

CORPUS **Corpus**
10 | 2011
Varia

Rising through the registers – A corpus-based account of the stylistic constraints on Light Verb Constructions

Arian Shahrokny-Prehn et Silke Höche



Édition électronique

URL : <http://journals.openedition.org/corpus/2110>
ISSN : 1765-3126

Éditeur

Bases ; corpus et langage - UMR 6039

Édition imprimée

Date de publication : 1 novembre 2011
Pagination : 239-257
ISSN : 1638-9808

Référence électronique

Arian Shahrokny-Prehn et Silke Höche, « Rising through the registers – A corpus-based account of the stylistic constraints on Light Verb Constructions », *Corpus* [En ligne], 10 | 2011, mis en ligne le 15 juin 2012, consulté le 01 mai 2019. URL : <http://journals.openedition.org/corpus/2110>

Rising through the registers – A corpus-based account of the stylistic constraints on Light Verb Constructions

Arian SHAHROKNY-PREHN, Silke HÖCHE
Leibniz Universität Hannover

1. Introduction

In this paper we aim at a usage-based description of stylistic facets of so-called Light Verb Constructions (LVCs), complex verbal phrases which seem to be all-pervasive in English:

- (1) Okay, well let's *have a listen* to your chest. (BNC; GYC 28)
- (2) As long as you don't mind me *having a little smoke*? (BNC; G0A 1164)
- (3) He suggested that they might *take a short stroll*, just a short one. (BNC; A7A 2924)

As will become clear from section 2, which provides a detailed account of form and meaning of these constructions, they offer a multiplicity of research questions, comprising syntactic, semantic, pragmatic, stylistic and sociolinguistic levels of descriptions. In this paper, we are going to focus on usage characteristics relating to the stylistic value of LVCs, which are –to some extent– stigmatized as colloquial forms. This evaluation is based on intuition-based descriptions by (non-) native speakers, and we aim to empirically dismiss such an insufficient stylistic categorization of the constructions investigated here.

The outline of the paper is as follows: Section 2 makes the reader familiar with the state of the art concerning approaches to LVCs. Section 3 provides an overview on opinions, claims and studies dealing with the central topic of this paper, i.e. register-specific characteristics of LVCs. Section 4 presents the methodology that we used for conducting our research, while in section 5 we explore whether LVCs are tied to specific

genres and to which extent these genres can be analysed in terms of colloquialness. The different subsections of chapter 6 then elaborate on the semantic classes preferred by specific genres and also raise the intriguing question how representative corpora really are. Finally, section 7 summarizes our findings and presents future perspectives.

2. A short history of LVCs

Being so ubiquitous, Light Verb Constructions have, of course, received considerable attention from linguists with highly diverse theoretical backgrounds (see, among many others, Labuhn 2001; Gradečak-Erdeljić 2004). It seems, however, that scholars shun extensive, coherent and empirically based investigations of these structures because of their seeming unsystematicity in the choice of available light verbs and the complements allowed by each of these. As Stein remarks (1991: 2), “the idiosyncrasies of some expressions seem to have prevented linguists from seeing the underlying pattern”. Little has changed since Stein’s implicit call for studies, one exception being Labuhn (2001).

While reference grammars of English written in the first decades of the 20th century acknowledge the existence of LVCs (Curme 1931; Jespersen 1954 [1942]), the first more systematic accounts had to wait until the 1960s (Nickel 1968). A very interesting and well-structured approach was provided by Anna Wierzbicka (1982), who, asking “why can you have a drink when you can’t have an eat?”, captures such forms as constructions carrying their own formal and semantic characteristics. The focus being put on *have a V*, she works out specifications in great detail for a collection of subforms of the pattern which can be captured by the following constructional schema:

Light Verb + *a* + (modifier) + verb stem

A few comments on the elements and terminology made use of are due here: Firstly, as regards the notion “light verb”, this term was originally used by Jespersen 1954 [1949] and is still applied, at times off-handedly, at times purposefully. The label is commonly done justice with scholars presenting the members of this verb class (i.e. *do*, *give*, *have*, *take*, *make*) as

semantically empty, delexicalized or desemanticized (see Allan 1998; Gradečak-Erdeljić 2004), being nothing more than a linking element. It is, however, disputable whether they are as light as they seem.

