UNIVERSITY OF MURCIA

International Journal of English Studies

IJES

http://revistas.um.es/ijes

What did Joseph Wright mean by *meaning*: The complexity of lexical semantics in the *English Dialect Dictionary Online*

MANFRED MARKUS
University of Innsbruck (Austria)

Received: 17/11/2019. Accepted: 04/03/2020.

ABSTRACT
EDD Online, the online version of Joseph Wright's English Dialect Dictionary, was completed by a project team at the University of Innsbruck in 2019. The sophisticated search-engine of the new interface 3.0 reveals the multi-faceted role of semantics in dialect words. Its complexity is due to both the fuzziness of lexical forms and the ambiguity of their meanings. This paper, beyond the theory-biased "complexity debate", supports the opinion that traditional regional dialects, qua low-contact varieties, have developed a higher degree of lexical complexity than high-contact varieties, i.e. pidgins and creoles, and, in terms of word formation, than the Standard variety of English. The paper first discusses the often polysemous or homophonous meanings of headwords, then of strings within word compositions and phrases. The lemmas also sometimes turn out to be (bound) morphemes or variants. A major aspect in this paper is the wealth of figurative meanings in dialect. This is simply due to the essential role of iconicity, that is, a result of the fact that dialect speakers ("people") want to "see" in their minds what they mean.

KEYWORDS: complexity, semantics, EDD Online, English dialectology, lexicology, meanings in EDD.

1. INTRODUCTION

In the *English Dialect Dictionary* (*EDD*), published in six volumes by Joseph Wright from 1898 to 1905, the author regularly refers to the various dialectal "meanings" of his headwords, no matter whether formal/structural criteria or semantically functional ones are involved. For example, the entry for COMB, sb.1, v. first lists ten word combinations, then three phrases and, at a distance, three derivations, all provided with explanations of their senses. Six phrases with the verb *to comb* are provided separately, as is the past participle form *kem't.* These essentially formal criteria of arrangement are presented on an equal level with ten definitions so that the total number of "meanings" amounts to fifteen.[1]

Address for correspondence: Manfred Markus. Faculty of Letters, University of Innsbruck, Innrain 51, A6020 Innsbruck, Austria; e-mail: manfred.markus@uibk.ac.at

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Print ISSN: 1578-7044; Online ISSN: 1989-6131
doi: 10.6018/ijes.403481

From a modern linguistic point of view, this concept of "meaning", with form and function of a word mixed up, is certainly questionable. However, a close investigation of the "meanings" in the *EDD*, now easily accessible via *EDD Online 3.0* (eddonline-proj.uibk.ac.at), shows that Wright's pre-modern terminology was, up to a point, justified by the complexity of his object of description: the English dialect lexis (including phrases) from 1700 to 1904.[2] Shortly after the publication of the *EDD*, Saussure developed his well-known theory of *signifiant* and *signifié* being kept categorically apart (Saussure, 1973: 99). Notwithstanding the remaining validity of this principle and the enormous impact that it was to have in the 20th and 21st centuries, the strict separation between lexemic form and meaning in a dictionary is more easily said than done. This paper argues that the structure of dialectal English lexis of the Late Modern English (LModE) period is subject to a remarkably high degree of semantic complexity, different from that in the English Standard. The paper breaks new ground because the "complexity debate" of the last two decades has mainly been a theoretical discussion, which has shown little interest in traditional dialects (cf. Miestamo, Sinnemäki and Karlsson, 2008; Sampson, Gil & Trudgill, 2009) and with only a marginal concern for semantics[3]. On the other hand, dialectology, in the old sense of areal or regional dialectology, has rarely ventured into (modern) semantics, nor has it found ways of studying English dialects synoptically, that is, without the neogrammarian focus on specific dialect areas and on narrowly defined linguistic phenomena. This paper tries to start filling this gap, less by taking part in the theoretical discussion[4], but mainly by using the "corpus" now available in the form of *EDD Online 3.0*.[5]

