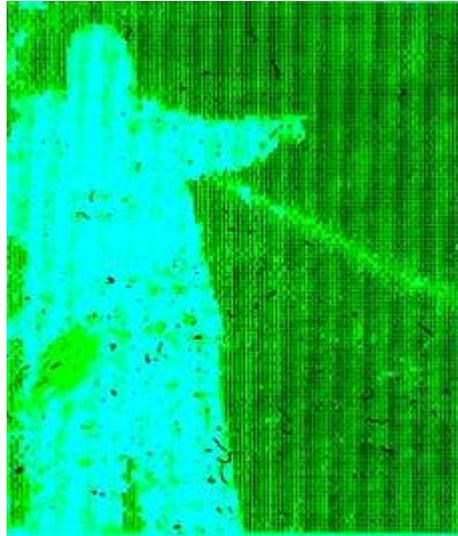


Södertörns högskola | Institutionen för kultur och kommunikation
Kandidatuppsats 15 hp | Engelska C | Vårterminen 2011



Lexis, Discourse Prosodies and the Taking of Stance

A Corpus Study of the Meaning of 'Self-proclaimed'

Av: Mikael Altemark
Handledare: Kristy Beers-Fägersten

Abstract

This study is concerned with the description of the semantic and pragmatic characteristics of the attributive adjective *self-proclaimed*, employing corpus-linguistic methodology to explore its meaning from user-based data. The initial query provided the material from which a lexical profile of the target word was constructed, systematically describing collocational data, semantic preferences, semantic associations and discourse prosodies. Qualitative analysis of sample concordances illustrated the role of the target word in expressing different kinds of meaning-bearing stances. The results demonstrate the importance of context and communicative functionality as constraints determining meaning, determining the discourse prosodies of *self-proclaimed* as one of either negation; accepted-positive and accepted-negative. Further, the analysis of *self-proclaimed* as a stance marker indicates the linking evaluative meanings of extended lexical units to the project of linguistic description of intersubjective stancetaking as a possibly fruitful venue for research

(Keywords: self-proclaimed, corpus linguistics, pragmatics, semantics, extended lexical unit, discourse prosody, stance, intersubjectivity)

Table of Contents

1	Introduction	4
2	Previous research.....	7
2.1	Corpus linguistics and meaning	7
2.1.1	Corpus, collocation, concordance.....	7
2.1.2	The Extended Lexical Unit.....	8
2.2	Evaluation	10
2.2.1	Stance.....	11
3	Methodology and Data	13
4	Results and discussion.....	Error! Bookmark not defined.
4.1	Quantitative results	15
4.2	Extended Lexical Unit	16
4.2.1	Discourse prosody	20
4.2.2	Distribution in text-types	23
4.3	Stance.....	25
4.3.1	Attitudinal stance	26
4.3.2	Epistemic stance	26
4.3.3	Intersubjective stance.....	27
5	Conclusion.....	29
6	References	33
7	Appendix	35

1 Introduction

In recent years, there has been an increasing interest in the study of the evaluative meanings and functions of language. This interest has not limited itself to self-expression in the form of subjective stance, but also touches on the relation-building aspects of intersubjective stance. This perception of language is highly indebted to the works of Bakhtin and Voloshinov (White 2003:92-93) and their dialogical view of communication. Here, language is deprivileged and relativised, expressing a “co-presence in the texts of the ‘voices’ of particular individuals” (Fairclough 2003:41). This means that, through adoption of a certain stance, speakers may report on, modalize, revise and evaluate propositions according to the communicative and rhetorical goals of the speaker in relation to a construed and constructed audience. Thus, *self-proclaimed* is dialogical in that it inherently refers back to a previous utterance, building on or polemicising against a previous proposition while at the same time drawing on communal values in encoding how hearers are expected to respond to the stance taken.

Language users realize stance through the deployment of linguistic resources. Many researchers (see, for example, Hunston and Thompson 2003; Englebretson 2007; Martin and White 2005) investigating these resources and processes have invoked Stubbs’ 1986 challenge in which he concluded that

whenever speakers (or writers) say anything, they encode their point of view towards it ... The expression of such a speaker's attitudes is pervasive in all uses of language. All sentences encode such a point of view ... and the description of the markers of such points of view and their meanings should therefore be a central topic for linguistics (qtd. in Englebretson 2007:17).

Corpus methodologies are prominent in research geared towards locating and describing such evaluative markers and one of the findings is that earlier criteria for identifying instances of evaluation have been limited (Hunston and Thompson 2003; Stubbs 2002). One important aspect has been the possibility of using corpus methodologies to describe differentiation of semantic and pragmatic encoding of evaluation (Channell 2003:40). This aspect can be exemplified by considering that lexical items like *monstrous*, *undoubtedly* or *betray* might all be said to clearly be words carrying a negative evaluation - there is an encoding of attitude on the semantic level. However, a word such as *master-*

mind, is less straightforward – here there is a presence of a pragmatic encoding in addition to semantic meaning. Some people might consider *mastermind* no more than a clear and conventional label one can use to distinguish one person from the next. That is, for them it is used mainly as a descriptive characteristic to define someone adept at planning and executing a project. At the same time, others might respond to it as clearly value-laden, implicating someone as a sinister planner of misdeeds of the terrorist kind. It has been the project of corpus semantics to attempt to map out and explain when and where semantic rather than pragmatic meaning dominates, and vice versa. In the following dictionary definitions of *self-proclaimed*, the same principle is seen at work - meaning of lexis can either be seen as primarily encoded on a semantic level or primarily encoded on a pragmatic level.

a) Merriam-Webster:

Self-proclaimed (adj), based on one's own assertion

b) Collins Cobuild:

1 adj **Self-proclaimed**, used to show that someone has given themselves a particular title or status rather than being given it by other people.

2. adj **Self-proclaimed**, used to show that someone says themselves that they are a type of person which most people would be embarrassed or ashamed to be.

The second entry is taken from a corpus-based dictionary, and exemplifies a definition produced through the examination and analysis of thousands of real-life instances of lexical items in context. The dictionary definitions will be discussed and exemplified further in section 4.1.1. Corpus linguistic methods has shown that our evaluative use of words and phrases is not fully open to conscious reflection, and that these results are often in contrast to results arrived at through intuition (Channell 2003:54). One outcome of this has been the questioning of pure divisions between syntax, lexis and pragmatics, followed by the development of new corpus methods for analyzing meaning. Sinclair came to consider the phrase as the normal carrier of meaning, rather than the word and his model of the *extended lexical unit* (henceforth ELU) has been used extensively in corpus studies of meaning. Findings strongly suggest that evaluative or attitudinal language use is contingent upon co-text while also indicating the often covert nature of such phenomena (Stubbs 2009; Stewart 2010).

The aim of this corpus linguistic study is to investigate the meaning and function of *self-proclaimed* from the vantage point of three linguistic perspectives. First, I will explore the

prevalence and contexts of occurrence of *self-proclaimed*. Second, I will analyze the where and when of the occurrences of semantic and pragmatic encoding of meaning in *self-proclaimed* while also addressing potential additional meanings to those present in Merriam-Webster and Collins COBUILD. Third, I will use the concepts of subjective and intersubjective taking of stance to explore the meaning of *self-proclaimed* on a discourse level.

