*Language Teaching* (2020), **53**, 203–214
doi:10.1017/S0261444819000107

CAMBRIDGE
UNIVERSITY PRESS

PLENARY SPEECH

# Vague language in business and academic contexts

Michael McCarthy

University of Nottingham, UK
E-mail: mactoft@cantab.net

**Abstract**

Business English and academic English may perhaps be thought of as areas of language use requiring precision of expression and a quest for specific and unambiguous meaning. However, corpus evidence shows that both types of language, in their spoken contexts, exhibit noticeable use of the kinds of vague expressions found in everyday conversation. In this lecture, I focus on one type of vagueness: VAGUE CATEGORY MARKING (VGM). This feature involves mention of an example or examples of something followed by reference to a broad, ad hoc category of which the chosen examples are seen as typical. References to categories commonly involve expressions such as *or whatever, and so on, or something (like that)*. In both business and academic English, vague category marking is an important projection of shared knowledge and shared identities. In business, vagueness is also a useful tool in delicate negotiations. In academic English, vague categories refer to bodies of assumed shared knowledge and are crucial in the pedagogic process of grafting new knowledge onto old. Subtle differences are drawn out by different types of vague category markers. I conclude with some implications for teaching in these specialised areas.

## 1. Being clear about vagueness

Vagueness is pervasive in everyday language. English offers numerous ways in which we can blur or make less precise the things we say and write. *Things* in what I've just said is a good example: what are those 'things'? Conversations? Essays? Emails? Tweets? Confessions? Promises? It could be all or any. By using *things,* I give myself a let-out from being challenged for having omitted important items, and I give *you* a bit of scope to conceptualise what kinds of 'things' I might be talking about. I may not even be sure myself what all these 'things' are. The noun *thing* is versatile; despite its apparent countability and its perfectly paired uncountable partner *stuff*, *thing* can be used of animate and non-animate entities and of countable and uncountable ones (as can *stuff*). Cutting (2006) has some interesting examples of the use of *thing*. English also offers an array of vague quantifiers when we cannot, or do not want to be precise about numbers and quantities, my favourite example being, from some years back, a police spokesman in London who announced to the press after a tumultuous street demonstration that 'approximately eleven people were arrested'. Such usage, odd when put under the microscope of linguistic scrutiny, passes generally unnoticed in everyday use. Other morphemes, words and phrases also allow us to de-focus from precise commitment to sharp-edged notions; you can be *hungry-ish*, an object can be *yellowy* in colour, and we can be *sort of tired* or find something *kind of irritating*. Such usage is normal, especially in conversational language; it is not 'lazy' or 'evasive', and normal conversation would be impossible without it. But what about the straight-talking world of business and the detailed and precise world of academia?

Vagueness is well-covered in the literature; Channell (1994) looks at a range of vague lexis, including quantifiers, and vague words such as *whatsisname* when the speaker can't find the right word. Cutting's (2007) edited volume has papers on the types of vagueness I have just illustrated and

Revised version of a keynote lecture given at the 3rd Languages in the Globalised World International Conference, Leeds Beckett University, Leeds, UK, 23 May 2018.

© Cambridge University Press 2019

https://doi.org/10.1017/S0261444819000107 Published online by Cambridge University Press

more. And vague language is relevant to the concerns of this conference on language in a globalised world, in that questions of world Englishes and the use of English as a lingua franca, along with the powerful position English occupies in international business and academic life, lead us to ask whether investigating vagueness in any single variety of English is valid and generalisable, not to mention whether vague expressions can be reliably translated across languages, a topic I will return to briefly later. In the English as a lingua franca field, Metsä-Ketelä (2012, 2016) looked at the kinds of vague expressions I want to examine today in an academic setting and found a greater frequency of usage of vague expressions among lingua franca users, though with a narrower range of items when compared with the native-speaker repertoire. Such measures of comparison often end in insoluble claims of overuse or underuse by non-native users. In the business sphere, Handford (2010, p. 145) finds little difference in the occurrence of language clusters (chunks) between native-user meetings and a sub-corpus of meetings where English is used as an international language and argues that it may be time to reappraise the first language (L1)/second language (L2) distinction in business and professional contexts. Comparisons like these may be inconclusive, but they at least alert us to the fact that areas of usage such as vague language and spoken language practices in general are questions to be addressed in more globally-located contexts of English.

In this talk, I would like to argue that vagueness is relevant wherever we look at conversational exchanges as a basic, human, universal communicative act. Vagueness may well take different forms in different languages and varieties, and, as I hope to show today, its distribution and functions are context-dependent, but none of that should distract us from the starting point that vagueness is a fundamentally human, interactive affair which creates and maintains bonds among groups by projecting and exploiting common ground and shared knowledge. For this reason, vagueness in language use should always occupy a prominent place in our descriptions and prescriptions, especially in the pedagogical domain. To ignore it or to marginalise it is to undervalue one of the core features of human communication.