Secondly, the use of the term “verb stem” in the description of LVCs is an innovation by Wierzbicka (1982). Although many authors describe the unit following the LV as a noun, Wierzbicka’s approach (see also Gradečak-Erdeljić 2004) seems highly plausible, if we take into consideration pairs such as *have a think* and *have a thought*, where in the latter case the predicate is complemented by a deverbal noun, while in the former case we find a verb stem. In fact, constructions involving light verbs are highly diverse, and those patterns incorporating verb stems as the complement of the light verb constitute only one particular group in a continuum ranging from verb stem (*have a quick drink*) via lexicalized verbal nouns (*have a cold drink*) to morphologically derived nominals (*have an argument*).¹

The focus of our study is on the first variant, which, due to the verbal character of the element following the light verb, preserves a highly processual character of the action or event verbalized in such a constructional frame. As can be taken from the above examples the same form, in this case *drink*, can instantiate two different kinds of complements for the light verb, evoking different ways of interpretation. In many cases, the modifying element is indicative of the particular type of reading of *drink*. While attributes such as *quick*, *friendly* etc. clearly imply a processual reading, modifiers like *hot*, *cold*, *long* or *cocoa* identify the following item as denoting a concrete object (i.e. the lexicalized noun variant). More difficult are instances such as *have a drink* or *have a look*, where the modifying slot remains empty. In the majority of cases, both interpretations are possible, and only contextual material can disambiguate the intended reading. Thus, manual selection and analysis of all potential target structures retrieved from a corpus

¹ The differences here are subtle. While the modifier *quick* clearly implies an event reading of *drink*, *cold* definitely modifies a concrete entity *drink*. In other cases these differentiations are not as straightforward, though.

is absolutely necessary in order to prevent a skewed set of data (see section 4 for methodological issues).

Research conducted on LVCs to date has discussed several, often interrelated issues: One major point of interest is the (non-)equivalence of LVCs and their corresponding simple verb variants. Thus, the question often pursued is: if existing at all, what is the specific meaning of LVCs as compared to their simpler counterparts, e.g. ‘We *had a quick look* at the map.’ vs. ‘We *quickly looked at* the map.’? What most researchers agree on is the temporally bounded reading that the construction imposes on the event to be verbalized. Moreover, Wierzbicka (1982), for example, notes a certain light-heartedness concerning the agent’s involvement in the action presented by *have a V*, and also a certain degree of pleasure or relief. *Take a V*, on the other hand, is ascribed an inchoative function, emphasizing the initiation of the event.

A second point of interest is in how far the allegedly semantically light verbs add meaningful aspects to the construction. Cattell (1984: 2) notes that these verbs seem to mean “little more than that a verbal action occurred”, and Huddleston and Pullum (2002: 290) add that the “main semantic content is located [...] in the noun functioning as the head of the direct object”. Linguists working in the Cognitive Linguistic paradigm have abandoned the tradition of considering the respective verbs as being meaningless, and instead seek the motivation of constructional meaning in the full forms of these verbs (e.g. Newman 1996).

Thirdly, regarding their functional contribution to the grammatical system of English, some authors mention the syntactic “flexibility” or “versatility” of LVCs. Here, the constructions are frequently presented as a means to open up slots for elaborate modification, as in ‘Lady Roscarrock *gave a sigh of pleasurable anticipation.*’ (BNC, EWH 1524). They are thus considered as a form which is semantically equivalent to the simple verb construction, which, however, permits the insertion of complex modifying material. Moreover, the constructions are said to facilitate coordination of eventive elements, as in ‘I

rushed off to *have a wash* and *a shave* [...].’ (BNC; EE5 396) (Renský 1966; Nickel 1968²).

A fourth issue, as formulated by Wierzbicka (1982: 757), is the existence of “considerable dialectal differences” in the usage of *have a V* and *take a V* constructions between American English and British English. A preference of *have a V* over *take a V* in BE is commonly reported, while the reverse situation is said to hold for AE. As we were able to show in an earlier study, however, both forms are well-established in both dialects and exhibit their own specific regularities (Höche, Shahrokny-Prehn 2009).

Finally, a notion which is acknowledged by several authors, but usually presented in a few lines without further ado or empirical evidence, is the assumption of substantial stylistic constraints on the *have a V* constructions insofar as they are “highly colloquial” (Wierzbicka 1982: 757). The focus of our research presented in this paper is put on this last claim, which we have empirically investigated on the basis of a large set of data retrieved from the Corpus of Contemporary American English (COCA). Before we get to the discussion of numbers and results of this investigation, we briefly summarize opinions and findings on the stylistic status of LVCs, both from a synchronic and diachronic point of view.