2 Previous research

2.1 Corpus linguistics and meaning

Corpus linguistics should be understood as a methodology rather than as a sub-discipline of linguistics. Biber (1999) identifies the defining character of the corpus-based approach to language study as consisting of four aspects: (1) It is empirical in its goal of identifying patterns in natural text; (2) natural text is studied in the form of large and principled corpora; (3) the corpora is processed by means of automatic and/or interactive computer technology; and lastly, (4) it is characterized by the mixing of quantitative and qualitative analytical methods (pp. 4-5). In other words, the aim is to utilize records of language performance in order to study repeated patterns “what is usual and typical” and to quantitatively and qualitatively examine this data. This goes beyond just establishing relative frequencies of words, rather, this constitutes an “inherently sociolinguistic” (Stubbs 2002:221) research project into more complex features of language use, opening up an area of empirical study of the function and meaning of words and phrases (ibid. 20).

2.1.1 Corpus, collocation, concordance

When searching for words or phrases in corpora the results are examined in the form of concordance lines. The concordance line is a way of presenting the resulting data of a corpus search, placing “all occurrences in the centre of the screen within a limited span, which is usually an arbitrary number of characters or words in a single line, or within sentence or paragraph boundaries”, allowing for them to be “read as a series of fragments, vertically, from top to bottom” (Stubbs 2009:179) and makes patterns visible and countable. However, when the analyst is faced with hundreds and hundreds of concordance lines, a third order is introduced

with the data subjected to automatic sorting, listing and statistical processing. The data is then more accessible to examination of typical **collocations**, aided by statistical methods for judging relative significance of association (Stubbs 2002:62, 2009:11). Collocation is here to be understood in the Neo-Firthian sense of “directly observable quantity that serves a purely descriptive purpose”. The quantities here referred to are the frequencies by which lexical items co-occur within a given distance (called **collocational span**) of the search word (called **node**) (Evert 2004:pp.15-19) – for example 4 words to the left of the node and 4 words to the right. However, there is also a need for statistical interpretation– if two words are sufficiently plentiful in a corpus, pure coincidence might account for their collocation. Thus, frequencies alone do not tell us what we are interested in, they need not say much at all about how “glued together” two words might tend to be (ibid. p. 20). The question of what kind of statistical method to be used will be further addressed in the methodology section.

When making the case for corpus methods, many linguists have dismissed intuition and introspection as unreliable. In studies of evaluative lexis claims like “without recourse to intuition, qualitative data show clear evidence of where there is an evaluative polarity to an item” (Channell 2003:54) are not uncommon. Such claims might be met with the objection that introspection, in fact, is vital for both search decisions and the following evaluation of resulting corpus data. Nevertheless, it is also true that corpus data introduces a degree of objectivity where there previously was only speculation (Stubbs 2009:117). Though a perfect corpora cannot be designed, bigger and better is luckily always around the corner, and methodology is continually improved; for example, Dilts and Newman (2006) propose algorithmic procedures for establishing positive/negative evaluation based on data obtained prior to corpus-based studies, while Evert (2005) addresses different approaches to the measurement of association scores.

2.1.2 The Extended Lexical Unit

The objectively observable existence of the phenomenon of collocation is central to the model of the **extended lexical unit** (henceforth referred to as ELU) proposed by Sinclair. The ELU is a semantic unit developed to explain the relation between the physical fact of collocation and the (unobservable) “psychological sensation of meaning” (Stubbs 2009:118). This type of meaning is illustrated by Sinclair's famous example of the phrase *set in*. He showed that *set in* consistently tended to collocate with words *rot*, *decay*, *bitterness*, which all exhibit negative or unpleasant qualities (Stewart 2010:11). Sinclair argued that this showed the importance of

co-text for meaning, and prompted him to conclude that meaning is a phrasal phenomenon (Dam-Jensen and Zethsen 2007:1616).

Inside the model of the ELU, Sinclair used the term **discourse prosody** to describe the compounded evaluative meaning of *set in* and its typical collocates. In following studies of discourse prosodies, some of the common characteristics are defined as being evaluative or attitudinal, hidden, contingent upon co-text, best revealed by computational methods and considered the functional choice linking meaning to purpose (Stewart 2010:30:160). In other words, it is understood in terms of “the real world: *where, when, why, and to whom*, [the lexical item] means what it does” (Gill 2009:4).

It has become clear that there is a problematic diversity in terminology undermining replicability and reliability of corpus research. This confusion can be traced to the diverging - but equally influential - perspectives on discourse prosody of Sinclair and Louw (for an exhaustive and interesting discussion, see Stewart 2010). For Sinclair, discourse prosody is derived from analysis of the node, its collocates and concordances together and the resultant description of function in discourse context. Louw instead focuses on what he terms *semantic prosody*, which corresponds to the interaction of co-occurrences of node with semantic preferences and the resulting “aura of meaning” that colors node and phrasal unit (Gill, 2009:4). Hoey (2005) addresses this terminological confusion, proposing the subsumption of both semantic preference and discourse prosody into a new common category. However, for Phillip (2009) this is a lumping together of functionally distinct abstract concepts. In this essay, Phillips *semantic association* will only briefly be covered, illustrating the polar connotative and attitudinal meanings derived from the semantic preference of lexical items. Thus, discourse prosody is reserved for modal pragmatic effects and function alone, following the approach of Stubbs and Sinclair.

Stubbs underlines that in the ELU, we have a well-specified linguistic unit, in which constituents and their relations are defined. Combining Phillips and Stubbs take on the concepts originally developed by Sinclair, a fully defined ELU model would consist of the following eight parts:

- 1) **Collocation**, the relation between the node and individual collocates (observable and countable).
- 2) **Colligation**, abstract grammatical categories co-occurring with the node.
- 3) **Semantic preference**, the co-occurrence of the node with words from certain semantic fields (giving evidence for topic or theme)

- 4) **Semantic association**, the description of connotations not necessarily shared by all members of a speech community
- 5) **Discourse prosody**, the hypothesizing on the functional choice linking meaning to purpose.
- 6) **Strength of attraction**, describing the significance of the relation between node and collocate.
- 7) **Position and positional mobility**, which go beyond co-occurrence to address relative and variable positions (*fight for fit* occurs less often than *fit for fight*).
- 8) **Distribution**, the analysis of restriction of discourse prosodies and semantic sets to registers or specialized text-types.

Thus, linguistic acts and human action pregnant with meaning is addressed here in three ways. 1 and 2 provide us with the means to describe “the observable behavior of a social group” (Stubbs 2009:125) while 3, 4 and 5 address the subjective meanings of speakers and hearers in context. Finally, 6, 7 and 8, address the question of probability, position and the extent to which the ELU applies in textual context.

2.2 Evaluation

One of the major findings of corpus linguistics is the realization that there is a strong connection between the lexical system and evaluation (Stubbs 2002:242-243) and this has brought with it implications for a number of fields, including critical discourse analysis (Baker, Griellatos et. al 2008), cross-cultural studies of stance moods (Precht 2003) and register-distribution of grammar (Conrad and Biber 2003). Evaluation has been systematized in a number of frameworks, and notably Hunston and Thompson (2003) operationalize it based on its role in discourse, resulting in a tripartite division of function: a) expressing the opinion of a speaker; b) constructing relations between speaker and hearer; and c) organizing a discourse (p. 6).

The first two aspects are of most direct relevance to a study of *self-proclaimed*. Acts of evaluation tie in to the general or specific communal value systems underlying the opinion expressed. Instances of evaluative language in text are evidence not only of a speaker's opinion on propositions or entities. In fact, the argument an individual makes for something being good, bad, credible, unlikely or novel is "only convincing, or even meaningful, when they contribute to and connect with a communal ideology or value system" (Hyland 2005:175). In

its relation construction function, evaluation might instead be used as a manipulative device, "exploited by speakers to express evaluative meaning covertly" (Hunston and Thompson 2003:8, 38) or hedge utterances ("our results suggests" would be an example of evaluative hedging encountered in an academic paper).