Today I want to focus on one particular type of vague expression in English and its use in two somewhat specialised contexts: academic speaking and spoken business English. The type of expression I want to look at is what are called VAGUE CATEGORY MARKERS, also known as *general extenders, extension particles, set marking tags*, *generalised list completers, vague tags* and *vague category identifiers*, among various other labels for the phenomenon (see Channell, 1994; Cheshire, 2007; Dines, 1980; Dubois, 1992; Jefferson, 1990; O'Keeffe, 2004; O'Keeffe & Adolphs, 2008; Overstreet, 1999, 2005; Overstreet & Yule, 1997). Example 1 below is a classic case from the CANCODE spoken corpus. CANCODE stands for 'Cambridge and Nottingham Corpus of Discourse in English'. For full details of the corpus, see McCarthy (1998). This corpus, and sub-corpora of it, are the main source of data for the present lecture.

Example 1

[context: the speaker is describing a riverside parking area much patronised by motor-cyclists. <$1> = speaker 1]

<$1> *There's like a little caravan that sells hot dogs and drinks* ***and things like that****.*

The vague category marker (which I will abbreviate from now on to VCM) is *and things like that*, and it is extremely common in conversational English, occurring 576 times in the five-million-word CANCODE spoken corpus. Along with items such as *or whatever, something like that, (and) that sort of thing, and stuff*, it forms part of a repertoire of fixed expressions whose job it is to refer implicitly to categories of people, things and ideas without giving an exhaustive (and undesirable, tedious and perhaps impossible) list of everything in the projected category.

https://doi.org/10.1017/S0261444819000107 Published online by Cambridge University Press

In the case of example 1, we are led to assume that the caravan sells a range of items of which hot dogs and drinks are seen as the most typical, often referred to as the EXEMPLAR(s). For the speaker to list the total list of all the items which the caravan sells would be an absurdity; it would most likely not be known or available to the speaker anyway and would be tiresome and irrelevant to the listener(s), who would rapidly lose the will to live. Instead, speakers choose one or two exemplars to trigger an ad hoc category in the listener's mind. We never know whether what the category is felt to include in the mind of the speaker is the same as in the mind of every listener, nor do we need to know; all is well, and the conversation can proceed to its goals as long as both parties have an adequate idea of what is included at the moment of utterance. In example 1, the combination of exemplar(s) and the VCM is deemed adequate for the listener mentally to create a sufficient idea of what the caravan is all about. A VCM is a projection or assumption of a high degree of shared knowledge with one's interlocutors and, as such, may be culturally determined, with the notion of a 'caravan' as a non-permanent café or snack-bar being not necessarily universal, yet unproblematic to those familiar with informal, cheap-and-cheerful sales outlets in popular visitor locations, quite a common sight in the UK.

O'Keeffe (2004, 2006) showed how VCMs can project different types or levels of shared knowledge. She looked at 138 examples of vague categories in a radio phone-in programme broadcast in the Republic of Ireland. Some of the categories referred to people, events and things which were probably only interpretable to Irish listeners, some were interpretable on a wider, European scale and some were global/universal (i.e. almost anyone would understand them). O'Keeffe suggested that VCMs which referred to Irish cultural and national matters and current affairs helped the radio listeners to identify as a group sharing common concerns. They manifested as an in-group, rather like the students in Cutting's (1999) study of how implicit language used over time functioned to cement the identity of a group of students on the same academic course.

VCMs in general conversational contexts have been well-investigated (e.g. Evison, McCarthy, & O'Keeffe, 2007; Jucker, Smith, & Lüdge, 2003; Overstreet, 1999; Overstreet & Yule, 1997), with attention focusing on the creation of ad hoc categories (i.e. categories that are not lexically fixed), hedging, politeness and other pragmatic features which I will return to later. Some scholars have looked at vagueness in more circumscribed and specialised uses of language such as radio phone-ins, the context for O'Keeffe's (2004) study, while Adolphs, Atkins and Harvey (2007) looked at vague language in health communication, Koester (2010) in office communication, and Cotterill (2007) in courtroom discourse. Less studied is the business context (though Handford, 2010 pays some attention to it) and the academic context, where limited research is reported in O'Keeffe (2006, pp. 133–136) and Evison et al. (2007), along with the lingua franca studies already mentioned (Metsä-Ketelä, 2012, 2016). Today I want to combine the business and academic contexts in a comparison. One reason for this is that they both represent somewhat special contexts of speaking, and both are concerned with creating and reinforcing group identities, though in different ways. Comparing them, I hope, will enable them to shed further light on each other and to tell us something further about VCMs in general. Business and academia are also two good places to locate a study on the basis that it may be tempting to think they would be less inclined to exhibit the use of VCMs than casual conversation among social contacts, friends and intimates, since their contexts are likely to be more formal and more goal-driven, and more precise and specific in how they refer to people and things.

