

doi: 10.15584/sar.2018.15.2.9

Béla LUKÁCS
University of Nyíregyháza
lukacsbela@vipmail.hu

ONE-WORD INTERJECTIONS AS DISCOURSE MARKERS IN FEMALE AND MALE SPEAKERS' ACADEMIC TALK: A CASE STUDY BASED ON THE MICHIGAN CORPUS OF ACADEMIC ENGLISH

Abstract: Through a corpus-based search, the present paper's objective is to attempt to reveal the differentiation of some interjections as discourse markers in the academic talks presented by speakers of both genders. By doing so, I offer a deepened insight into why female and male speakers resort to the use of these interjections. On the basis of an online corpus of the University of Michigan, I present my findings according to several other criteria. The corpus is compiled of a huge set of academic talks by students and staff of the university. Finally, using Jakobson's functions, I classify and further analyse these discourse markers. The data attained through searching this corpus take me a step further in casting some light on the ratio of the interjections used by the two genders, and the possible reasons for the uneven distribution of interjections in terms of female and male use.

Key words: interjections, discourse markers, corpus, genders, Jakobson's functions

Introduction

In my paper, I seek to uncover the "relative frequency with which men and women use certain features of language" (O'Grady 2007: 715) in the case of the interjections considered as discourse markers in academic speech. I assumed that even academic speeches included interjections to some extent, and thus a ratio of their usage between the two sexes might be formed; however, in order to prove and support my assumptions, a corpus-based search needed to be performed to

obtain some data. To this end, we need to examine if interjections may even be regarded as discourse markers.

Interjections are a part of speech which has various definitions (Matamala 2009: 486). Rowe and Levine define interjections as follows: “[i]nterjections are not vital part of the sentence grammatically. They can be removed and not alter the grammatical structure of the sentence. *They are used to express feelings*” (Rowe and Levine 2016: 105, emphasis mine). My selection of the interjections was based partly on their latter statement, and partly on the autonomy of the interjections, i.e. the fact that they “do not enter into constructions with other word classes” (Quirk 1985: 74). Despite that, others, for example Cuenca, quoted after Matamala (2009: 487), consider even the structures such as *Good morning* as interjections.

In his comparison of colloquial and technical language, Petőfi (1973: 254) provides two remarks:

[...] it is obvious that the colloquial use of language presupposes the knowledge of much less detailed and much less precise definitions than the use of particular technical language does. The problem arising here is, that while the requirements concerning the ‘precision’ of the definitions in the technical languages are given within the particular ‘theories’, no such requirements can be formulated with respect to the use of colloquial language.

Petőfi’s (1973) observations also support my observation that colloquial speech events tend to be informal and the speech acts themselves are usually less precisely formulated, which leads to the inclusion of interjections in my research data. Interjections convey important messages relevant to a speaker’s current utterance. They usually have special significance to speakers in terms of giving some extra emotional colouring to their oral communication, and of managing the discourse between the participants in several ways. Here I offer a list of the functions of interjections, which is by no means complete:

1. expressing emotions, feelings;
2. initiating, keeping up or concluding a discourse or topic;
3. dominating the discourse;
4. drawing others into discourse;
5. bringing the discourse around to or steering it away from something;
6. avoiding the discourse;
7. resuming the discourse;
8. encouraging the discourse.

Some of Jakobson’s communicative functions bear similarities to the above taxonomy, as we can see in Table 1:

Classification	Strongest Factor	Function	Examples
Referential	Context	descriptions, contextual information	<i>Our business hours are 9am-5pm, Monday through Friday.</i>
Emotive	Addresser	interjections/expressions of emotional state	<i>Oh, man... Awesome! Whew!</i>
Conative	Addressee	concerned with commanding; vocative or imperative addressing of the receiver	<i>Go on, open it! Shoo. Get out of there. Check this out.</i>
Phatic	Contact	concerns channel of communication; performs social task as opposed to conveying information; to establish, prolong, or discontinue conversation	<i>Hey! Mmmhmmm...How about that? Really? No way.</i>
Metalinguistic	Code	requires language analysis; using language to discuss language	<i>Noun, adjective, code-switching: Water is a non-count noun, right?</i>
Poetic/Aesthetic	Message	involves choosing words carefully; the art of words, often self-reflective	<i>But, soft! What light through yonder window breaks?</i>

Table 1. Jakobson's functions (Tribus 2017: 4)