2. WHAT IS LEXICAL SEMANTICS?

Lexical semantics is "the study of the meaning of words, phrases, and lexemes, especially in sets rather than in isolation" (McArthur, 1992: 600; similarly, Yule, 2006: 100). This brief quotation from an authoritative handbook article naturally ignores the problem of the fuzziness involved in the terms used in the definition: meaning, words, phrases, etc. Beginning with the first term and trying to grasp lexical meaning, quite a number of technical keywords come to mind in line with different scholarly approaches that have been adopted in recent semantics: features, roles (such as agent), sense relations (such as synonymy, antonymy etc.), different form-sense relations (such as homophony and polysemy), prototypes, figurative extension (i.e. metaphors, metonymies, meronymies), lexical gaps and, in a historical approach, the issue of semantic change. Moreover, semantic issues across languages and dialects are liable to increase the complexity of lexical semantics. There is no need here to elaborate on the details of all these concepts and to list the names of scholars affiliated with them. However, Fenk-Oczlon & Fenk (2008: 56) have convincingly argued that a small lexicon and a low complexity in phonology, morphology and syntax, as found in pidgin and creole languages, may be compensated for by "a tendency to

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
Print ISSN: 1578-7044; Online ISSN: 1989-6131

homonymy and polysemy" of expressions (56) as well as "high context sensitivity" (57) and "non- conventionalised metaphors" (58). This paper argues that such compensatory features amending "deficits" of low complexity on some linguistic levels may be even more typical of "low-contact" varieties, i.e. of traditional (English) dialects.

Anticipating and rejecting the fundamental tentativeness of semantics, Ludwig Wittgenstein, in his *Philosophical Investigations* (1934: 43), coined his *tabula rasa* definition that was to become famous: "The meaning of a word is its use in language". ('Die Bedeutung eines Wortes ist sein Gebrauch in der Sprache'.) In a more pointed form, Humpty Dumpty had previously offered Alice his famous "explanation" of semantics (in Lewis Carroll's *Through the Looking-Glass*, 1871: 267): "When I use a word (...) it means just what I choose it to mean — neither more nor less".

This paper suggests that "meaning" in language may meander between rather subjective Humpty-Dumpty definitions and the semantic theoretical constructs elaborated for standardised languages. Within certain niches of the language system (McArthur's "sets"), such as flora/fauna, onomastics, kinship and colour terms, structuralist, cognitive and psycholinguistic constructs have proved helpful, and, up to a point, this should hold true for the systematic quality of dialects. On the other hand, in low-contact varieties of English (i.e. traditional dialects) semantic complexity seems particularly extreme and, in my opinion, of a special kind. The main purpose of this paper is to address this dialectal complexity and to trace some of the patterns in it.

3. COMPLEXITY IN LOW-CONTACT (VS. HIGH-CONTACT) VARIETIES OF ENGLISH

As mentioned above, the complexity of English as a language system has been a controversial scholarly issue. The debate started around the beginning of this millennium. In view of the many English varieties worldwide now under discussion, Trudgill as early as 2001 convincingly drew the typological line between high-contact and low-contact varieties of English, a distinction again applied in view of complexity by Kortmann and Szmrecsanyi (2009). High-contact varieties are the various post-colonial varieties of English, that is, pidgins and creoles, as well as the non-Standard urban varieties in the British Isles. Pidgins and creoles, and to some extent urban varieties, have been the output of learning strategies of adults and, thus, of strategies of simplification. Low-contact (rural) varieties, on the other hand, are the traditional regional dialects of English, acquired by native-speaker children in line with natural L1-acquisition and marked by various processes of "complexification" (Trudgill, 2001: 371).[6]

Wright's *EDD* covers the period from 1700 to 1904, a time when dialects (*qua* regionalects) were still in common use and were relatively unaffected (compared to the subsequent period) by

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
Print ISSN: 1578-7044; Online ISSN: 1989-6131

the repercussions of industrialisation and by sociolectal influences. One may therefore feel encouraged to expect dialectal complexity in its many facets. To describe these in detail within this pilot paper is impossible, the more so since linguistic complexity cannot be limited to grammatical features (which is what research so far has primarily focused on, cf. Kortmann & Szmrecsanyi, 2009: 3, 15), but, naturally, also refers to semantic and pragmatic aspects of the language system and of language use, respectively. The following section provides a short survey of the dialectal complexity of semantics.