## 2. English VCMs: items and types

I am committed to corpus-based research wherever possible, especially when it comes to investigations of the spoken language, which it's very difficult for any of us to be objective about, since it all happens in real time, online, so to speak. In this respect, native users possess no particular advantages over non-native users. I believe that corpora can provide empirical evidence to either support or challenge preconceived ideas about language use, and corpora have served me well over the years. However, there are difficulties in searching for VCMs using corpus-analytical software, as opposed to, say, searching

https://doi.org/10.1017/S0261444819000107 Published online by Cambridge University Press

for articles, modal verbs or prepositions, where we know what we are looking for because they are already established as closed classes. Fortunately, two sources of information enable us to garner reliable data. Firstly, previous researchers have already provided lists of VCMs from their own data which can form the basis of searches in new datasets. Secondly, the corpus linguist proceeds from the reasonable hypothesis that VCMs are likely to be routinised and formulaic to a great extent and not created on the fly, and so lists of recurring chunks of various lengths automatically generated by frequency-counting provide a necessary back up and checking mechanism on existing lists. Lists of chunks from the two contexts, business and academic, may expose potential VCMs not commonly found in other conversational contexts or, vice-versa, VCMs established in other contexts may occur more, or less frequently in the two specialised contexts, or not at all. All this can be said against the background that VCMs from casual conversational contexts are likely to provide a suitable benchmark, in the sense that casual conversation has been used successfully over the years as a benchmark in many other analyses of different contexts of speaking (Drew & Heritage, 1992; Sacks, Schegloff, & Jefferson, 1974).

Based on my own work and work done in collaboration with other researchers, and after checks on other people's material, I will refer to the list of the most frequent VCMs in social conversation given in Evison et al. (2007), which used a one-million-word sub-corpus of social and intimate conversations compiled from the larger, five-million-word CANCODE spoken corpus. Evison et al. found that 11 VCMs occurred more than ten times per million words (ten occurrences being the cut-off point for inclusion in the study). Top of the list were variations of the impersonal pronouns *something, anything* and *everything* in the strings *and/or [something/anything/everything] (like that)*, where round brackets indicate optional members of the string and square brackets indicate alternatives. Next came *(and/or) stuff (like that)*, which was only 60% as frequent as the top string. Frequency then declined till *and all the rest of it* and *(and) this that and the other* showed 12 and 11 occurrences, respectively, hitting the baseline. So, the frequency range is considerable, with a distinct preference for VCMs based around *thing*.

However, Evison et al. (2007) conflated some items (bringing together, for example, impersonal pronouns *something, anything, everything*, and combining *kind* and *sort*), items which could equally well be treated as separate ones. They were conflated because they follow the same pattern with just some internal variation. This conflation was okay for the purposes of the publication in which it was placed, but there is good reason to reshape the list, given that some linguistics have argued for the significance of the distinction between VCMs that start with *and*, as opposed to those which start with *or*. The labels given to these are ADJUNCTIVE (those introduced by *and*, as in *and things like that*) and DISJUNCTIVE (those introduced by *or*, as in *or whatever*). Overstreet (1999, pp. 3–4), relates the two types to distinct pragmatic functions connected with politeness strategies; Aijmer (2015) suggests that disjunctive VCMs may offer the potential for a looser interpretation of the projected category. My own recent collaborative cross-linguistic research on VCMs in English and Russian (Malyuga & McCarthy, 2018) has shown that the adjunctive-disjunctive distinction is an illuminating one, and so I will bring it into service here since we are comparing two types of specialised English, business and academic, and it may prove useful as a way of illustrating more subtle similarities or differences in the data. This means that *and things like that* will be treated differently to *or things like that*, not just in terms of distribution but in potentially different functions.

## 3. Business data

My business data today come from CANBEC (Cambridge And Nottingham Business English Corpus), a just over 900,000-word corpus of spoken business English. Recordings for the corpus were made at meetings from 2001 onwards in 26 companies, in small and larger industrial and service businesses mostly involving middle- or upper-management UK speakers, and about 10% of the speakers were non-native users, though it is not my present aim to separate these out in any way. The businesses in CANBEC include manufacturers of industrial equipment (e.g. cranes and lifting gear), pharmaceuticals, service industries (e.g. the hospitality industry, financial services, consultancy and IT). Meetings

https://doi.org/10.1017/S0261444819000107 Published online by Cambridge University Press

varied in the numbers of speakers; some were meetings between different companies, some were internal. Topics included everyday problems and procedures, schedules, decisions, logistics, sales and marketing, human resources and so on. The corpus also includes 'academic business discourse' (for example, lectures given at a university business school). Michael Handford, one of the primary movers of the CANBEC project, gives full details of the contents of the corpus and a comprehensive analysis of it in his book on business meetings (Handford, 2010).