Tribus's (2017) table of the Jakobsonian functions, complemented by examples, suggests that it is the emotive and phatic functions that primarily belong to the category of interjections as discourse markers. The speaker aims to contribute to the conversation in one way or another (to control it, give some emotional colouring to the discourse, etc.). For this reason, they may be regarded as the speaker's strategies at the same time. My taxonomy is obviously not complete, as it does not include all the possible alternatives, because the kinds of such alternatives are presumably subject to individual selection as well. This is why, in addition to the context and background knowledge, one needs to bear in mind this feature of interjections for a correct understanding of the speaker's intention. This statement is much more relevant to the lexical items with less distinct meanings than to those with clearer, more distinct meanings, e.g. nouns. On the other hand, words such as *uh*, *yuck*, *oh* do not have clear-cut meanings – like, for example, nouns, which indicate “entities” (O’Grady 2007: 722) – and they are rather markers of one’s current emotions.

O’Grady (2007: 712) defines discourse markers as “expressions that bracket utterances in discourse, separating one “unit of talk” from a previous one (e.g. *well*, *y’know*).” This definition covers a huge verbal and non-verbal domain, an enormous set of linguistic entities that may come under the heading of discourse markers. Fraser (1999: 931) suggests the following definition: “a class of lexical expressions drawn primarily from the syntactic classes of conjunctions, adverbs,

and prepositional phrases”. From the perspective of my paper, the key concept is *primarily*, because it implies that other linguistic entities may also be regarded as discourse markers. Therefore, what I wish to argue is that interjections may also act as discourse markers.

In terms of taxonomy, Table 1, offered by Tribus (2017) is similar to the classification proposed by Cuenca, also setting his theory’s foundation through “adapting Jakobson’s classical functions” (Matamala 2009: 487), and who “proposes differentiating expressive, conative, phatic, metalinguistic and representative interjections” (Matamala 2009: 487). Out of the two, Cuenca’s taxonomy is more detailed and specific, thus fitter for my purpose of classifying and analysing the interjections acting as discourse markers. Using Matamala’s (2009) observations, I have prepared Table 2 by way of a short summary:

Expressive interjections	Expressive interjections express the speaker’s feelings, for example, <i>Good God</i> or <i>wow</i> .
Conative interjections	Conative interjections are units used by the speaker in order to produce an effect on the listener, such as <i>please</i> or <i>hey!</i>
Phatic interjections	Phatic interjections show that communication has been established. Two subtypes can be found. The first contains prototypical units like <i>good morning</i> , <i>hello</i> , <i>bye</i> or <i>thanks</i> . The second are the units between phatic and metalinguistic interjections, expressing agreement, disagreement, etc., such as <i>allright</i> , <i>fine</i> or <i>OK</i> .
Metalinguistic interjections	Metalinguistic interjections are used as discourse markers (Schiffrin 1987). They are the key elements which demarcate units of speech in conversational exchanges, e.g. <i>look</i> , <i>well</i> or <i>listen</i> .
Representative interjections	Finally, representative interjections are onomatopoeic interjections, such as <i>miaow</i> , <i>gobble</i> , <i>gobble</i> or <i>cock-a-doodle-doo</i> .

Table 2. Cuenca’s taxonomy of interjections (Matamala 2009: 487-488).

Describing the Michigan Corpus of Academic Spoken English from the viewpoint of my research

In contrast to the features of colloquial speech, described in one of the previous paragraphs of his paper, academic talks – by nature – are well-formulated and coherent verbal products in terms of both their form and content. Usually, they are composed at home, with ample time provided for compiling the required material; still, when it comes to delivering such talks to the students, interjections are almost certain to appear. If we consider these interjections as discourse markers, this statement is proven by the following: “[s]peakers are not generally aware of discourse markers but they are important signals in discourse” (O’Grady 2007: 546).

I performed my search in the Michigan Corpus of Academic Spoken English (MICASE). The corpus includes 152 speech events, totalling 1,848,364 words.

The categories of these academic speech events are highly versatile; searches may be performed on the basis of several criteria, including – besides the speakers’ gender – the interactivity level of discourses, language proficiency level of the speakers (e.g. native, near-native North American English), first tongue, faculty and discipline subjects, age of speakers, position of speakers (student, staff or other), etc. This versatility makes MICASE ideal for my purposes, namely to reveal the ratios between female and male speakers in terms of using discourse markers during their academic speech events.

My search does not cover the issues of the signals of group identity (O’Grady 2007:547). The corpus is, however, fit for that purpose as well, because there is a wide range of first languages to select from.