3. A history of colloquialness?

Concerning stylistic nuances, many authors emphasize the rather colloquial flavour carried by LVCs. The following quotes are to illustrate this common tendency:

All these cases [LVCs] indicate a reluctance in **colloquial speech** to predicate by means of a full verb, since this method is felt as too formal, too scientific, precise. In **colloquial language** there is here as elsewhere a tendency to more concrete forms of expression. A noun seems nearer to

² See, however, Shahrokny-Prehn, Höche 2010, who show on the basis of LVCs as represented in the BNC that only a very small set of utterances including instances of the construction actually serve the coordination of events.

popular feeling than the more abstract verb.
(Curme 1931: 22, our emphasis)

The construction type [i.e. LVC] provides greater weight than the corresponding SV [simple verb] type, especially if there are no optional adverbials, and is often preferred to the SV construction **in informal language**. (Quirk *et al.* 1985: 751, our emphasis)

One counter-statement is found in Stein (1991: 26), who, suggesting that *have a V* expresses “personal courtesy, personal attention and care for another person” plus some higher degree of politeness, critically questions the validity of the construction’s common assignment to colloquial registers. Serious doubts on the restriction of LVCs to informal contexts are also cast by Claridge (2000). While she does not question the more frequent use of multi-word verbs in spoken language, she emphasizes her conviction that “this type has thoroughly permeated all levels of the language nowadays. Academic writing in the humanities, for instance, can bear witness to this [...]” (*ibid.*: 104).

An interesting point is made by Dixon (2005: 461). Although he agrees on the usage of LVCs as colloquial forms from a synchronic point of view, he does, however, point at records of the use of the constructions in Middle English, which, of course, have as their source highly elaborate literature and therefore cannot be deemed colloquial:

These constructions tend to carry an overtone of friendliness and intimacy, and are found far more frequently in colloquial than in formal styles of English. Some examples are found in the older literature, e.g. *give a cry* from 1300, *have a run* from 1450, but these are comparatively rare. Note, though, that **very little of premodern literature reflected colloquial usage**. (*ibid.*, our emphasis)

Indeed, the quotation base of the Oxford English Dictionary online (2nd ed. and later additions) (*OED*) contains only few examples of the constructional frame discussed here from the Middle English period.

Rising through the registers – A corpus-based account of the stylistic constraints on Light Verb Constructions

As several authors note, the pattern is much more typical of Modern English with its tendency towards analytic structures (Jespersen 1949; Nickel 1968; Brinton, Akimoto [eds] 1999). Yet, complex predicates involving (equivalents of) *have*, *take*, *give* and *make* were already in use in Old English, and there is a growing body of literature on corpus-based studies of this linguistic phenomenon (Brinton, Akimoto [eds] 1999; Claridge 2000 *inter alia*).

Without going into detail regarding individual studies, we can state that a hard and fast statement on the stylistic constraints underlying the use of LVCs seems impossible. While the issue appears to concern the majority of authors exploring these structures, their findings are rather heterogeneous. Concerning the diachronic analyses, the different results of the studies summarized here are obviously owed to differences in the formal definitions the authors apply to the structures that they are interested in, the make-up of the corpora that they consulted for their investigations, and the stages of English considered. Similarly diverging is the picture emerging from synchronic descriptions, which, although well-balanced corpora have been available for decades now, is still predominantly depicted on largely intuitive grounds.

The corpus study reported on in the following section sets out to remedy some of the shortcomings of synchronic depictions of LVCs with special emphasis on the stylistic categories that they preferably occur in. It is based on a large set of data retrieved from a well-balanced corpus and incorporates empirical methods in order to validate our claims.