4. TYPES OF SEMANTIC COMPLEXITY IN DIALECTS

Artificial languages, such as computer languages, and certain sections of natural languages, for example, nomenclatures and technical terminology, show a tendency to a one-to-one correspondence between form and meaning. However, this principle suggesting an ideal world of denotation does not hold for the greater part of natural languages and even less for their dialects. Here are the main "troublemakers" causing semantic fuzziness:

polysemy: two or more semantically related words with the same form (to walk – to walk the dog)
homonymy: two or more apparently unrelated meanings coming with the same form (*bank*)
bound morphemes: word-partial forms with an identifiable meaning, in the EDD sometimes lemmatised when frequent (e.g. a-)
free morphemes: their role in the formation of compounds/combinations and phrases (in the *EDD* mostly within entries)
variants: in the *EDD* either classified as headwords of autonomous articles or listed within entries
lexical gaps in Standard English, filled inconsistently in dialects (cf. Fischer 2000)
figurative expansion: the literal meaning of a lexeme is transferred to something contextually or psychologically connected with it (metonymy; metaphor; meronymy)

In essence, the complexity of dialect lexemes consists in the fuzziness of both factors of the semantic relationship: form and meaning. The form may not be that of a typical word, but may depend on "cotext" (its contiguous surroundings) and be linguistically attributable to some pattern, for example, of word formation. Alternatively, meaning may not clearly be identifiable but may be ambiguous, or hard to identify (lexical gaps). The hypothesis of this paper is that English dialects as a whole have more such semantic imponderables than the English Standard. The reason seems quite simple: dialects always have had to function within their regional micro-contexts and are mainly used in spoken communication, where tone, mimicry, gestures and, above all, pragmatic familiarity are supportive communicative means. Structural ambiguity is disambiguated by the strong role of situational features and by the small-world limits of (regional) dialects as used in relative isolation from each other.

Evidence for this structural complexity is, however, difficult to come by. Previous (mostly neogrammarian) dialectological studies have focused on dialect geography of singular dialects

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
Print ISSN: 1578-7044; Online ISSN: 1989-6131

(e.g. in McMahon, 1994: 226-232) or on singular word forms (e.g. Hickey, 2017, in a paper on *grand* in Irish English). *EDD Online* allows for a more comprehensive approach, tracing not a specific dialect as opposed to others, but dialect as a whole, i.e. as a linguistic "genre", in the way literature has developed different "genres".

5. THE COMPLEXITY OF HEADWORDS IN *EDD ONLINE*

EDD Online is based on a digitised version of the dictionary text and allows access to a great number of linguistic features by way of its sophisticated software/interface. The often extremely substantial entries, sometimes extending over pages, have been analysed by an Innsbruck research team under the directorship of Manfred Markus according to modern linguistic criteria, among these semantics and pragmatics.[7] The basic search mode refers to the headwords or strings within headwords. As Figure 1 shows, semantic complexity already begins here.

Figure 1: Interface with retrieval list and entry window after a search for *head* as part of headwords.

The search for the (implicitly truncated[8]) string *head* produces 132 lemmas, listed alphabetically. Wright's criteria of lemmatisation are not entirely transparent, but the fact is that many of the headwords are compounds, combinations, phrases or even spelling variants. The *OED* (online version), by comparison, has only five lemmatised headwords containing the string

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
Print ISSN: 1578-7044; Online ISSN: 1989-6131

head. Obviously, the two dictionaries have pursued different policies of lemmatisation, with the *OED* apparently containing relatively more material within its entries.