Malyuga and McCarthy (2018) used VCM data from CANBEC to compare with VCMs in the business and professional sub-corpus of the Russian National Corpus. We were unable to do a straight statistical comparison because the English and Russian corpora are not matched in compilation or annotation and searchability, and so we came at it from the angle of translatability, posing the question: how translatable are VCMs in similar contexts across languages? In the study we used a baseline list of VCMs drawn from conversation (as given in Evison et al., 2007) and from other sources in the literature as a check and backup. I will refer today to the business English VCM list from Malyuga & McCarthy. Although that list and the Evison et al. list couldn't simply be placed side-by-side (because of the conflation of items in the latter), it is clear that there is almost complete overlap in the actual items appearing in the top ten, but with variations in individual frequency. Notably though, VCMs appear to be only half as frequent in the business data as in the social conversations (1293 per million words and 2940, respectively). This can be explained by the very high degree of (projected) shared knowledge that is typical of casual conversation among friends, family and acquaintances. As Evison et al. (2007, p. 146) put it, the closer the relationships among speakers, 'the greater the shared space they can exploit'.

Malyuga & McCarthy's top four business items are:

*or something (like that)*
*or whatever*
*and whatever*
*etcetera*

There is a steady drop-off in frequency after these. For the moment I'll just ask you to keep in mind that the top two are disjunctive (*or*-VCMs), with the third being the adjunctive (*and*-VCM) counterpart of the second. Firstly, I'd like to look at the different levels of shared knowledge the VCMs draw upon in business meetings. In this and the following examples, the exemplar is underlined and the VCM is in bold.

Example 2

<$1> *That's why whenever supermarkets introduce their own brand, they have to do massive price reductions and lots of incentives like, you know, <u>buy one, get ten free</u>* ***or whatever****.*

This is a good example because it contains the second most frequent VCM in the business data, but also because it is prefaced by *like* and *you know*, which emphasise the appeal to shared knowledge. In this case the category is interpretable by anyone familiar with supermarket shopping and tempting special offers, albeit tinged with a bit of sarcastic humour (*ten free*). In the next example, the reference is to a much narrower world, that of the individual company.

Example 3

<$1> *Any teams which have ongoing lines like <u>toothpaste</u>* ***and all that that sort of thing****, they will work on the new way …*

Only insiders can know this company's full range of products; we can make intelligent guesses as to what might be included here (shampoo? dental floss? a narrow/broad range of personal hygiene items?) but what matters is that the listeners know which teams they are in and whether they will

https://doi.org/10.1017/S0261444819000107 Published online by Cambridge University Press

be affected by the 'new way' of working, so this is quite an in-group level of shared knowledge. The next example comes from a negotiating meeting between two companies, where logistics are being discussed. Once again, we see the reinforcing *you know*.

Example 4

<$1> *I guess I I'm keen that the knowledge is retained within the erm, you know, the visibility of the production schedules* ***or whatever***.

This is a reference broadly interpretable by anyone familiar with the business world and production schedules; I personally wouldn't have a clue what a production schedule looks like. This is a wider projection of group identity: one company does not need to explain to the other what a production schedule is; it's part of their shared world. But we also find categories that anyone and everyone can make sense of, as in the next example.

Example 5

<$1> *So it can be left probably until June, July,* ***something like that***.

This, generally speaking, is how the VCMs work in the business data: just as in O'Keeffe's (2004) analysis, different levels of knowledge are called upon and different types of identities are projected (people in general, the business community, a particular team in a particular company, etc.). We can show these different projected identities, from the narrowest to the broadest, as in this diagram:

within-company –> across companies –> the business community –> broad, non-business

## 4. Academic data

Let's turn for a while to academic data. We'll come back to the business data when we take a more bird's-eye view of the VCMs in our two datasets and in our closing remarks. The academic data for today is a sub-corpus of 460,000 words of academic encounters (e.g. lectures, seminars, tutorials) taken from the five-million-word CANCODE spoken corpus. The process of searching was similar to that of the business data: use an existing list as a starting point, once again calling on Evison et al.'s (2007) list compiled from social conversation data and checked against literature sources. Total occurrences were normalised to per one million words to make them more readily comparable with the business list and the conversational list, which are also normalised to one million words.

Once more there is considerable overlap between the academic data, the conversational data (but again, let's not forget the conflation of items in that earlier list) and the business data, with a couple of interesting differences. Here I'll just focus on comparing business and academic. The top four VCMs in the academic data, with the business data side-by-side, are:

| Business | Academic |
|---|---|
| *or something (like that)* | *or something (like that)* |
| *or whatever* | *and so on* |
| *and whatever* | *or whatever* |
| *etcetera* | *etcetera* |

*Or something (like that)* is number one in both lists, and *or whatever* comes second in business and third in academic. What is different is *and so on*, in second place in the academic list, but number ten in the business list. In the academic list, *and so on* is closely followed in fourth place by its Latin equivalent *etcetera* (and its spelling variants in the transcripts *et cetera* and *etc.*); it is also fourth

https://doi.org/10.1017/S0261444819000107 Published online by Cambridge University Press

place in the business list, so there's a good match there. Why *and so on* should not be as frequent in the business data as in the academic is difficult to explain, and I just leave it as an interesting nugget.