2.2.1 Stance

Conrad and Biber (2003:58) use the terminology of stance to analyze linguistic items conveying "meanings associated with the writer's attitude" towards propositions and entities. These evaluative meanings are grouped into three categories. First, that of **epistemic stance**, which indicates evaluation on the certainty or source of information. Second, we have **attitudinal stance**, which indicates the emotional or judgmental assessment of the speaker. The third category, **style stance** (indicating the manner in which the proposition is made) will not be addressed in this essay. Here we return to the Collins COBUILD definitions of *self-proclaimed*. In the first instance, *self-proclaimed* is "used to show that someone has given themselves a particular title or status rather than being given it by other people." This can be interpreted as an example of relation-constructing stance; it is used to modalize the speaker's view on the epistemic status of the source of information. In other words, *self-proclaimed* is used by the speaker to adjust the degree of his or her commitment to probability or likelihood of a proposition. In the second instance, *self-proclaimed* was "used to show that someone says themselves that they are a type of person which most people would be embarrassed or ashamed to be". Here, there is a more clearly attitudinal evaluative stance involved, a reflection of the speaker's view on what is good and what is bad (Hunston and Thompson 2003:4).

However, viewing evaluation as a concept primarily for analyzing subjective stance does not sufficiently take into account the fact that "the basic reason for advancing an opinion is to elicit a response of solidarity from the addressee" (Hunston and Thompson 2003:143). This intersubjective dimension of evaluation is explored by White (2003) and Martin and White (2005). They map the resources available to speakers in their positioning of themselves towards a previous proposition and in "[anticipating] the responses of actual, potential or imagined readers/listeners" (2005: 92), and link meaning to context and rhetorical effects rather than to specific grammatical forms. In this essay, I will only address the most basic components of the system proposed by Martin and White: the category of **heteroglossia** and its component functions of **dialogical contraction** and **dialogical expansion**. In contrast to **monoglossic** statements, heteroglossia implies recognition of the existence of multiple voices and

views on a proposition. However, just because other voices are acknowledged, it does not mean that the intersubjective functionality of heteroglossic utterances is constant, but rather, we can find both active openings for alternative voices (dialogical expansion) as well as “acts to challenge, fend off or restrict the scope” (White 2003:262) of this possible multiplicity (dialogical contraction).

Such a conception of evaluation and stance, then, strongly emphasizes the social rather than the personal, a process of linguistic engagement with socially determined values rather than truth-values. However, inside the framework of corpus methodology, a workable linguistic analysis is problematic, because

stance is a meaning, a type of meaning, or several types of meaning, rather than a form ... identifying stance entails more than simply locating those forms [thus] interpreting the role of stance in discourse entails a deeper understanding of the discourse as a whole than can be obtained from looking at the immediate co-text of an individual lexical item (Hunston 2007:27).

Thus, the highly contextual determination of meaning is an important aspect of stance, and all the more important because we need to be able to address intersubjectivity when analyzing meaning. Hitherto, this has not been translated in a working program for research (Du Bois 2007:140) and thus needs to be addressed in methodology and analytic frameworks. This project constitutes a huge undertaking, and is clearly outside the scope of this essay. Nevertheless, in following through an exploratory description of the basic categories of dialogic contraction and expansion, this essay aims to contribute a more complete grasp of the fascinating interaction between lexis, meaning and function.

3 Methodology and Data

Modern computational technology makes possible the use of powerful tools for processing large databases of natural text, revealing lexical, semantic and pragmatic patterns. Today, there are very large on-line corpora available for public use. One of these, the *Corpus of Contemporary American English* (Davies, Mark. (2008-)¹ was chosen for the collection of data.

The preliminary query run was for *self-proclaim** in all available text genres - spoken, academic, magazine and news. The * is used to catch any potential significant variations of the **lemma** *self-proclaimed* – that is all the realizations of a word. For example the lemma *take* can be realized as *takes, took, taking*. Even though no lemmatic variation worthy of note, the preliminary run instead revealed some technical limitations of COCA; a substantial number of concordance lines from the spoken register could not be displayed. Because of this, results from spoken registers were thereafter excluded.

A quantitative analysis of the results for the query *self-proclaimed* then established a total frequency of occurrence in the corpus, the relative frequential distribution in different genres, as well as in the least and most common text-types. The second stage in the process was the construction of a profile of the ELU. Firstly, this allows for summarizing and systematically presenting information in a coherent way, in order to assist the analyst in uncovering significant patterns. Secondly, it allows for comparisons, with the purpose of starting from observed word-forms and working towards hypotheses on communicative function (Stubbs 2002:84; 2009:124). The modified ELU profile which is used in this study excludes two of the levels described in section 2.1.2. As the node-word always appeared as a premodifying adjective, the colligative relations were not considered relevant. In addition, the construction of a positional frequency table was left out due to constraints of space.

The search query consisted of the target word *self-proclaimed*, with a collocate span of 4:4. In order to assess the significance of the relation between node and collocate, a numeric

¹ (COCA): 425 million words, 1990-present Available online at <http://www.americancorpus.org>.)

value was calculated using the **mutual information** (MI) score.² The MI test is one of the most applied tests in corpus linguistics, and is a concept borrowed from information theory. MI scores are a way of grading the non-homogeneity of the observed values and compare this numerical value to a table of the frequencies which could be expected in typical language use. In other words, what is presented is the coefficient of the association strength - with a value of 0 if there is a complete independence between the lexical items, with increasing value indicating a stronger significant correlation (Evert 2005:77). In order to give due weight to frequent cases, a **cut-off point** is set, setting a threshold which excludes the “noise” of collocates with low significant cooccurrence (Stubbs 2003:84). In this case, the cut-off point set to 3.0, the default value in COCA. Patterns and trends were then identified by inspecting the resulting data in the form of lists ordered by total frequency of occurrence as well as MI value strength.

Semantic preferences were arrived at by examining the initial 20 most significant collocations and their concordances. Then the collocates were grouped into subsets according to common semantic features. These results were then checked for reliability by examining an additional 100 random concordances. Randomly selected concordances of the top 20 collocates were excluded in this process.

Next, together collocate and semantic preferences constituted the starting point for speculation on possible semantic associations activated. Categorization and identification of discourse prosodies was done via manual inspection and analysis of the concordances of the 20 most significant collocates. Further processing consisted of grouping results of semantic preference and discourse prosodies according to text-type and subjecting these to counting. This then enabled the analysis of semantic preference, discourse prosody and register-type distribution.

The third level of analysis follows the frameworks of Conrad and Biber (2003), White (2003) and Martin and White (2005). Here, a number of example concordances extracted from the results obtained in the corpus search are used for the qualitative exploration of different aspects of stance.

² The variant of MI used by the COCA is the following algorithm:

$$MI = \log \left(\frac{AB * \text{sizeCorpus}}{A * B * \text{span}} \right) / \log (2).$$

3.1 Quantitative results

The following table gives us an overview over the total frequencies of the target word in the COCA, presented in descending order of rate of occurrence.