Selecting discourse markers

As I assume, almost every oral utterance contains some interjections (recorded talks for certain purposes, e.g. audiobooks, are perhaps exceptions), and indeed, I have also found a number of them in MICASE, which is in line with what I presupposed. Table 1 shows the set of perhaps the most frequent interjections applied as discourse markers.¹ The column Interjections as Discourse Markers shows the one-word interjections (queries) I wrote in the browser of MICASE. Female and male speakers’ data are also presented, with the interjections’ absolute frequencies (i.e. numbers of tokens searches brought in MICASE) and PM (frequency per million) values.

My search has returned three lexical items that are even less frequent than the other 17: *nuh-uh*, *nuh-huh*, and *yuh-huh*. Despite the fact that, for example, *nuh-huh* was coined in the 1920s, as seen in *OD* (“Origin 1920s. Imitative. Compare *uh-huh*, *uh-uh*”), and similarly, *nuh* is the product of the 1940s, as indicated in the same online dictionary (“1940s. Representing a colloquial pronunciation of *no* [interjection]”), these words are undoubtedly rare in the MICASE corpus.

Nuh-uh and *nuh-huh* are probably identical in terms of meaning, but they are spelled differently in MICASE. The same seems to apply to *yuh-huh*, *uh-uh* and *uh-huh*. The latter statement seems validated by another online dictionary’s entry “*yuh huh* in agreeance; approval” (*UD*).

As my table shows, there is a considerable difference between the female and male uses of interjections as discourse markers. Female speakers, according to my results, used 844 interjections more than male ones did. This phenomenon might have been caused by the possibility that female speakers wanted to add some emotional colouring to their talks. This is validated by a well-known

¹ The selection is arbitrary: primarily, I attempted to compile it by searching for the items I assumed to be the most frequent.

phenomenon of online communication: women tend to use emoticons more frequently than men do: “Tossell et al. found that females send more messages with emoticons while males use a more diverse range of emoticons” (Chen et al. 2017: 3), or words expressing their feelings. As Xiufang (2013: 2) writes, “[u]sing more adjectives to describe things and their feelings can show that women are more sensitive to the environment and more likely to express their emotions with words, which makes women’s language more interesting than men’s sometimes”.

	Interjections as Discourse Markers	Female Speakers		Male Speakers		Hits	
		Freq.	PM	Freq.	PM	#	PM
1.	<i>ah</i>	89	48.15	124	67.08	213	115.23
2.	<i>argh</i>	1	.54	0	0	1	.54
3.	<i>err</i>	0	0	1	.54	1	.54
4.	<i>hooray</i>	3	1.62	0	0	3	1.62
5.	<i>nuh-uh</i>	3	1.62	0	0	3	1.62
6.	<i>nuh-huh</i>	1	.54	0	0	1	.54
7.	<i>oh</i>	2,268	1,227.03	1,662	899.17	3,930	2,126.20
8.	<i>ooh</i>	54	29.21	30	16.23	84	45.44
9.	<i>oops</i>	36	19.47	20	10.82	56	30.29
10.	<i>phew</i>	0	0	1	.54	1	.54
11.	<i>uh</i>	5,880	3,181.19	10,835	5,861.94	16,715	9,043.13
12.	<i>uh-huh</i>	1	.54	0	0	1	.54
13.	<i>uh-uh</i>	0	0	1	.54	1	.54
14.	<i>um</i>	11,195	6,056.71	6,631	3,587.49	17,826	9,644.20
15.	<i>wow</i>	100	54.1	65	35.16	165	89.26
16.	<i>yay</i>	17	9.2	4	2.16	21	11.36
17.	<i>yeah</i>	5,338	2,887.96	4,773	2,582.28	10,111	5,470.24
18.	<i>yuck</i>	2	1.08	0	0	2	1.08
19.	<i>yuh-huh</i>	1	.54	0	0	1	.54
20.	<i>yummy</i>	2	1.08	0	0	2	1.08
Total:		24,991	13,520.60	24,147	13,063.98	49,138	26,584.59

Table 3. The one-word interjections taken from MICASE

The numbers of the total frequency of each interjection allow us to set up three quite distinctive groups. Their distribution is, however, rather uneven: the most frequent ones are saliently frequent, and the least frequent ones are saliently infrequent. In Table 3, the most frequent interjections are the following: *oh* (2,268; 1,662), *uh* (5,880; 10,835), *um* (11,195; 6,631), and *yeah* (5,338; 4,773). Except for *uh*, all of these were uttered by female speakers more frequently. All these four words are short monosyllabic words, which makes them ideal interjections and discourse markers at the same time, because, although these words are so short, speakers can use them to express themselves in a number of ways in a conversation, e.g. through intonation.