Overall, the pattern is similar in the academic and business data: both show a steady fall-off in frequency from the top four (all with over 100 occurrences per million in both datasets) till we get to those just above the baseline of ten occurrences. VCMs in the academic data are less than half as frequent as in the social conversation data but only a little less frequent than in the business data (1120 per million compared with 1293 per million, respectively). The picture is clear and unsurprising: conversation has a much higher incidence of VCMs for the reason already mentioned, i.e. the greater shared space that is available in talk between acquaintances, friends and intimates.

## 5. Adjunctive and disjunctive VCMs

It is interesting how the separation of the adjunctive and disjunctive types brings disjunctive VCMs floating to the top of the list in both business and academic. So, we ask ourselves, why should speakers prefer *or something (like that)* and *or whatever*? The potential adjunctive form *and something (like that)* occurs zero times in both datasets. That's not to say that adjunctive forms don't figure prominently: we've seen high ranks for *and so on* and *and whatever*, while the variations of *(and) (all) that/those sort(s)/type(s)/kind(s) of thing(s)* feature as number five in the academic list and six in the business list. The interesting thing about disjunctive *or*-VCMs is that they leave open the door for a looser, wider interpretation of the category being evoked than the adjunctive *and*-VCMs, a claim we find support for in Aijmer (2015). Let's consider some examples.

Example 6

[Academic: seminar on commercial law]

<$1> *For example, we give minors special protection in the law and people under sixteen* ***or whatever*** *…*

Here the lecturer makes a broad reference to how the law protects young and/or vulnerable groups. We can see this as an appeal to the kinds of groups most adults could understand. It would not seem, at this point, relevant to express the people referred to as a closed, legally-defined group – it is offered as a fairly loose category, simply as a way of exemplifying special protection under the law. We get a similar type of reference in example 7.

Example 7

[Academic: medical lecture on land-mine injuries]

<$1> *… because of the fact that tanks move best on roads. These are going to be placed mainly in highways, streets, crossroads* ***or whatever****.*

Here the lecturer is discussing where land-mines are typically placed with the purpose of disabling or destroying military tanks: the listeners are referred to general knowledge of road and street landscapes rather than any medical, technical or military class of entities – the category is again quite loose, a sort of 'fill the gaps for yourself' message, for which disjunctive VCMs seem well-suited.

In example 8, the disjunctive VCM is used in the sensitive context of a sales negotiation between two companies. Speaker <$1>emphasises the open-endedness of what is being proposed; it's not the end or anything comparable to a closed situation. It's a *kind of* starting level, *just kind of to start from* and *something* for them to work towards; as the turns unfold <$1>seems to be reinforcing a vague open-endedness with regard to where things might lead.

https://doi.org/10.1017/S0261444819000107 Published online by Cambridge University Press

Example 8

[Business: sales negotiation meeting between two companies]

<$1> *What we've got really is kind of a starting level. This isn't the end of it* ***or anything like that.***
<$3> *Mm.*
<$1> *It's just kind of+*
<$2> *Right.*
<$1> *+to start from+*
<$2> *Okay.*
<$1> *+and something for us to work towards.*
<$3> *Mm.*

The listeners acknowledge the VCM, albeit only minimally, and seem ready to go ahead. There is a good reason for including more context like this and not just looking at single lines of talk, especially when we get into the detail of trying to understand the subtle differences between adjunctive and disjunctive. What we have to remember is that VCMs do not ESTABLISH categories or refer to existing, lexicalised categories (for example, the category of 'mammals' or 'vehicles'); VCMs PROJECT POSSIBLE categories, and such categories may be presented as open or closed to a greater or lesser degree. It is for the listener(s) to decide what the speaker means, and listener response will tell us about how the VCM is received, so, to fully explain and understand how VCMs work, we have to look at turn-taking. VCMs and their meanings are a joint enterprise, an example of co-construction which we can best observe in the progression of turns (see Clancy & McCarthy, 2014).

Turning now to adjunctive VCMs, we can see some subtle differences. Here the categories projected seem more delineated, offering less opportunity for open-ended interpretation. Example 9 has two VCMs (in bold) and in both cases, the exemplars offer limited scope for open-endedness; in each case potential open-endedness is circumscribed by the words preceding: *(getting people) to get back to me* and *start dates*, respectively. There aren't that many alternative courses of action or notions that I can think of which fall into these two categories. *And things like that* really does seem to mean 'things *like* that'.