Table 1. Tokens of *self-proclaimed* in the COCA

COCA	Frequency	Total Words	Rate of occurrence
News	220	86670479	2.54 per million
Magazine	182	90292046	2.02 per million
Academic	127	85791918	1.51 per million
Spoken	106	90065764	1.18 per million
Fiction	38	84965507	0.45 per million
Total	673	437785714	1.54 per million

Self-proclaimed is not a commonplace item in the COCA, with only 673 tokens, giving it an occurrence rate of just over 1.5 tokens per million words. It is most common in the news category, followed by magazine publications. Academic and spoken registers end up as third and fourth respectively. In fiction, *self-proclaimed* is a very uncommon word with only 38 tokens. It is interesting to note that the rate of occurrence of *self-proclaimed* is slightly more than 3.5 times higher in written than in spoken text.

Further, *self-proclaimed* is strongly associated with written public discourse in the form of news editorials, special interest and news/opinion. The three text types with the highest rate of occurrence are editorial news (6.21 per mil), African-American interest magazines (5.08 per mil) and news/opinion magazines (3.57 per mil). The rate of occurrence is remarkably low in natural sciences and in fiction in general. The three text types with the lowest frequency of occurrence consist of academic publications relating to science/technology, movie scripts - both 0.22 per mil - and medical journals (0.18 per mil).

3.2 Extended Lexical Unit

The first step of constructing an ELU profile consists of establishing collocational frequencies and significant relation strength to the node-word. Table 2 below presents frequential data to the left and MI values to the right.

Table 2. Collocates of *self-proclaimed*

Ranked by frequency			Ranked by MI value				
	lexis	freq	MI		lexis	freq	MI
1	capital	12	4.86	1	anarchist	5	9.87
2	experts	10	4.98	2	mastermind	5	9.66
3	king	10	4.25	3	prophets	4	8.19
4	leaders	6	3.47	4	nut	5	7.25
5	boy	6	3.02	5	freak	4	7.23
6	anarchist	5	9.87	6	prophet	4	6.92
7	mastermind	5	9.66	7	socialist	5	6.51
8	nut	5	7.25	8	rap	4	6.13
9	socialist	5	6.51	9	lover	4	5.63
10	expert	5	4.70	10	experts	10	4.98
11	prophets	4	8.19	11	capital	12	4.86
12	freak	4	7.23	12	expert	5	4.70
13	prophet	4	6.92	13	liberal	4	4.31
14	rap	4	6.13	14	king	10	4.25
15	lover	4	5.63	15	artist	4	3.63
16	liberal	4	4.31	16	mission	4	3.61
17	artist	4	3.63	17	leaders	6	3.47
18	mission	4	3.61	18	leader	4	3.38
19	leader	4	3.38	19	member	4	3.05
20	member	4	3.05	20	boy	6	3.02

A clear pattern is present in the frequential table with a sharp spike of high frequencies, followed by an exponentially descending curve of distribution. The frequently most common collocates, *capital* (12), *experts* (10) and *king* (10) represent a tight cluster, with almost 30% of total occurrences.

However, both the singular and plural form of the collocates *leader*, *prophet* and *expert* are present in the results, and a merging together of these shows an even clearer the tendency. The exponential drop-off is sharper and the 5 most common collocates now represent 50% of all occurrences. This might suggest their co-occurrence with *self-proclaimed* as being parts of

conventional or idiomatic phraseology. The remaining majority of the collocates then constitute a slowly descending tail of 12 tokens making up the remaining 50% of the total.

Moving on to MI-value-sorted results, it is interesting to note that almost half of the infrequent collocates here score very high. Among the top three we have *mastermind* and *anarchist*, with extreme MI values of almost 10, followed by *prophet* (8.19). A group of collocates occurring only 4-5 times also score a very high MI - between 5.55 to 7.28. If *prophet* was combined with *prophets*, the whole upper field of very high MI values would be dominated by infrequent collocates. Further, we can observe that the three most frequent collocates, *capital*, *king* and *expert(s)*, are all grouped into a field of moderate MI values ranging from 4 to 5. The MI values of *leader* and *leaders* align almost exactly.

Semantic preferences entails examining collocates in relation to each other and grouping those with semantic commonalities into semantic sets.

Table 3. Semantic preference

Semantic set	cause/ideal	enthusiasm	authority	centrality	skill
collocates	anarchist (5)	nut (5)	leaders (6)	capital (12)	experts (10)
	socialist (5)	lover (4)	leader (4)	king (10)	expert (5)
	liberal (4)	freak (4)	prophets (4)		mastermind (5)
	mission (4)		prophet (4)		artist (4)
	member (4)				
Sum	22	13	18	22	24

The most common semantic set is *skill*, which includes a number of professional competencies, followed by 22 tokens each for positions of supremacy/centrality and causes/political ideals. Positions of authority in political, military and religious contexts is placed fourth with 18 tokens while identities taking an enthusiastic stance toward a subject or field lag behind at only 13 tokens. The largest lexical variation can be found in *cause/ideal*. The two collocates – *boy* (6) and *rap* (4) - did not fit into any of the semantic sets but will be addressed later.

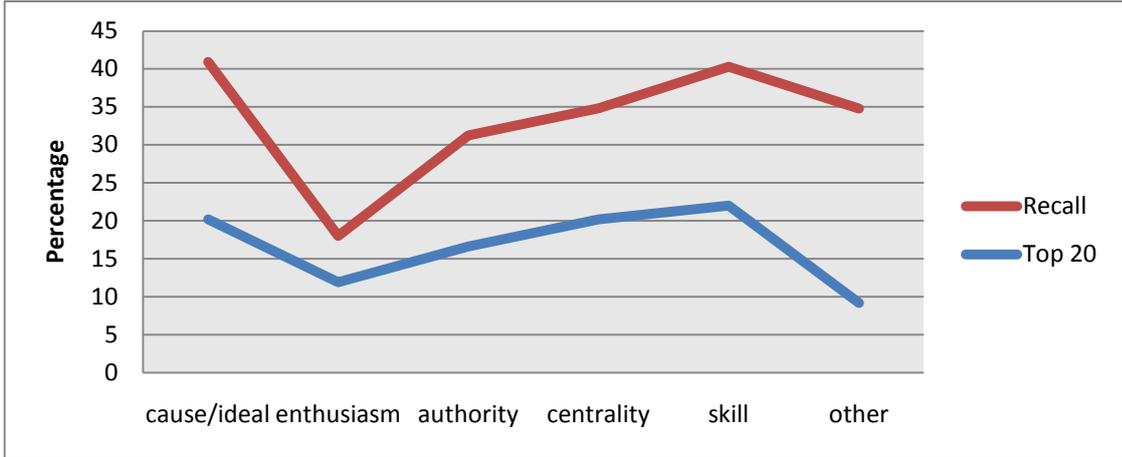
Precision and **recall** are terms used in information theory to assess reliability. The first measures relevance of the data already present, in this study assured through calculation of MI values. However, recall relates to the measure of relevant data missed or not present. In Stubbs’ model, this is solved by the retrieval of a number of random concordance lines in or-

der to catch varied lexical realizations not present among the most common collocates (Stubbs 2002: 70, 89). This was operationalized in the examination of 100 random concordance lines. Of those, 18 already occurred in the initial list of significant collocations and therefore excluded. The results are here illustrated by example phrases and presented in descending order of frequency (quotation marks in the original): 17 instances of concordances pointed to people being self-proclaimed followers of a cause or adherents of a political ideal and exemplified by: *environmental activist*; “*Black Panthers of Hip-Hop*”; “*neo-Marxist*”; *guardian of America’s thinking*. 15 persons were described as self-proclaiming themselves to possess skills or special competence in some way: *spies and secret agents*; *psychic*; *Mr. Fix-it*; *sorcerer to the stars*. 12 entities are described in terms of claiming to have political, military or moral authority. These include *spokespeople*; “*candidate for change*”; *sphere of influence*. Another 12 relate to people and entities with absolute properties; someone being the best, the centre, the origin and so on. Examples are: *poetic father of French romanticism*; *home of the blues and birthplace of rock-and-roll*; *best-dressed player in the NBA*. The fewest instances found, with only 5 occurrences, were the self-proclaimed enthusiasts termed: “*river guy*” or “*dirt guy*”; *golf addict*; *turkey fanatic from Illinois who now resides in Florida*.