4. Methodology

During earlier research we compiled a list of possible LVCs by simply starting with a corpus search for *{take} a* and *{have} a* (for the British National Corpus (BNC) curly brackets –in the case of the COCA square ones– indicate a lemmatised search). Following Wierzbicka's formal criteria of the LVC we thinned out the results, ending up with a list of around 80 different verb stems that could be part of an LVC. Excluded were words that clearly did not fit the verb stem / zero-derivation parameter, e.g.

have an argument where the post-LV component is a morphologically derived noun. Although this level of selection is quite straightforward, several instances had to be excluded later on when it became apparent that a potential LVC was identical in form with a simple verb + object string, e.g. *had a look on his face* (V+O) as opposed to *had a look at his face* (LVC).³ These instances had to be manually selected and excluded. As in the example given, prepositions sometimes provided helpful clues. In cases where even a close analysis of the respective contexts was not enough to clear up ambiguity (e.g. *have a view*), these, too, were excluded as LVCs.

Furthermore, we also took into account that LVCs can occur with internal modifiers, e.g. *{take} a close look*, and accordingly placed search inquiries along the lines of *{have} a * look* in order to find all instances of a potential LVC string containing one (or multiple) modifiers (the asterisk being the placeholder for wildcards in the COCA). These findings were also put under scrutiny in order to exclude any non-LVC strings. In several instances, the internal modifiers made the process of exclusion easier since the verb stem of the LVC prefers a certain set of modifiers while it is very seldom used with certain others that can be attached to the formally identical deverbal noun. For example, *close* as a modifier is much more likely to occur with the verb stem *look* in the LVC than with the deverbal noun; on the other hand, *surprised* will co-occur with *look* in the ‘facial expression’ setting where we are not dealing with an LVC.

Finally, we categorised the verb stems based on semantic classes following and partly adapting Levin’s (1993) terminology. Such a step is useful for the detection of semantic idiosyncrasies of the construction and the identification of possible subconstruction. Thus, all of the analyses that we conduct for our data, including the present one, are sensitive to semantic categories. As will become apparent in this study, different genres of texts display different affinities for particular semantic verb groups.

³ The difference here is not only in the preposition. While the former refers to an expression on someone’s face, the latter encodes looking at someone’s face.

Rising through the registers – A corpus-based account of the stylistic constraints on Light Verb Constructions

Departing from our previous use of the BNC as our research corpus, for our present venture we chose the Corpus of Contemporary American English (COCA). At the time of writing the COCA contains 400 million words divided equally among five different genres, namely SPOKEN (SPOK), FICTION (FIC), MAGAZINE (MAG), NEWS, and ACADEMIC (ACAD), with 20 million words added to the corpus each year, i.e. 4 million for each genre. The equal distribution of text enables us to undertake a detailed analysis of the occurrence of the LVC in different genres and test the aforementioned statements that it is a construction preferred by if not restricted to the colloquial register.

5. A question of genre?

In order to provide a large enough basis for our analysis, we chose the *take a V* construction for our corpus investigation since it is observably more frequent in AmE (19,522 tokens) than the *have a V* construction (4,034 tokens). Our first step, then, was to search for the *take a V* string in the corpus, and after meticulously sifting through the data and counting out the instances for the different genres, the result appeared almost effortlessly simply before our eyes.

Table 1. Distribution of take a V across the different genres of the COCA

	SPOK	FIC	MAG	NEWS	ACAD	TOTAL
TAKE a V	8143	6341	2657	1769	602	19522
Percentage	42	32	14	9	3	100

The numbers clearly show that in the case of *take* as the LV the construction is favoured by the SPOKEN genre (8,143 instances) with FICTION trailing closely behind (6,341 instances), the remaining three genres constituting a good quarter of all instances (MAG 2,657, NEWS 1,769, ACAD 602).

One of the difficulties in coming to a reliable conclusion about the stylistic properties of the construction is that style and genre are not interchangeable. Yet, the division that we are dealing with in the COCA is based on different genres.

Is it possible –if only for efficiency’s sake– to interpret the occurrences in one of the genres as indicative of colloquial style? The short answer is “no”; the long answer will be attempted in the following.

The general question we have to ask –and answer– first is this: how ‘colloquial’ exactly are the different genres? First of all, for a construction to even enter a corpus like the COCA it has to be more than purely colloquial since it would simply not be represented if it were. The make-up of corpora in general and the COCA in particular severely limits the inclusion of elements that are purely colloquial. For spoken language to even be included in a corpus it has to, first of all, be recorded by someone, either for the explicit purpose of corpus-building or for some other reason. Since it is still impossible to incorporate large amounts of spoken text into a corpus automatically, spoken segments are relatively small in comparison. Furthermore, due to legal reasons, persons to be recorded have to be aware of the process and agree to it. For scientific reasons, recordings are usually made in a very controlled environment, i.e. a specific group of people is recorded in certain pre-specified situations. All these factors clearly limit –and sometimes exclude– the occurrence of colloquial language in a corpus. However, if we can say that it is unlikely for colloquial language to be included in a corpus like the COCA, and at the same time observe a relevant number of occurrences of LVCs in this corpus, then we can come to the cautious conclusion that the use of LVCs is certainly not restricted to colloquial language.