Example 9

[Business: internal planning meeting]

<$2> *So I I'm still kind of chasing them up and [1 sec] getting people to get back to me* ***and stuff*** *so it's not a hundred percent accurate at the moment+*
<$1> *Right.*
<$2> *+like we're, start dates* ***and things like that*** *but -*
<$1> *Right.*

Example 10 takes us back to the academic context, where a supervisor is advising a research student.

Example 10

[Academic: literary stylistics supervision]

<$1> *Mm. But er yeah. Have you done anything on the intentional fallacy and the affective fallacy* ***and that sort of thing?*** *Well that that's quite important because I don't think you can ever demonstrate what a writer's intentions are.*

https://doi.org/10.1017/S0261444819000107 Published online by Cambridge University Press

I understand neither of these 'fallacies', I do admit, but I imagine there is probably not an endless list of fallacies to deal with in literary stylistics and criticism if we're trying to get at a writer's intention. Here the student is expected to understand the accepted context of *fallacies* within the discipline, and the supervisor's advice is limited to that sub-domain within the discipline, *that sort of thing*, and not 'other sorts of things'. *Or whatever* would have suggested quite a different, much wider domain which the student might investigate. We have a similar situation in example 11: *and so on* refers to a set of human organs which the lecturer assumes the listeners will be attuned to; indeed, it would seem oddly inappropriate in this situation to use *or whatever*, which could suggest 'any old organ, it doesn't matter'.

Example 11

[Academic: School of Medicine lecture]

<$1> *Before we go on to the third lecture where we talk about the liver and pancreas* ***and so on****, …*

In the section on the business data, I referred to the different levels of shared knowledge that speakers project (within a company, the business community, the world in general). In the academic world, we have something similar, with, as we have seen, appeals to internal disciplinary knowledge in examples 10 and 11. But we also find references to knowledge that may exist on a wider level. I find example 12 much easier to interpret than 10 or 11.

Example 12

[Academic: archaeology field trip]

<$1> *Okay. So, as you go around actually in spots like this where the actual material is exposed you see the bank is made up of obviously soil and stones* ***and so forth****.*

Grit, bits of rock, sand, all come to my mind, but not a great deal else, and certainly in the actual environment where the field trip is happening, the local geology might limit further what I am imagining the bank to be composed of, and the students are informed that what is there is *obviously* there. Nonetheless, I have an easier time with this than with stylistics fallacies or human organs. Example 13 takes us into a sociolinguistics seminar where discussion is taking place on statistics and variables and averages in research. The lecturer cites an example from TV weather forecasts.

Example 13

<$1> *Erm on the weather forecasts the last couple of years instead of like drawing the symbols for er weather fronts* ***and stuff*** *which they think they're obviously too clever for people to follow. Especially at breakfast telly. They do percentages, don't they? So, they say the percentage of rain and it's like twenty percent seven percent* ***and all this****.*

Although both VCMs limit interpretation to what people normally see on televised weather forecasts, they are relatively easy to grasp precisely because most people are familiar with the conventions of such forecasts. This is one way, and a good one, of teaching: draw analogies with everyday experience. So, we have a parallel with what happens in the business discourse: VCMs range from the narrow to the broad in what they evoke.

However, it is the goals which are often crucially different. Business and academic discourse do share generic activities (presenting information, analysing, reporting) but they also have distinct spheres of

https://doi.org/10.1017/S0261444819000107 Published online by Cambridge University Press

activity. In business, a good deal of the data is to do with planning, selling, negotiating, decision-making and so on. The academic data rests on the purposes and goals of academic institutions: the display of knowledge, the sharing of it and nurturing of critical thinking, building curiosity and creativity, learning through interaction with teachers and peers, apprenticeship and gatekeeping of the world of the professions and of scholarship. The stakes are high in both worlds – you produce and sell or sink, you learn and pass the exams or sink. Both worlds thrive on the building of rapport and on being able to engage in and/or understand talk in all its subtleties. If you do not know what is meant by *production schedules or whatever* and if you do not know what the *liver and pancreas and so on* means, you may be in trouble.

In both worlds, the business and the academic, participants can expect different problems in processing written documents; VCMs will probably not be one of them. In spoken contexts, however, understanding and being able to use VCMs appropriately may mark the boundary between success and failure. In the academic data, it is overwhelmingly the teachers who use VCMs, often as a convenient shorthand to avoid lengthy lists. Students are assumed to know what the lists contain, but, as always, it is not what the lists are projected to contain in the abstract sense of taxonomies or existing lexical categories, but in the ad hoc sense of what is contextually implied by any single VCM, determined by what is relevant to that moment of utterance and that context and purpose. In both cases, business and academic, VCMs present a pedagogical challenge, one which English for academic purposes and business English teaching should not shy away from, especially in courses where speaking and listening are considered important. Chan (2018, p. 29) lists a number of spoken contexts in which business people typically interact: 'participating in discussions, interacting with others, managing the floor, showing solidarity, mitigating face threats, and so on' (herself using a classic VCM indicating the shared knowledge that English for special purposes teachers interested in her paper may be expected to have). Natural data in these contexts abound in VCMs functioning in the ways I have illustrated.