With the additional understanding gained from the random concordance lines, I am now able to address the collocates unaccounted for in table 3. As for the collocate *boy*, the concordances turn out to refer to such differing entities as *mamma’s boy*, *inexperienced boy* and *bad boy*. *Bad boy* could possibly fit into the category of identities relating to skill – a bad boy is someone involved in the criminal profession, but is also used to describe people as being hip or in. Without knowledge about the larger context, it is hard to establish pragmatic meaning - just like in the case of the more idiosyncratic *self-proclaimed radster Saucer Boy*. Moving on to *rap*, there is a duplicate result in COCA for “*grandfather of rap*” as well as a Japanese artist talking about *his bond with rap’s self-proclaimed “Louis Vuitton don*. These two might be used as metaphoric expressions for people having preeminent influence in the field of culture, making them fit neatly into the semantic set of centrality.

In order to gain a clearer picture of the precision in the initial creation and grouping of collocates, the percentage of the total for each semantic set for both the top 20 and the 100 random concordances was compared. Table 4 shows that they largely follow same tendencies.

Table 4. Comparison of recall and top 20 percentage of total



In total, about 25% of the recall collocates did not fit into any of the semantic sets established earlier, which is much higher than the 10% of the 20 top collocates. It might be argued that the more typical collocates like *expert* and *leader* (which here were removed from the classification of random concordances) are more common because of their idiomatic quality. We might here consider Stubbs reiteration of Searle’s argument that “within the class of idiomatic sentences, some forms tend to become entrenched as conventional devices for indirect speech acts” (qtd. in Stubbs 2002:59). However, as we can observe from table 4, even if more collocates were unclassifiable; the general trends are confirmed, if averaged out.

The following table exemplifies five polar attitudinal values - borrowed from Martin’s (2003) appraisal framework for classifying evaluative judgements (p. 156) – illustrating different possible semantic associations in relation to *self-proclaimed* and semantic preferences exhibited.

Table 5. Semantic association

Aspect	Positive	Negative
Normality	Avant-garde	Peculiar
Capacity	Capable	Incapable
Veracity	Authentic	Deceitful
Tenacity	Inspired	Arrogant
Propriety	Good	Bad

Semantic association describes “secondary meanings - associations, evaluation and connotations - which are fundamentally psychological in nature and therefore not necessarily shared by all members of a speech community” and thus they may vary from one individual to another depending on circumstance (Gill 2009:4-5). In other words, one person’s self-proclaimed anarchist might be a freedom-fighter while the same person is another one’s terrorist. This can also be exemplified by the *self-proclaimed typewriter-nut* found in the data. Depending on person, the resulting connotations of the phrase might be either associations of someone quite peculiar or relating it to inspiring acts of proudly displaying dedication, no matter what the community thinks. Semantic association, then, does not constitute the functional character of an ELU, but the individual psychological affects and attitudes the language user might contribute towards it.

3.2.1 Discourse prosody

As discussed previously, discourse prosodies are evaluative, hidden and contingent upon context. Identifying discourse prosody, intuition is used to formulate “a hypothesis about the conscious and intentional social action.” However, this is not straightforward and even though “the hypothesis is based on observable data [...] improved formulations - and counterexamples - are always possible. (Stubbs 2009:126).” With this established, the three types of discourse prosodies abstracted from the data follows, exemplified with two concordances each. The discourse prosodies identified are then compared and discussed in relation to the dictionary definitions brought up in introduction.

3.2.1.1 Denial

Under the term **denial** we find people evaluated as acting in a way not in agreement with the conventionally positive values of claimed identity, and the discourse prosody carries with it a

force of negation of such a claim. This discourse prosody corresponds to the first definition given in Collins COBUILD, exemplified in the following examples (1-2)

- (1) said to me, " ADD has become a growth industry. " Not every self-proclaimed expert knows ADD from ABC. For instance, depression can cause someone to be
- (2) of Supervisors. " Our city politics have become a joke, and it is self-proclaimed leaders removed from reality like Alioto who hope to perpetuate it. # DIETRICH von

In (1), *self-proclaimed* is the core of an ELU exhibiting a discourse prosody of denial via unfavorable ratification of identity claim. Because ADD is hard to diagnose, it is something best left to real (medical) experts, who in addition (unlike *self-proclaimed* ones) do not see ADD as a goldmine. This pattern is repeated in (2). City politics is not supposed to be a joke, and leaders are not supposed to be removed from reality. The discourse prosody is effectively one of denial of the claim to leadership because the self-proclaimer is evaluated as an illegitimate or un-recognized leader.

3.2.1.2 Accepted-negative

In the third and fourth example, a discourse prosody corresponding to the second definition given in Collins COBUILD is present i.e.

- (3) allies defeat the advocates of central planning that today it is difficult to find a self-proclaimed socialist or communist. Those sympathetic to central planning now promote their policies under
- (4) on the Greg Dobbs show, where he could count on softball questions from a self-proclaimed liberal host. # Mr. DiGiacomo is understandably afraid to face tough questions about why

Example (3) above illustrates a discourse prosody of **accepted-negative**. The claim to identity is never challenged, but rather taken in and charged with negative evaluation. The allied victory over central planning is set up as resonating with communal values, indicating the negative evaluation of those who would advocate authoritarian non-market solutions. For the speaker, the people who still denote themselves socialists or communists confirm his or her negative evaluation by the very fact that they deviously try to slip their policies in through the back door. The same goes for (4), where the political orientation of a talk-show host is ac-

cepted, but then used to show that liberal is to be understood as negative. It is suggested that it is only to be expected that a talk-show host who claims to be a liberal would be “soft” instead of critical.

3.2.1.3 Accepted-positive

In this section, a previously undescribed discourse prosody of **accepted-positive** is presented, exemplified in the following concordances:

- (5) have adjacent offices, each equipped with a portable crib and changing table. A self-proclaimed neat freak, Paxon says he's the more meticulous diaper changer, although he
- (6) executive who keeps finding time for sports. He is, after all, a self-proclaimed "sports nut. " Undaunted by a money-losing stint as owner of the Attack

Being a meticulous diaper changer and having good equipment is certainly positively in a magazine like Good Housekeeping – the source of example (5). Even though a qualifying *although* is introduced, there is no serious challenge to the claim of being a neat-freak. Thus, the functional meaning becomes one of positive acceptance. In (6), the identity of sports nut is dependent on being both an undaunted executive and a person who still finds time for sports, a claim readily reported and accepted. This accepted-positive discourse prosody is not described in either of the two dictionaries consulted. The corpus data, however, suggests that it is too common and consistent to dismiss as expectable aberrations.