In order to be able to include spoken language and at the same time allow for a continuing expansion of the corpus, the compilers of the COCA have to rely on pre-transcribed text, meaning that the SPOKEN element consists almost entirely of transcripts of TV programmes. It does not contain recordings of ‘spontaneous’ speech along the lines of other corpora like e.g. the BNC. In some cases this peculiarity is of only limited consequence. Especially in shows like *Jerry Springer* et al., speech production is often spontaneous and unscripted and can therefore be included as truly spoken language without serious problems. However, in other instances like news programmes

we are entering shallow waters. What in general can be referred to as the Anchorman Syndrome (and which will be elaborated in a special case study below) refers to instances where a particular speaker has a peculiar catchphrase, e.g. ‘*I’m Keith Olberman. Good night and good luck.*’ Additionally, due to the specific set-up of the news programme certain formulations / formulaic expressions will turn up more frequently than in other forms of spoken communication (‘*Good evening and welcome to Dateline / 20/20 / Crossfire.*’; ‘*Let’s take a look...*’). None of these phenomena corrupts or devalues the findings that can be made. Still, we have to carefully consider which instances are indeed markers of a more general trend and which are too special to be used in support of these trends.

Furthermore, although news programmes, daily TV shows and radio broadcasts may not necessarily be scripted, they frequently use a style that is certainly not identical with the colloquial, adhering to rules and regulations of their trade. In short, to equate *spoken* with *colloquial* is a notion that in truth cannot be upheld.

Additionally, can the occurrence of LVCs in the Fiction genre be explained as an imitation of real spoken language by an author and therefore be used as a justification for the overall colloquialness of the construction? Even if the language of a novel appears colloquial it has to be regarded as highly artificial since the author will in all likelihood have spent a considerable amount of time on the construction of a certain dialogue, down to the level of what individual characters are supposed to represent and how their language can be used to support this impression. Additionally, an author’s impression of what counts as colloquial does not necessarily coincide with real, i.e. naturally occurring colloquial speech. However natural a verbal exchange in a novel might appear it can actually not be regarded as spontaneous speech.

Summing up, the occurrences of LVCs in the COCA cannot be ascribed to colloquial language. Although this result can only show that these constructions are not restricted to colloquial language, their usage in highly formal contexts such as news programmes supports the assumption that they are

indeed not ‘highly colloquial’ as is often claimed. In what follows, a more detailed analysis will be undertaken in order to come to a more adequate conclusion of the distribution of Light Verb Construction across the different genres.

6. The eye of the beholder

As has been shown above, the condensation of empirical data into one neat chart, however charming that might appear, covers up rather than unveils the intricacies of the item under scrutiny. Accordingly, in what follows a more fine-grained analysis will be attempted. For starters let us have a look at a pre-condensed version of the above chart.

Table 2. Distribution of take a V broken down by genre and semantic class

Semantic class	SPOK	FIC	MAG	NEWS	ACAD	TOTAL
visual perception	7053	1723	1355	1080	439	11658
consumption	179	2970	348	212	40	3750
motion	288	708	610	299	71	1977
bodily hygiene	239	655	239	133	38	1303
sensory perception	266	25	9	3	2	305
pause	50	74	52	30	9	216
egestion	1	140	15	0	0	156
try verbs	59	44	22	9	0	134
bodily contact	1	0	0	0	0	1
amusement	0	0	0	0	1	1
joint speech activity	0	0	0	0	0	0
others	7	2	7	3	2	21
TAKE_a_V (AmE)	8143	6341	2657	1769	602	19522

It can be seen quite clearly that VERBS OF VISUAL PERCEPTION encompassing the verb stems *look*, *view*, *glimpse*, *glance*, *peep*, and *peek* constitute more than half of all instances of the *take a V* construction. Even more interesting, though, is the observation that it is responsible for over eighty percent (86.6%) of all hits within the SPOKEN genre. In a way, the above chart does not so much demonstrate the presence of light

verbs in the different genres in the COCA but more the omnipresence of the *take a look* construction in this particular genre. In a first case study, the VISUAL PERCEPTION-group will be put under closer observation while the rest of the LVCs will be tackled separately and used as a point of reference.