Given the range of VCMs and their variants, it probably makes sense to focus on the most frequent ones and to raise awareness of the subtle differences that can be generated by disjunctive and adjunctive types. As in so many cases of the teaching of speaking, a language awareness approach is to be recommended (see McCarthy & Handford, 2004), since native users may not perceive VCMs to be part of their repertoire in either L1 or L2, such is the liminal zone that so many common features of speaking seem to occupy in the minds of users and in pedagogical materials. Chan (2017) goes further and advocates the use of authentic transcripts of business English in class to raise awareness of interactional features, including VCMs, and reports a discussion with students on one quirky example, 'da da da things like this' (p. 77). She received positive feedback from her students on the use of transcripts, which they felt to be real and authentic. VCMs are at the heart of communication, and, in these goal-driven contexts, they serve to push the talk towards achieving the professional, educational and institutional goals in a very efficient and economical fashion.

## 6. Postscript

The exigencies of a keynote lecture, the warm welcomes and generous introductions and the severe time constraints on the delivery of the lecture and questions and comments following it meant that, inevitably, some material prepared for this talk never made it to the actual event. However, perhaps I may beg of the editor of this journal space to add what would have been part of the total argument if time had allowed.

Despite some internal variation and options to add/omit elements, VCMs display the characteristics of chunks, or lexical bundles, or fixed expressions (the terminology varies among linguists). Some of the core strings occur with notable frequency: *and stuff (like that)* occurs 67 times in one million words of social conversation in the CANCODE corpus, while *or anything (like that)* occurs 60 times. Their nature as chunks is further reinforced by their syntactic behaviour. The most common words around which VCMs are built include *thing, stuff*, the impersonal pronouns *something, anything, everything* and the demonstrative pronoun *that*, as in *He wants to put our cranes* ***and that*** *in the catalogue*. *That* is singular, yet it seems it can refer back to a plural entity (*cranes*). *Thing* is a countable

https://doi.org/10.1017/S0261444819000107 Published online by Cambridge University Press

noun; *stuff* is uncountable. One might expect, therefore, that plural exemplars might be followed by plural VCMs, that countable exemplars might attract VCMs with *thing(s)*, and uncountable exemplars VCMs with *stuff*. On the contrary, VCMs seem to be extremely flexible in terms of what types of exemplars they pair with. The VCM *(and) (all) that sort of thing* follows singular count nouns four times (e.g. *mailshot, serial number*), plural count nouns in 14 cases (e.g. *agreements, labour costs*), and clauses in seven cases (e.g. *if you pick up any mistakes*), along with one case of a non-count noun (*toothpaste*, as seen in example 3, above). The VCM *and stuff (like that)* pairs with singular and plural count nouns (e.g. *brochure, codes*), non-count nouns (e.g. *double entry accounting, software*) and clauses (e.g. *talk to clients*). This is one of the basic characteristics of chunks: they have no internal grammar, and so nothing outside of them can alter their status as fixed (albeit with some internal lexical variation). Malyuga and McCarthy (2018) found a similar syntactic flexibility in exemplar-VCM pairings in Russian professional discourse, along with a reasonable level of translatability, and one might suppose the same to apply to other languages. To some extent all this makes pedagogy easier since syntactic matching seems of little importance to speakers; what matters is the projection of people, things and ideas based on the exemplar(s) and the projection of shared worlds and shared identities. In this respect, VCMs are by no means peripheral; they give us a lens through which to examine the very core of practices in professional and academic life.

## References


Adolphs, S., Atkins, S., & Harvey, K. (2007). Caught between professional requirements and interpersonal needs: Vague language in healthcare contexts. In J. Cutting (Ed.), *Vague language explored* (pp. 62–78). London: Palgrave Macmillan.

Aijmer, K. (2015). General extenders in learner language. In N. Groom, M. Charles, & S. John (Eds.), *Corpora, grammar and discourse. In honour of Susan Hunston* (pp. 211–234). Amsterdam: John Benjamins.

Chan, C. (2017). Investigating a research-informed teaching idea: The use of transcripts of authentic workplace talk in the teaching of spoken business English. *English for Specific Purposes*, *46*, 72–89.

Chan, C. (2018). Proposing and illustrating a research-informed approach to curriculum development for specific topics in business English. *English for Specific Purposes*, *52*, 27–46.

Channell, J. (1994). *Vague language*. Oxford: Oxford University Press.

Cheshire, J. (2007). Discourse variation, grammaticalisation and stuff like that. *Journal of Sociolinguistics*, *11*(2), 155–193.

Clancy, B., & McCarthy, M. (2014). Co-constructed turn taking. In K. Aijmer & C. Ruehlemann (Eds.), *Corpus pragmatics: A handbook* (pp. 430–453). Cambridge: Cambridge University Press.