The word with the ultimate frequency in Table 3 is *um*, which is perhaps the interjection most commonly uttered in all kinds of registers, including the register of lectures. Definitely, this statement applies not only to MICASE. *Um* occurs when one is looking for the correct words or expressions – maybe this is why *um* has the highest frequency in my list. The second most frequent interjection, *uh*, may also be associated with a similar context: thinking aloud or trying to find out what to say. On the basis of these sentences, both *um* and *uh* play a significant role in any discourse in terms of cognitive processes.

The second group consists of six interjections: except *ah*, they all were used more frequently by women.

Interjections as Discourse Markers	Female Speakers	Male Speakers	Numbers of Hits
<i>ah</i>	89	124	213
<i>ooh</i>	54	30	84
<i>oops</i>	36	20	56
<i>ooh</i>	54	30	84
<i>wow</i>	100	65	165
<i>yay</i>	17	4	21

Table 4. Five interjections with medium frequency

The rest of them, shown in Table 5, occurred only occasionally:

Interjection	Female Speakers	Male Speakers	Numbers of Hits
<i>argh</i>	1	0	1
<i>err</i>	0	1	1
<i>hooray</i>	3	0	3
<i>nuh-uh</i>	3	0	3
<i>nuh-huh</i>	1	0	1
<i>uh-huh</i>	1	0	1
<i>uh-uh</i>	0	1	1
<i>phew</i>	0	1	1
<i>yuck</i>	2	0	2
<i>yuh-huh</i>	1	0	1
<i>yummy</i>	2	0	2

Table 5. The eleven least frequent interjections

The ratio is uneven here as well, in terms of the usage by genders: no more than 3 out of 11 interjections were uttered by male speakers. The data seen in Table 6 are really surprising to me, because it is the category with the fewest members which exhibits the largest frequency number: metalinguistic interjections occur 34,542 times in the corpus; type 2, phatic interjections, 7 times; and the most numerous category, the expressive interjections – 14,589 times.

The distribution ratios of use by female and male speakers are the following:

expressive interjections: 7,910:6,679;
 phatic interjections, type 2: 6:1;
 metalinguistic interjections: 17,075:17,467.

The above data suggest that female speakers used the expressive and phatic interjections significantly more times than the male speakers did; the metalinguistic interjections, however, were more frequently used by male speakers.

	Type	Female speakers	Male speakers	Total #
Expressive interjections	<i>ah</i>	89	124	213
	<i>argh</i>	1	0	1
	<i>hooray</i>	3	0	3
	<i>oh</i>	2,268	1,662	3,930
	<i>ooh</i>	54	30	84
	<i>oops</i>	36	20	56
	<i>phew</i>	0	1	1
	<i>wow</i>	100	65	165
	<i>yay</i>	17	4	21
	<i>yeah</i>	5,338	4,773	10,111
	<i>yuck</i>	2	0	2
	<i>yummy</i>	2	0	2
		7,910	6,679	14,589
Conative interjections	-			
Phatic interjections, type 1	-			
Phatic interjections, type 2	<i>nuh-uh,</i>	3	0	3
	<i>nuh-huh</i>	1	0	1
	<i>uh-huh</i>	1	0	1
	<i>uh-uh</i>	0	1	1
	<i>yuh-huh</i>	1	0	1
		6	1	7
Metalinguistic interjections	<i>err</i>	0	1	1
	<i>uh</i>	5,880	10,835	16,715
	<i>um</i>	11,195	6,631	17,826
		17,075	17,467	34,542
Representative interjections	-			

Table 6. A Classification of Interjections according to Cuenca’s taxonomy (Matamala 2009: 487-488).

Conclusions

I have searched 20 interjections in the MICASE corpus, and my search has revealed that these words were more frequent in women’s utterances than in male talks. This observation was in line with my assumption, but others have

also come to similar conclusions, e.g. Xiufang (2013), Zhenpeng et al (2017), Witmer and Katzman (1997), to name but a few. “Witmer and Katzman (1997), however, looked at the percentage of the males and females using emoticons within 3000 messages posted on public newsgroups and special interest groups, and found that women were significantly more likely to use emoticons more frequently than men” (Fullwood et al. 2013:3-4).