While roughly two thirds of all instances of the non-VISUAL PERCEPTION LVCs fall into the FICTION category and only 14% into the SPOKEN, in the case of the VISUAL PERCEPTION-group the situation is reversed. In order to better understand this ‘deviant’ behaviour of the latter group of LVCs, a case study focussing on the *take a look* construction will be undertaken below. We chose this particular construction since it is by far the most frequent with 11,333 of 11,658 tokens, thereby constituting 97% of the whole VISUAL PERCEPTION-group.

6.1 Case study: take a look

At least part of the reason for the high percentage of occurrences of *take a look* in the SPOKEN register can be found in the specific make-up of this category. As mentioned above, the SPOKEN part of the COCA contains TV (or radio) programmes from the different networks, featuring news or political programmes like *PBS Newshour*, *CNN Crossfire*; general interest shows like *NBC Today*, *ABC’s Good Morning America*; and talk shows like *Oprah* or *The Jerry Springer Show*.

One feature that is of special interest for our present research takes full swing in programmes like these, namely the Anchorman Syndrome. The presence of a host / anchor / presenter interacting with an unseen –and often not even physically present– audience promotes the usage of the almost formulaic expression *take a look*. In some instances, specific visual material is presented following this statement: ‘*Take a look at this surveillance video / the weather chart / their on-camera confession / the polls.*’, or more generally, the audience is invited to view a video clip. Correspondents or segments of the running programme are introduced who will *take a look* at a specific event, certain people or a particular topic. In short, there is a lot to be looked at in programmes like these.

In order to get an idea what can be looked at when we are not dealing with TV shows let us now consider another

genre of the COCA in which this construction turns up. In FICTION, the broad notion of presenting visual material is retained although it now refers to much more concrete entities like objects, documents or body parts.

In all likelihood, the difference is not so much what is looked at but rather by whom. In the novel a specific setup or constellation of people is necessary in order for this construction to work. Usually more than one person has to be present and a shift / focus of attention on a specific entity has to be initiated. In the TV show these prerequisites are a given factor since there are always at least two parties present, namely the host and audience (even if the latter is in front of the TV set), and the former's function as presenter already creates a situation that is very susceptible to this specific construction. The host frequently invites viewers to consider certain material, be it a video-clip that follows, statistics, polls, or something else. Therefore *take a look* is commonly used in the setting of a TV show, especially news programmes.

6.2 *The curious incident of the non-existent construction*

So far, we have been talking about a genre-specific tendency of usage. What follows now will, on the one hand, present a further case study that will elaborate on the above passage, and on the other hand will also make it apparent that although we should "trust the text", as Sinclair so famously put it, we cannot blindly put our faith into the representativeness of even the largest of corpora.

In her paper, Wierzbicka (1982: 795) claims that while *take a look* is quite possible as an LVC the phrase **take a listen* (whether the asterisk is well-founded or not will be shown shortly) will not appear since it is incompatible with the functional content of *take*, namely an initial movement of the body-part involved in the action.⁴ Tempted by this bold statement, we undertook a corpus search in the COCA for the so-called non-existent phrase and were indeed rewarded. We found that *take a*

⁴ Brugman (2001) comes to a different conclusion about the semantic content of *take*, employing force-dynamic notion in her analysis. We will, however, not go into detail on this point.

Rising through the registers – A corpus-based account of the stylistic constraints on Light Verb Constructions

listen, a prominent member of verbs of sensory perception, occurs 252 times (internal modifier included), which is definitely more than can be expected for an allegedly ungrammatical phrase even when taking into account the odd error in production, e.g. a speaker confusing *look* and *listen*. In itself, this result is quite elating, empirically proving the relativity, if not to say fallibility, of intuitively made statements.⁵ However, even corpus data sometimes have to be taken with a grain of salt.

The phrase almost exclusively appears in the Spoken part of the COCA (245 of 252 hits) which is quite odd when we take into account that even *take a look* as a very biased construction does occur in the other genres in at least a third of all instances. When we look at the distribution of the construction across time, another peculiarity becomes apparent.