3.2.1.4 Other

Under this category, we have concordances either lacking discourse prosody or in which the co-text fails to provide the minimal information needed to say anything for certain about the present or non-presence of discourse prosodies. This is an example of usage corresponding to the Merriam-Webster definition. Two examples are used to illustrate instances of this:

- (7) Today the self-proclaimed Granite capital of the world is filled with granite homes, granite signs, granite
- (8) 're in total agreement. -Ed. BOURBON OF PROOF I read your piece on self-proclaimed radster Saucer Boy (The Truth, October 2008), and I have a

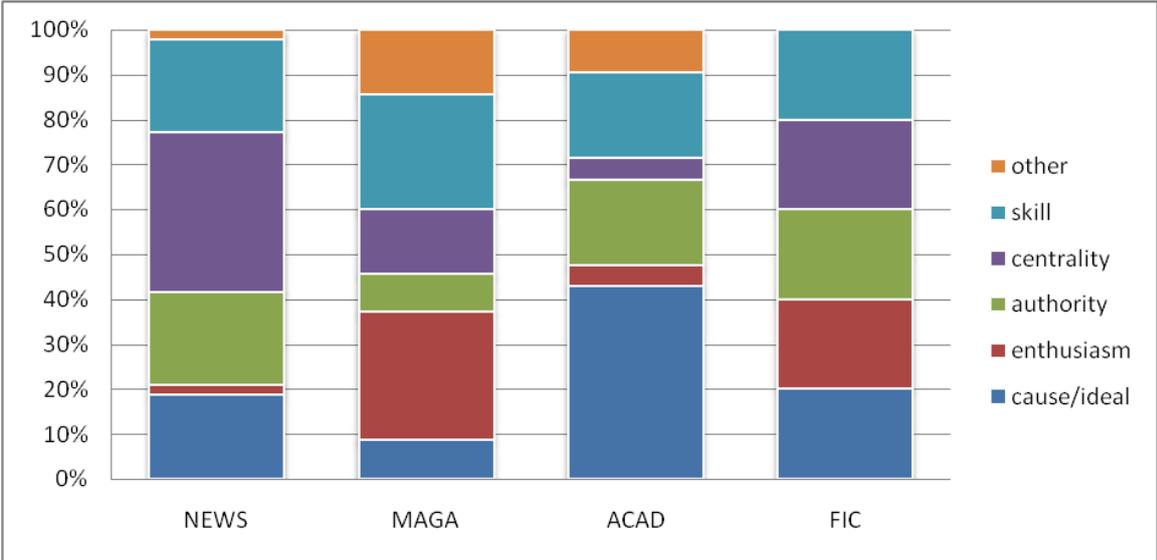
The other category thus gathers cases of *self-proclaimed* when it is either used in a neutral reporting manner or when the concordance line is unable to give any significant clues to a clear evaluative function. Example (7) exemplifies the first possibility, as it contains no indication of a certain attitude being expressed with regards to either granite or the fact that the entity proclaims itself the Granite Capital. (8), on the other hand, is an example of a concordance line characterized by creative or metaphorical language. Here, an understanding of the meaning of the propositions is highly reliant on previous real-world and pragmatic knowledge. It also shows the technical limitation of an un-expanded concordance line, because while some evaluation is expressed, the format cannot capture it.

After classifying all 109 concordances of the 20 most common collocates, it emerges that negative discourse prosodies tend to dominate. Grouped together, the denying and negatively evaluating discourse prosodies of 4.2.1.1 and 4.2.1.2 make up slightly more than 45% of the total. *Positive* discourse prosodies account for about 32% while *other* (unclassifiable or lacking prosody) make up just over 25%.

3.2.2 Distribution in text-types

In the last stage of constructing the lexical profile, analysis of distribution in text-types shows how widely our description applies. Table 6 makes it possible to discern some clear trends in regards to how often certain semantic fields are addressed in certain texts.

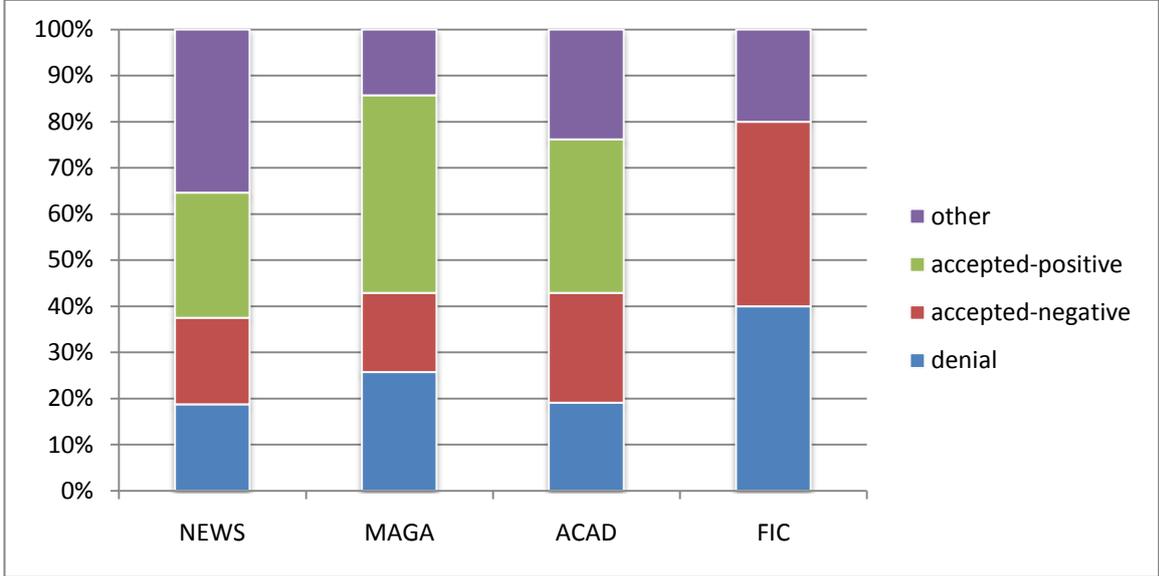
Table 6. Register-type distribution of semantic sets



There is a very even distribution of the semantic set relating to skills in all registers. This might be accounted for by the dominance of a highly conventional phrase like *self-proclaimed expert*, displaying its use in all represented publication types. In magazines, there is a large preference for semantic sets relating to *enthusiasm*, with a corresponding lack of it in news. Academic publications, surprisingly, features a domination of *cause/idea*. On further examination, it turns out that academic titles relating to foreign policy and art are responsible for many of these occurrences. The relative ordinariness of the *other* category in magazines can be attributed to the large variation in lexical realization of phrases. Many have metaphorical or other creative characteristics which makes them hard to classify from the concordance line. For fiction, there is an almost perfectly random distribution, perhaps indicating a tendency in literary works to equally address the aspects of human existence encompassed in the semantic sets.

Below, the patterns relating to discourse prosody distribution in either news, magazine, academic or fiction type genres are visualized.

Table 7. Distribution of local prosodies



In the context of the news category the highest total ratio (about 35%) of concordances were either unclassifiable or lacking discourse prosody. This might be attributed to either news-type registers being neutrally reportive stories and/or often requiring a high degree of pragmatic background knowledge for successful decoding. Almost as common are discourse pro-

sodies of *accepted-positive*, at about 30%, with the rest evenly distributed between negative criticism and denial.

In the magazine type publication register, *accepted-positive* is the dominating tendency, representing over 40% of use. This could be related to the common life-style type stories portraying individuals with unusual interests or narratives as interesting and successful. In line with this reasoning is the relative scarcity of negative criticism (15%). However, magazine type publications also have the highest ratio of denial out of all publication types, indicating that discourse prosodies are divided between celebrating and blocking identity claims. Also noteworthy is the very low occurrence of the *other* category. This could indicate a discourse prosodic scarcity and be taken to support the hypothesis of a very strongly polarized division of evaluation in magazine type publications.