However, Wright, too, has only lemmatised a minor part of the various types of word formation, which means that the greater part is "hidden" within entries. An *EDD* query for compounds with *head*, for example, produces 249 matches. The results for combinations, derivations and phrases amount to 269, 22 and 182, respectively. In sum, there are 722 combining lexemes or phraseologisms including the string *head*. Drawing a provisional general conclusion here, one may tend to interpret this basic challenge of a quantity-based concept of complexity as evidence of the greater productivity of everyday words, such as *head*, in dialect.[9] Similar findings are found for headwords with *foot* (44), *face* (22), *hand* (80), *house* (53) and many others. By comparison, the *OED* has only three lemmas for *foot*, two for *face*, two for *hand*, and five for *house*. However, one should consider that the *OED*, apart from these headword results, also allows for a search routine which simultaneously includes all types of word formations and phrases. The overview of Table 1, for the sake of a transparent comparison, presents the figures for all the *EDD* words and types of word formation (+ phrases), contrasting the total numbers with those of the *OED*.

	comp.	comb.	deriv.	phrases	headwords	tot. EDD	tot. *OED*
head	249	269	22	182	132	**854**	**202**
foot	71	171	12	76	44	**374**	**97**
face	86	51	4	40	22	**203**	**52**
hand	99	234	8	217	80	**638**	**150**
house	221	87	2	59	53	**422**	**227**

Table 1. Quantitative results for five test words, with inclusion of types of word formation and phrases.

The differences between the *EDD* and the *OED*[10] (see the last two columns, in bold) are suggestive enough to arouse scholarly curiosity but do not provide full evidence, given the different sizes of the two dictionaries – six vs. 22 volumes (last printed versions), or some 75,000 vs "more than 600,000 words" (see *OED* home page). In order to make up for this lack of argumentative force, one could relate the "total" figures of Table 1 to the overall role that *head* and the other keywords play in a full text search: 7,249 (*EDD*) vs. 17,554 (*OED*) for *head*, 2,480 vs. 7,726 for *foot*, 3,000 vs. 11,572 for *face*, 7,498 vs. 16,851 for *hand*, and 5,707 (incl. 3,226 in quotations) vs. 20,785 for *house* (plus 38,273 in quotations). These figures reflect the clearly larger role of the keywords as such in the *OED*, apparently mostly as part of quotations or definitions.[11] Given the full-text figures, the numbers of Table 1 appear all the more striking: in spite of its far larger size, the *OED*, compared with the *EDD*, lists only about half to a fourth of the words containing our randomly selected sample strings. The logic behind these results is

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
Print ISSN: 1578-7044; Online ISSN: 1989-6131

simple: not knowing let alone using all the hundreds of thousands of words that English provides[12], dialect speakers would tend to use the common ones more frequently and more flexibly – in compounds, derivations, combinations and phrases.[13]

Of course, the sum of all dialect words included in the *EDD*, on which our quantitative analysis is based, is a theoretical construct beyond the concrete speaker's competence. However, *langue*, the Chomskyan *competence* of the ideal native speaker, is no less a mere theoretical construct. This paper, fair enough, focuses on English dialect as a theoretical construct and on the motivation of English dialect speakers in general, rather than on features of isolated dialects or individual dialect speakers. If this approach is valid, our first type of semantic complexity can be defined as the meaning of base words to be subject to frequent modification in dialect because of their integration in some compositional word or phrase pattern. Word formation and phraseology, then, are the first domain of semantic complexity.

6. THE COMPLEXITY OF SYNONYMS (AND THEIR COLLOCATIONS)

Pickl (2013: 63), among others, has shown that lexical fields are subject to dialectal variation in different ways, depending on the affinity of the practices involved in such fields to spatial limitation or expansion. Lexical fields/terms of emotional involvement, for example, encourage distributional irregularity and diversity. Without going into detail on this theory, we can now test it by checking terms of definition that are liable to be affected by the emotionality of language users, e.g. *woman*.

