of Pragmatics*, *Intercultural Pragmatics*, *Discourse & Communication*, *Text & Talk*, *Discourse, Context & Media*, *Journal of Business Communication*, *Journal of English for Academic Purposes*, and *English for Specific Purposes*. She authored the monographs *The Language of Business Studies Lectures. A Corpus-assisted Analysis* (John Benjamins, 2007) and *The Rhetoric of Financial Discourse. A Linguistic Analysis of ICT-mediated Disclosure Genres* (Rodopi, 2013).

**Author's address:** belinda.crawford@unipi.it

**Acknowledgements:** This research has been financed by the Italian Ministry for the University (PRIN 2015 no.2015TJ8ZAS).

## References

- Bianchi F. 2012, *Culture, corpora and semantics. Methodological issues in using elicited and corpus data for cultural comparison*. ESE Salento University Publishing, Lecce. <http://siba-ese.unisalento.it/index.php/culturecorpora/issue/view/1111> (16.02.2018).
- Bligh D. 2000, *What's the use of lectures?* Jossey-Blass, San Francisco.
- Chesterman A. 2004, *What is a unique item?*, in Gambier Y., Shlesinger M., and Stolze R. (eds), *Doubts and directions in translation studies. Selected contributions from the EST Congress, Lisbon 2004*, John Benjamins Amsterdam, pp. 3-13.
- Clay T. and Breslow L. 2006, *Why students don't attend class*, in "MIT Faculty Newsletter", XVIII [6-7]. <http://web.mit.edu/fnl/volume/184/breslow.html> (03.10.2016).
- Coleman J.A. 2006, *English-medium teaching in European higher education* in "Language Teaching" 39, pp. 1-14.
- Collins L. 2011, *How can semantic annotation help us to analyse the discourse of climate change in online user comments?* in "Linguistik Online" 70 [1]. <https://bop.unibe.ch/linguistik-online/article/view/1743/2958> (12.01.2018).
- Crawford Camiciottoli, B. 2005, *Adjusting a business lecture for an international audience: A case study*. in "English for Specific Purposes" 24 [2], pp. 183-199.
- Crawford Camiciottoli, B. 2007, *The Language of Business Studies Lectures: A corpus-assisted analysis*. John Benjamins, Amsterdam.
- D'Oliveira C. and Lazarus J.S. 2016, *MIT OpenCourseWare. A leader in open education*, in "The Bridge" 46 [23], pp. 12-23.
- DiPiro J.T. 2009, *Why do we still lecture?* in "American Journal of Pharmaceutical Education" 73, p. 1.
- Flowerdew J. 1994, *Research of relevance to L2 lecture comprehension*, in J. Flowerdew (ed.), *Academic listening* Cambridge University Press, Cambridge, pp. 7-29.
- Jaskot M. and Ganoshenko I. 2015, *Culturemes and non-equivalent lexis in dictionaries*, in "Cognitive Studies/Études cognitives" 15, pp. 115-124.
- Lerman S.R. and Miyagawa S. 2002, *Open courseware: A case study in institutional decision making*, in "Academe" 88 [5], pp. 23-28.
- Limbach B.J. and Waugh W.L. 2005, *Questioning the lecture format*, in "The NEA Higher Education Journal", Fall, pp. 47-56.
- Martinez S. 2014, *OCW (OpenCourseWare) and MOOC (open course Where?). Proceedings of OpenCourseWare Consortium Global*. [http://cdlh7.free.fr/OCWC\\_2014/Final\\_papers/Paper\\_16.pdf](http://cdlh7.free.fr/OCWC_2014/Final_papers/Paper_16.pdf) (15.01.2018).
- Miller L. 2002, *Towards a model for lecturing in a second language*, in "Journal of English for Academic Purposes" 1, pp. 145-162.
- Mulligan D. and Kirkpatrick A. 2000, *How much do they understand? Lectures, students and comprehension*, in "Higher Education Research & Development" 19 [3], pp. 35.
- Nicolae A.O. 2015, *Multicultural dialogue: Translating culturemes. The Proceedings of the International Conference Literature, Discourse and Multicultural Dialogue*. (Arhipelag XXI Press, Tîrgu Mureş, Romania, pp. 215-222.
- Nord C. 2014, *Translating as a purposeful activity: Functionalist approaches explained*. Routledge, Milton Park.
- Ooi V.B.Y. 2009, *Computer-mediated language and corpus linguistics*, in Kawaguchi Y., Minegishi, M. and Durand J. (eds.), *Corpus analysis and variation in linguistics*, 1, John Benjamins, Amsterdam, pp. 103-120.
- Pamies A. 2017, *The concept of cultureme from a lexicographical point of view*, in "Open Linguistics" 3[1], pp. 100-114.
- Parini J. 2004, *The well-tempered lecturer*, in "The Chronicle of Higher Education" 50, B5. Retrieved from <https://chronicle.com/article/The-Well-Tempered-Lecturer/7998/> (05.12.2016).
- Pott A. and Baker P. 2012, *Does semantic tagging identify cultural change in British and American English?*, in "International Journal of Corpus Linguistics" 17 [3], pp. 295-324.
- Poyatos F. 1976, *Analysis of culture through its culturemes: Theory and method*, in Rapoport A. (ed.), *The mutual interaction of people and their built environment*, Mouton, The Hague, pp. 265-274.
- Rayson P. 2008, *From key words to key semantic domains*, in "International Journal of Corpus Linguistics" 13 [4], pp. 519-549.
- Rhoads R.A., Berdan J. and Toven-Lindsey B. 2013, *The Open Courseware movement in higher education: Unmasking power and raising questions about the movement's democratic potential*, in "Educational