The most balanced distribution of categories was found in the academic articles featuring the node word. Even if *accepted-positive* stands out, the margin is lesser than 10%. Compared to the common conception of academic type articles as non-modal, this is an interesting result. However, acceptance might tend to involve evaluations of epistemic status (see section 4.3.2). This resonates with findings of Conrad and Biber (2003) on the relative distribution of epistemic stance in academic registers.

There were many instances of concordances intuitively clearly being evaluative, but for which a discourse prosody could not be established. This is not surprising since meanings are, after all, distributed “across more than one word” and even across a whole discourse segment. Further, they “often express speakers’ relations to other people, and may depend on assumptions and worldviews” (Stubbs 2002:105) and thus reside too far on the pragmatic end of the spectrum of meanings for the concordance to make sense to the analyst. While corpus linguistics is concerned with probabilities and typicalities, the results also indicate the need for corpus methodology to be assisted by additional qualitative in-depth methods of analysis like those developed in frameworks addressing meaning as expressed in stance.

3.3 Stance

Under this heading, two of the stance categories defined by Conrad and Biber (2003) – epistemic stance and attitudinal stance, as well as the intersubjective stance functions introduced by White (2003; 2005), will be considered. When a person takes an epistemic stance, they are indicating their position on the source or quality of information – “to what degree is this

statement likely to be true”? Attitudinal stance expresses the subjective evaluative opinion of the speaker towards a proposition. In the intersubjective stance, the concern is with addressing and evaluating the proposition of a previous speaker, as well as with achieving certain rhetorical effects by tending to, in the words of Precht, the “relationships between the speakers, the relative status of the speakers, and the presentation of self” (2003:241).

3.3.1 Attitudinal stance

Attitudinal stance relates the position of a speaker towards a proposition or state of things, being exemplified by the following concordance:

- (9) on the Greg Dobbs show, where he could count on softball questions from a self-proclaimed liberal host. # Mr. DiGiacomo is understandably afraid to face tough questions about why

Here, *self-proclaimed* acts first and foremost as a device modalizing evaluation already present in the co-text. There is a strongly negative stance taken towards Mr. DiGiacomo, which is intensified by the evaluation of the host of the show he appears on. The understanding is that offering softball questions to people who should be strongly challenged is what you could expect from a liberal, and if this is accepted, the attitudinal stance taken towards Greg Dobbs, who voluntarily calls himself a liberal, is very negative indeed.

3.3.2 Epistemic stance

In (10) an example of epistemic stance is introduced, exemplifying a concordance in which the stancetaking is focused on the position of a speaker about what the source of information is,

- (10) campaign to open Stonehenge was Arthur Pendragon, founder of the Loyal Arthurian Warband and self-proclaimed reincarnation of King Arthur. He picketed the stones with " Excalibur " in hand

Here *self-proclaimed* is used in a manner focusing on the source of information, Arthur Pendragon. He is the founder of an organization, and also claims to be the reincarnation of a mythical king. Removing the word *self-proclaimed* in the text, the force of the proposition would be very much different, in that it would support a controversial position rather than just reporting on it; it assumes a “a reportive epistemological stance ... used as a means for speak-

ers to avoid assessing the degree of their authority over information since the information derives from the speech act (and therefore the epistemological assessment) of the original (reported) speaker” (Mushin 2001:151). Thus, the commitment of the speaker towards the truth-value of a reincarnation of the legendary British leader is effectively modalized in transfer from speaker to reported speaker.

3.3.3 Intersubjective stance

Until now, stance has been considered from the subjective self-expressive perspective. Under the heading of intersubjective stance, *self-proclaimed* is instead considered in the dialogic light of what words and phrases accomplish functionally – i.e. how speakers use them to position themselves in relation to previous utterances and in anticipation of (and construction of) a perceived audience. These aspects are pointed out and analyzed in (11) and (12):

(11) all available facts before pronouncing on a weighty subject. Readers feel abused when a self-proclaimed expert or talking head announces loudly a position totally inconsistent with the facts known to

(12) be earning with you, right? I'm sorry, but he's a self-proclaimed anarchist who lives in fucking Malibu who made a few good movies in the seventies

Self-proclaimed always perform a dialogical function because it is lexically encoded in the same way as an evaluative function is semantically encoded in a word like *betray*. In the *self-* it typically refers back to the utterance of a previous speaker - of the 673 instances of *self-proclaimed*, only one concordance line indicated a speaker using *self-proclaimed* to refer to him- or herself. However, even though (11) is an example of dialogicity, the space open for dialogue is contracted at the same time - both in the reference to a previous speaker and for the constructed audience. The previous speaker is identified with a talking head, a pundit - hardly an identity conventionally associated with the marks of an expert. Simultaneously, the dialogical options of the audience are contracted as well. A strong sense of solidarity is invoked by the speaker by addressing conventional values: no one likes to be told what to think. And if the one responsible for this is not even an expert, but a self-proclaimed expert, the dialogical space is all the more contracted.

As for (12), here *self-proclaimed* yet again injects a dialogic element to the co-text. Someone attests to being an anarchist, and the co-text takes on an attitudinal stance of negative evaluation towards this identity. Despite this hostile stance, the identity is accepted, and

the dialogical content towards the previous speaker amounts to an expanding space of “I accept his claim, but I do approve of it”. In regards to the audience, a dialogical expansion is present as well. It invites a negative evaluation of the subject of the utterance, but it still does not totally block alternative views. This is evidenced in how the authorial voice proceeds to present evidence for its case, for which there would have been no need if the construed audience was not assessed as needing additional persuasion. One of the arguments mobilized is the fact that the person reported on indeed calls himself anarchist, although *self-proclaimed* here is not absolute like in (11), but a modal device in dialogical expansion.

4 Conclusion

Through the examination of user-based data, it has been shown that the meaning of a lexical item like *self-proclaimed* cannot be considered as limited to only the node-word itself, because, as stated by Sinclair,

so strong are the co-occurrence tendencies of words, word classes, meanings and attitudes that we must widen our horizons and expect the units of meaning to be much more extensive and varied than is seen in a single word (qtd. In Dam-Jensen, Zethsen 2007:1616)

Thus, analysis of recurrent patterns shows that word-meaning should to be considered a linguistic feature straddling the classic division of language into syntax, semantics and pragmatics. In taking up this Sinclarian view of meaning in the study of *self-proclaimed*, several interesting results were observed. Collocational data confirms that *self-proclaimed* exhibits a semantic preference for both certain collocates and certain semantic sets. That is, language users do not use this attributive adjective to describe just whichever random person; they use it to describe certain entities which can be grouped into certain clusters that are charged with meaning.

Also, the concept of discourse prosody extends an understanding of *self-proclaimed* through placing its meaning as function in context. This is illustrated by comparing the meaning found in corpus and non-corpus-based dictionaries with our results. It is shown that if no discourse prosody is present, the meaning of *self-proclaimed* “defaults” to the denotational meaning seen in the Merriam-Webster dictionary. However, there is more to a word, in this case three kinds of discourse prosodies. First, there is one of negation; second, there is one of accepted-negative and third; there is one of accepted positive. Intriguingly, the first two correspond to the definitions given in the Collins COBUILD dictionary. The third has not previously been systematically described, and, in contrast to the others, it has a communicative function of positive evaluation. In addition, we can also conclude that both discourse prosodies and semantic preferences exhibit context-dependent patterns; they tend to occur more often in certain registers and text-types than in others.