The construction indeed appears to be on the move increasing in frequency from 8 hits in the first five years (1990-1994) to more than 20 times as much in present day American English (2005-2009). An even more fine-grained analysis reveals that there is a steep increase in hits between 2004 (9) and 2005 (40) that continues through 2006 (47) and becomes erratic afterwards. How can such behaviour of a very specific construction be explained?

As can indeed be expected from the discussion of *take a look* above, *take a listen*, too, appears almost exclusively in news shows and similar programmes. Here, it is used as an introductory phrase aimed at the audience that, interestingly enough, precedes video clips more often than audio clips. In this respect the construction fulfils the same function as *take a look* since it is not restricted to auditory material. Why exactly this construction is preferred in some cases cannot be conclusively answered. However, a closer look at the data reveals intriguing details.

From the information the COCA gives about individual texts we can discern that in almost half of all instances (113 of 252) the source is a CNN broadcast. Even more specifically,

⁵ We cannot, however, rule out an on-going language change that only became observable in the intervening years between Wierzbicka's article (1982) and the first occurrence in the COCA (1990).

half of the CNN hits come from a particular news programme, namely *Nancy Grace*. All of the *Nancy Grace* hits (56) fall into the period from 2005 to 2008 (2005 being the year in which the programme was released). What this means is that a third of all hits within a five-year period (2005-2009) can be attributed to one particular source. This source provides more instances than were accounted for in the preceding period in total. The overall effect that this has, and which can be seen quite clearly in the diagram, will be referred to as the ‘Nancy Grace Spike’.

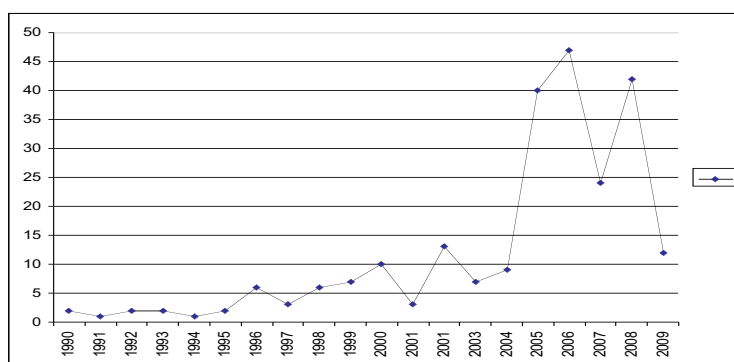


Figure 1. Occurrences of *take a listen* in the COCA

Apart from the overall increasing trend –if one is so bold as to call it thus– of the *take a listen* construction, the exponential increase in the use of the construction from 2000 onwards indeed appears to be triggered by one individual, namely Nancy Grace. We might speculate that the presence of this highly peculiar phrase in her programme worked to speed up an already existing development, yet we will not be able to provide definite proof for this theory.

However, the conclusion that can be drawn from this short episode is that even with large corpora like the COCA (400 million words and increasing) we cannot blindly trust the numbers. In instances where we are searching for particular items that are not highly frequent, a single individual can actually make a difference.

7. Conclusion

Our investigation into genre- or style-specific preferences of LVCs has raised important issues concerning corpus organization, text categories in the corpus, analytical tools, etc. We first of all have to admit that there is no one-to-one relationship between genre and style. The postulation of categories itself makes assumptions concerning which texts can be grouped together, and yet that does not necessarily mean that these texts share sufficient, let alone all vital attributes like, e.g., level of formality. The labelling of categories poses another problem or at least invites discussion. How representative are news shows and TV programmes of spoken language in general? A genre like FICTION is by no means homogenous and encompasses a variety of texts that can range from highly informal to very formal. In a sense, genre cannot tell us anything about formality. Ultimately, it would need a line-by-line analysis in order to determine whether a particular setting is formal or informal, taking into account text domain and demographic factors such as age, sex, social class etc. of both speaker and audience.