Figure 2: Terms of the semantic field *woman*, with one of the retrieved headwords opened.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
Print ISSN: 1578-7044; Online ISSN: 1989-6131

The very number of matches in Figure 2 (831) is enormous. A search for the plural form *women* would add another 135 passages of definitions. Given this overwhelming number of nearly a thousand results, with 858 headwords concerned, users may prefer to focus on terms for women under a certain aspect, such as time or place. Figure 2 shows at the right top that eight search filters are available, one of which is time spans. To free our search from the diachronically conditioned part of the complexity, for example, we could limit it to the short time span from 1900 to 1904. Naturally, in a dialect dictionary one could also focus on a certain dialect area. Moreover, there are filters *parts of speech*, *phonetic*, *etymology*, *usage labels*, *sources* and *morphemic* as well as a button called *last result* (in the search box). The latter device could be applied in view of the adjectives collocating with woman/women in definitions. In the selected example of Figure 2 (BESOM), there is a reference to "a woman of loose or slovenly habits". This definition could lead users to suspect that *loose* was not a rare connotation or collocate of *woman*, a suspicion that is, indeed, confirmed by the *last-result* routine (60 matches[14]). Figure 3 provides an example of this "piggy-back" mode.

Figure 3: Search for *woman* as a term of definition plus a "last-result" search for *loose*, again as a term of definition.

One can repeatedly apply the *last-result* mode, which invites serendipitous questions, raised by scholars' curiosity, on typical connotations (see, after *woman*: *gossip*, *old*, *young*, *married/unmarried*, etc.).

Apart from connotations and collocations, a wealth of synonyms testifies to semantic complexity. There is a very direct way of finding synonyms in *EDD Online*. The filter *usage label*, in its sub-filter *semantics*, allows for the activation of the keyword *synonymy*, which covers various specific keywords and abbreviations such as *synonymous*, *equivalent*, *equiv*. etc. Figure 4 shows an extract of the 210 headwords that are marked in the *EDD* as synonyms.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
Print ISSN: 1578-7044; Online ISSN: 1989-6131