For semantic preferences, there is a strong tendency of collocations belonging to the semantic set of cause/ideal to appear in academic type publications. Almost as strong is the pattern of a centrality-related semantic preference in news type publications. Identities expressing a stance of enthusiasm are common in magazine type publications but rare in all other types - except fiction, where an almost perfect random distribution of semantic sets is found. Another interesting finding is the genre-straddling tendency of phraseological items like *self-proclaimed expert* to be evenly distributed across all text-types.

For discourse prosodies, the trends are not as clear or strong as for semantic preferences. Nevertheless, I have identified three interesting patterns:

- The tendency of occurrence for the denial-type prosody remains close to constant across the three most common categories of publication. However, when grouping together both prosodies expressing different kinds of negativity we notice that *self-proclaimed* tends to feature negative prosodic meanings to a high degree rather than positive ones, especially in academic and news-type publications.
- The accepted-positive prosody is the single dominant form in magazine type publications, while accept-negative prosodies are at their least common here.
- The sizeable occurrences of unclassifiable concordances in journalistic news and academic publications leads us to conclude that, here, *self-proclaimed* often either is used as a “neutral” reportive adjective, with no prosody at all, or occurs in complex co-texts and contexts where the evaluative meaning is highly dependent on contextual pragmatic knowledge.

These results raise further question; about who is evaluated, in what manner and in what context. A comparative corpus study of near-synonymical lexis could shed further light on these considerations. This would entail constructing similar lexical profiles for words like self-styled, self-described and self-professed, which then could be used in a comparative corpus study. A first rationale would be that this could help clarify how language users modalize their evaluations. A second rationale revolves around the need to contribute to a theory of social action, underlining the relevance of Stubbs’ speculation that “if descriptions of evaluative meanings and cultural stereotypes can be made reasonably precise, it seems plausible that

this will tell us about the important meanings expressed in a discourse community” (Stubbs 2002:216)

Lastly, the concept of stance was introduced. With stance, corpus studies of ELU is linked to an approach to evaluation which shares some aspects, but also adds parameters making it more useful for analyzing evaluation in larger discourse units. Another reason for analysing stancetaking is the limits of the concordance; you can only get as far in classifying more pragmatically encoded linguistic features from discourse fragments. *Self-proclaimed* was exemplified as a carrier of evaluative meaning working together with other linguistic resources in different kinds of stancetaking. This allowed for an initial sketching out of the role of *self-proclaimed* as a linguistic resource understood inside a discourse semantic framework. This orients us to addressing “meaning in context and towards rhetorical effects, rather than towards grammatical form” with alternative opinions, points of view and value judgments providing the Bakhtinian backdrop for a bringing together of “lexically and grammatically diverse selection of locutions” in discourse semantic categories (Martin and White 2005:94). As argued by Hunston, the problematic character of stance makes quantitative figures useful only as points of departure for stance investigations (2007:28). However, corpus studies have revealed stance markers previously unknown to us. It also shows them to often be phraseologically realized. If this is accepted, then *self-proclaimed expert* is a good example of a conventional phrase which also might act as a marker of stance.

In conjunction to this question of interfacing corpus studies of meaning and intersubjective stance, it might be fruitful to consider White’s argument for the need to look beyond “commonalities in lexico-grammatical structuring or affordance and [instead] consider commonalities in rhetorical effect” (2005:208). For example, much like how co-text and context interact in determining the discourse prosody of a word (see the example of Stubbs’ *cause* in Stewart 2010:22-23), there is a similar pattern in a verb such as *to claim*, for which “the rhetorical potential ... may vary systematically under the influence of different co-textual conditions, and across registers, genres and discourse domains” (Martin and White 2005:103). In other words, in both the linguistic features of discourse prosody and stance function, there is a strong influence of context and co-text on the meaning of the lexico-grammatical units. This implicates the need for new approaches to studying them. These considerations are echoed in the speculation of Hunston and Thompson that such an endeavor would “show many grammatical features in a relatively unfamiliar light, and bring out connections between apparently unrelated phenomena” (2003:22). Such intriguing claims, together with the results presented

in this study, encouragingly suggest the productiveness of further studies linking discourse prosodies with the developing field of discourse semantics.

5 References

- Baker, P, Gabrielatos C et. al. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & Society*. 19:3. 273-306.
- Biber, Douglas. (1996). Corpus-Based Investigations of Language Use. *Annual Review of Applied Linguistics*. 16.115-36
- Conrad, R and Biber, D. (2003) Adverbial Marking of Stance in Speech and Writing. In: Hunston, S. & Thompson, J. (Eds.) *Evaluation in Text. Authorial Stance and the Construction of Discourse*. Oxford University Press, Oxford.
- Dam-Jensen, H, and Zethsen, K.K. (2007). Pragmatic patterns and the lexical system - A reassessment of evaluation in language. *Journal of Pragmatics*. 39. 1608–1623
- Du Bois, John W. (2007) The stance triangle. In: Englebretson, R. (Ed.) *Stancetaking in discourse: Subjectivity, evaluation, interaction*. John Benjamins Publishing Co., Amsterdam.
- Evert, S. (2005). The Statistics of Word Co-occurrences: Word Pairs and Collocations. Ph.D. thesis, University of Stuttgart
- Fairclough, Norman. (2003). *Analysing Discourse: Textual Analysis for Social Research*. London and New York: Routledge
- Gill, Philip. (2009) Why prosodies aren't always present. Insights Into the Idiom Principle. Retrieved 6 June 2011 <ucrel.lancs.ac.uk/publications/cl2009/317_FullPaper.rtf>
- Hunston, S & Thompson, J. (2003) Evaluation: An Introduction. In: Hunston, S. & Thompson, J. (Eds.) *Evaluation in Text. Authorial Stance and the Construction of Discourse*. Oxford University Press, Oxford.
- Hunston, Susan. (2007) Using a Corpus to Investigate Stance Quantitatively and Qualitatively. In: Englebretson, R (Ed.) *Stancetaking in Discourse*. John Benjamins Publishing Co. Amsterdam.
- Hyland, Ken. (2005). Stance and Engagement: a Model of Interaction in Academic Discourse. *Discourse Studies*. 7:2. 173-192.x

- Martin, J.R (2003). APPRAISAL Systems in English. In: Hunston, S. & Thompson, J. (Eds.) *Evaluation in Text. Authorial Stance and the Construction of Discourse.* Oxford University Press, Oxford.
- Mushin, Ilana (2001). *Evidentiality and Epistemological Stance. Narrative Retelling.* John Benjamins Publishing Company: Philadelphia, PA, USA
- Precht, Kristen (2003). Stance Moods in Spoken English: Evidentiality and Affect in British and American Conversation. *Text, 23: 2.* 239-257.
- Stewart, Dominic (2010). *Semantic Prosody. A Critical Evaluation.* New York and Oxon: Routledge.
- Stubbs, Michael (2002) *Words and Phrases. Corpus Studies of Lexical Semantics.* Oxford: Blackwell.
- Stubbs, Michael (2009). The search for Units of Meaning: Sinclair on Empirical Semantics. *Applied Linguistics*, vol. 32: 2. 115-137.
- White, P.R.R. (2003). Beyond modality and hedging: A Dialogic View of the Language of Intersubjective Stance. *Text, 23:2.* 259-284.
- White, P.R.R. and J.R. Martin (2005). *The Language of Evaluation: Appraisal in English.* Hampshire and New York: Palgrave MacMillan.
- Widdowson, H. G. (2000). On the Limitations of Linguistics Applied. *Applied Linguistics 21/1.* 3-25.

6 Appendix

See Appendix1.docx