However, some claims concerning the level of formality of LVCs can be made on the basis of their distribution across different genres. LVCs occur in all the different genres of the COCA, especially in SPOKEN and FICTION. LVCs of the Visual Perception type are particularly frequent in the SPOKEN section of the COCA, which consists of news programmes and TV shows. For at least some of the TV shows and for the news shows in general it can be said that they employ a level of formality that is ‘more formal than colloquial’. Here, specific LVCs such as *take a look* are used within a very structured setting and, indeed, appear to have attained the status of fixed phrases. The particular functional content of the LVC in this case, i.e. the speaker’s invitation to participate in the reception of information, points towards an area that future research will elaborate on: the connection between LVCs and politeness, or, as pointed out by Stein, the expression of care and personal attention (1991, see above). Anticipating some of the findings of our work in progress, we shall mention here one particularly striking observation: Both *have a V* and *take a V* have a

particular affinity for the adhortative *let's* construction, also known as “first-person imperative” (Quirk *et al.* 1985), which conveys encouragement and expresses suggestions for shared activities or actions. The highly frequent combination of the *let's V* and *have / take a V* can thus be taken as an indicator of the use of these LVCs for creating and maintaining interpersonal bonds and processes.

References

- Allan Q. (1998). “Delexical verbs and degrees of desemantization”, *Word* 49 : 1-17.
- The British National Corpus*, version 3 (BNC XML Edition). (2007). Distributed by Oxford University Computing Services on behalf of the BNC Consortium. URL: <http://www.natcorp.ox.ac.uk/>
- Brinton L. & Akimoto M. (eds) (1999). *Collocational and Idiomatic Aspects of Composite Predicates in the History of English*. Amsterdam / Philadelphia : John Benjamins.
- Brugman C. (2001). “Light Verbs and Polysemy”, *Language Sciences* 23 : 551-78.
- Cattell R. (1984). *Composite Predicates in English*. North Ride, New South Wales : Academic Press Australia.
- Claridge C. (2000). *Multi-Word Verbs in Early Modern English*. Amsterdam : Rodopi.
- Curme G. (1983 [1931]). *A grammar of the English Language*. Vol. 2 : *Syntax*, reprint. Essex, Conn. : Verbatim.
- Davies M. (2008-). *The Corpus of Contemporary American English* (COCA) : 410+ million words, 1990-present. Available online at <http://www.americancorpus.org>.
- Dixon R.M.W. (2005). *A Semantic Approach to English Grammar*, Rev. and enl. 2. ed. Oxford : Oxford University Press.
- Gradečak-Erdeljić T. (2004). “Periphrasis in the coding of events”, in B. Lewandowska-Tomaszczyk & A. Kwiatkowska (eds) *Imagery in Language*. Festschrift in Honour of Professor Ronald W. Langacker. Frankfurt (Main) : Peter Lang, 431-443.

Rising through the registers – A corpus-based account of the stylistic constraints on Light Verb Constructions

- Höche S. & Shahrokny-Prehn A. (2009). "Have a go at *have a V*- Iconicity at the Form-Meaning interface", Conference paper presented at *Interfaces in Language 2*. Canterbury : University of Kent, May 2009.
- Huddleston R. & Pullum G.K. (2002). *The Cambridge Grammar of the English Language*. Cambridge : Cambridge University Press.
- Jespersen O. (1954 [1942]). *A Modern English Grammar*. Part VI. *Morphology*. London : Allen and Unwin.
- Labuhn U. (2001). *Von Give a Laugh bis Have a Cry. Zu Aspektualität und Transitivität der V+N-Konstruktion im Englischen*. Frankfurt (Main) : Peter Lang.
- Levin B. (1993). *English Verb Classes and Alternations*. London / Chicago : University of Chicago Press.
- Newman John. (1996). *Give. A Cognitive Linguistic Study*. Berlin / New York : Mouton de Gruyter.
- Nickel G. (1968). "Complex Verbal Structures in English", *IRAL* 6 : 1-21.
- Oxford English Dictionary*. Online version. www.oed.com.
- Quirk R. *et al.* (1985). *A Comprehensive Grammar of the English Language*. London : Longman.
- Renský M. (1966). "English Verbo-Nominal Phrases", *Travaux linguistique de Prague* 1 : 289-299.
- Shahrokny-Prehn A. & Höche S. (2010). "Under construction. Rethinking the Functionality of Light Verb Constructions", *Analysing Data > Describing Variation. Papers from the 28th International Conference of the Spanish Society for Applied Linguistic*. Vigo : Universidade de Vigo (Servizo de publicacións), 843-853.
- Stein G. (1991). "The Phrasal Verb Type 'to have a look' in Modern English", *IRAL* 29 : 1- 29.
- Wierzbicka A. (1982). "Why can you Have a Drink when you can't *Have an Eat?", *Language* 58 : 753-799.