- Theory” 63 [1], 87-110.
- Vladiou M. 2011, *State-of-the-art in Open Courseware initiatives worldwide*, in “Informatics in Education” 10, pp. 271-294.
- Wilson A. and Thomas J. 1997, *Semantic annotation*, in Garside R., Leech G. and McEnery, T. (eds.), *Corpus annotation. Linguistic information from computer text corpora* Longman, London, pp. 53-65.
- Zhu W. and Flaitz J. 2005, *Using focus group methodology to understand international students’ academic language needs: A comparison of perspectives*, in “TESL-EJ”, 8, p. 4. Retrieved at <https://files.eric.ed.gov/fulltext/EJ1068108.pdf> (09/25/2017).

## Annex. Screenshot from Yale Open course “The American Revolution”

HOME » HISTORY » THE AMERICAN REVOLUTION » HIST 116 - LECTURE 1 - INTRODUCTION: FREEMAN'S TOP FIVE TIPS FOR STUDYING THE REVOLUTION

### HIST 116: The American Revolution

#### Lecture 1 - Introduction: Freeman's Top Five Tips for Studying the Revolution

<< Previous Session | [Next Session](#)

>>

##### Overview

Professor Freeman offers an introduction to the course, summarizing the readings and discussing the course's main goals. She also offers five tips for studying the Revolution: 1) Avoid thinking about the Revolution as a story about facts and dates; 2) Remember that words we take for granted today, like “democracy,” had very different meanings; 3) Think of the “Founders” as real people rather than mythic historic figures; 4) Remember that the “Founders” aren't the only people who count in the Revolution; 5) Remember the importance of historical contingency: that anything could have happened during the Revolution.



##### Lecture Chapters

1. [Introduction: Is the War Part of the American Revolution? \[00:00:00\]](#)
2. [Reading Materials for the Course \[00:08:24\]](#)
3. [Freeman's Tips One and Two: Facts and Meanings \[00:13:45\]](#)
4. [Freeman's Tip Three: The Founders Were Human, Too \[00:22:13\]](#)
5. [Freeman's Tip Four: The Other Revolutionaries \[00:31:33\]](#)
6. [Freeman's Tip Five and Conclusion \[00:37:48\]](#)

TRANSCRIPT	AUDIO	LOW BANDWIDTH VIDEO	HIGH BANDWIDTH VIDEO
------------	-------	---------------------	----------------------