Nevertheless, it would be unfair to claim that their findings are exclusive, because others give an account of remarkably different observations, e.g. Fullwood et al. (2013: 1), claiming that: “[a]lthough women were more likely than men to use emoticons, there was no difference between the sexes in the range of emoticons used. The fact that men expressed a similar range of emoticons to women implies a general convergence towards female expression in mixed-sex communication contexts”.

Xiufang (2013: 3-4), however, writes that “[c]onventional wisdom leads us to believe that females are more emotionally expressive than males”. When it comes to emotions, in my view, the use of interjections as discourse markers, and that of emoticons, are both motivated by the emotional dispositions of either gender, and, to be more precise, those of the individual speakers. I believe that this is the logic behind the different use of interjections as discourse markers.

Regarding Jakobson’s functions, it is the category of metalinguistic interjections that have resulted in the highest frequency value at the end of the analysis of the interjections. This means that the speakers quite frequently used this kind of interjection to “demarcate units of speech in conversational exchanges” (Matamala 2009: 488). My search, nonetheless, offers the data showing that female speakers used distinctively more interjections than male speakers did.

References

Books and articles:

- Fraser, Bruce (1999) “What Are Discourse Markers?” *Journal of Pragmatics* 31(7); 931-952.
- Fullwood, Chris, Lisa Orchard and Sarah Floyd (2013) “Emoticon Convergence in Internet Chat Rooms.” *Social Semiotics* 23(5); 648-662.
- Matamala, Anna. (2009) “Interjections in Original and Dubbed Sitcoms in Catalan: A Comparison.” *Meta* 54(3); 485-502. [doi: 10.7202/038310ar]
- O’Grady, William, John Archibald, Mark Aronoff and Janie Rees-Miller (2007) *Contemporary Linguistics: An Introduction*. New York: Bedford/St. Martin’s.
- Petőfi, János S. (1973) “Towards an Empirically Motivated Grammatical Theory of Verbal Texts.” [In:] János S. Petőfi and Hannes Rieser (eds.) *Studies in Text Grammar*. Dordrecht: D. Reidel Publishing Company; 205-275.
- Quirk, Randolph, Sydney Greenbaum, Geoffrey Leech and Jan Svartvik (1985) *A Comprehensive Grammar of the English Language*. London: Longman.
- Rowe, Bruce M., Diane L. Levine (2016) *A Concise Introduction to Linguistics*. [Fourth Edition]. New York: Routledge.

- Schiffrin, Deborah (1987) *Discourse Markers*. Cambridge: Cambridge University Press.
- Tossel, Chad C., Philip T. Kortum, Clayton Shepard, Laura H. Barg-Walkow, Ahmad Rahmati and Lin Zhong (2012) "A longitudinal study of emoticon use in text messaging from smartphones." *Computers in Human Behavior* 28(2); 659-663.
- Tribus, Angela C., (2017) *The Communicative Functions of Language: An Exploration of Roman Jakobson's Theory in TESOL*. [Unpublished MA thesis].
- Xiufang, Xia (2013) "Gender Differences in Using Language." *Theory and Practice in Language Studies* 3(8); 1485-1489.

Online sources:

- Chen, Zhenpeng, Xuan Lu, Sheng Shen, Wei Ai, Xuanzhe Liu, Qiaozhu Mei (2017) "Through a Gender Lens: An Empirical Study of Emoji Usage over Large-Scale Android Users." [Available at:] <https://www.researchgate.net/publication/316985738/download> [Accessed 12 July 2018]
- Michigan Corpus of Academic Spoken English (2017) [Available at:] <https://quod.lib.umich.edu/cgi/c/corpus/corpus?c=micase;page=simple> [Accessed 28 June 2018] (*MICASE*)
- Oxford Dictionaries* (2018) Oxford University Press. [Available at:] <https://en.oxforddictionaries.com> [Access 4 July 2018] (*OD*)
- Urban Dictionary* (1999-2018) [Available at:] <https://www.urbandictionary.com> [Accessed 4 July 2018] (*UD*)
- Witmer, Diane F. and Sandra Katzman (1997) "On-line Smiles: Does Gender Make a Difference in the Use of Graphic Accents." *Journal of Computer-Mediated Communication* 2(4); n.p. [Available at:] <https://onlinelibrary.wiley.com/doi/full/10.1111/j.1083-6101.1997.tb00192.x> [Accessed 10 July 2018]