Cotterill, J. (2007). 'I think he was kind of shouting or something': Uses and abuses of vagueness in the British courtroom. In J. Cutting (Ed.), *Vague language explored* (pp. 97–114). London: Palgrave Macmillan.

Cutting, J. (1999). The grammar of the in-group code. *Applied Linguistics*, *20*(2), 179–202.

Cutting, J. (2006). Spoken grammar: Vague language and EAP. In R. Hughes (Ed.), *Spoken English, TESOL and applied linguistics* (pp. 159–181). Basingstoke: Palgrave Macmillan.

Cutting, J. (Ed.). (2007). *Vague language explored*. London: Palgrave Macmillan.

Dines, E. (1980). Variation in discourse and stuff like that. *Language in Society*, *9*(1), 13–33.

Drew, P., & J. Heritage (1992). *Talk at work: Language use in institutional and work-place settings*. Cambridge: Cambridge University Press.

Dubois, S. (1992). Extension particles, etc. *Language Variation and Change*, *4*(2), 179–203.

Evison, J., McCarthy, M., & O'Keeffe, A. (2007). 'Looking out for love and all the rest of it': Vague category markers as shared social space. In J. Cutting (Ed.), *Vague language explored* (pp. 138–157). London: Palgrave Macmillan.

Handford, M. (2010). *The language of business meetings*. Cambridge: Cambridge University Press.

Jefferson, G. (1990). List construction as a task and resource. In G. Psathas (Ed.), *Interaction competence* (pp. 63–92). Lanham MD: University Press of America.

Jucker, A., Smith, S., & Lüdge, T. (2003). Interactive aspects of vagueness in conversation. *Journal of Pragmatics*, *35*, 1737–1769.

Koester, A. (2010). *Workplace discourse*. London: Continuum.

Malyuga, E., & McCarthy, M. (2018). English and Russian vague category markers in business discourse: Linguistic identity aspects. *Journal of Pragmatics*, *135*, 39–52.

McCarthy, M. (1998). *Spoken language and applied linguistics*. Cambridge: Cambridge University Press.

McCarthy, M., & Handford, M. (2004). 'Invisible to us': A preliminary corpus-based study of spoken business English. In U. Connor & T. Upton (Eds.), *Applied corpus linguistics: A multidimensional perspective* (pp. 167–201). Amsterdam: Rodopi.

Metsä-Ketelä, M. (2012). Frequencies of vague expressions in English as an academic lingua franca. *Journal of English as a Lingua Franca*, *1*(2), 263–285.


https://doi.org/10.1017/S0261444819000107 Published online by Cambridge University Press

Metsä-Ketelä, M. (2016). Pragmatic vagueness: Exploring general extenders in English as a lingua franca. *Intercultural Pragmatics*, *13*(3), 325–351.
O'Keeffe, A., & Adolphs, S. (2008). Response tokens in British and Irish discourse: Corpus, context and variational pragmatics. In K. P. Schneider & A. Barron (Eds.), *Variational pragmatics: A focus on regional varieties in pluricentric languages* (pp. 69–98). Amsterdam: John Benjamins.
O'Keeffe, A. (2004). 'Like the wise virgins and all that jazz' – using a corpus to examine vague language and shared knowledge. In U. Connor & T. Upton (Eds.), *Applied corpus linguistics: A multidimensional perspective* (pp. 1–20). Amsterdam: Rodopi.
O'Keeffe, A. (2006). *Investigating media discourse*. London: Routledge.
Overstreet, M. (1999). *Whales, candlelight, and stuff like that*. New York: Oxford University Press.
Overstreet, M. (2005). And stuff *und so*: Investigating pragmatic expressions in English and German. *Journal of Pragmatics*, *37*(11), 1845–1864.
Overstreet, M., & Yule, G. (1997). On being explicit and stuff in contemporary American English. *Journal of English Linguistics*, *25*(3), 250–258.
Sacks, H., Schegloff, E., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*, 696–735.

**Michael McCarthy** is Emeritus Professor of Applied Linguistics, University of Nottingham. He is author/co-author/editor of 53 books, including *Touchstone, Viewpoint, the Cambridge Grammar of English, Grammar for Business*, *English Grammar Today, From Corpus to Classroom* and *Academic Vocabulary in Use*. He is author/co-author of 113 academic papers. He is co-founder of the CANCODE spoken English corpus and the Cambridge And Nottingham Business English Corpus (CANBEC) spoken business corpus. He has lectured in 46 countries and has been involved in language teaching and applied linguistics for 52 years. His current interests include the spoken English of learners.

**Cite this article:** McCarthy, M. (2020). Vague language in business and academic contexts. *Language Teaching*, *53*(2), 203–214. https://doi.org/10.1017/S0261444819000107.

https://doi.org/10.1017/S0261444819000107 Published online by Cambridge University Press