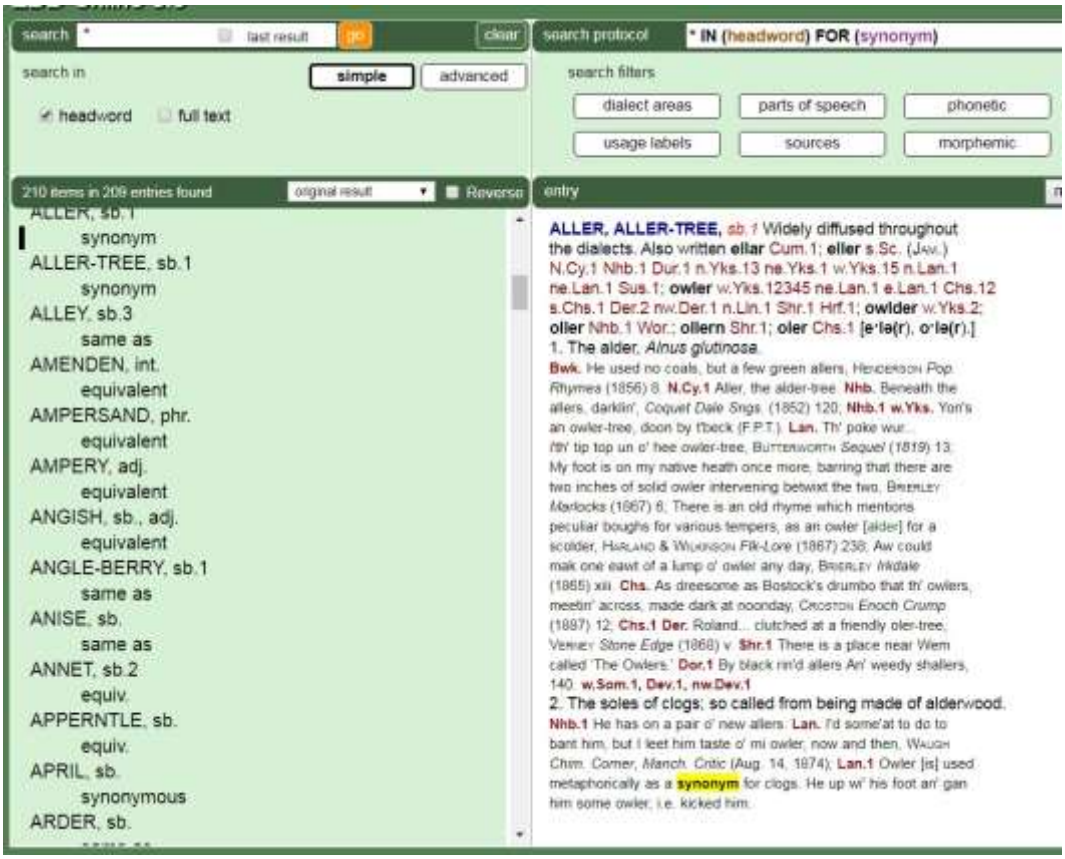

Figure 4: Search for synonyms via semantic filter.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
Print ISSN: 1578-7044; Online ISSN: 1989-6131

The selected example, ALLER, is a synonym for "clogs" (metonymically from *alder*, '[the wood of] the tree'). The question arising of where Wright geographically traced this synonym and for what time, can easily be answered by the addition of the filters *dialect area* (e.g. English counties) and *time span*. The dialect areas are visualised by the *map* option (Figure 5).

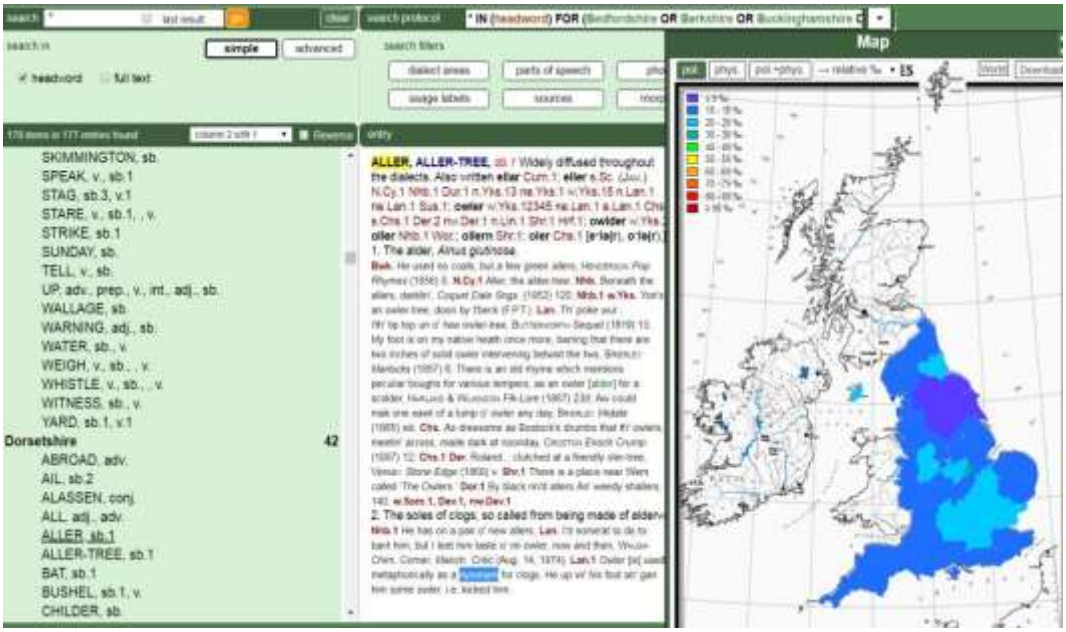

Figure 5: Search for synonyms combined with *English counties*, with map of distribution of *aller*.

These few examples of connotational and collocational productivity as well as lexical richness as a result of synonymy may suffice to demonstrate the ubiquity of variation by synonymy in English dialects. The issue of synonymy has, not surprisingly, crossed our line of thought because different synonymous variants favour different collocates. The lexeme *aller* of Figure 5, for example, has produced six combined lexemes, which are all semantically connected with the alder tree. In-depth studies on this lexical diversity now seem easily feasible. The references in this section to some of the available search routines of *EDD Online 3.0* may serve as a first step encouraging further, more detailed studies.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
Print ISSN: 1578-7044; Online ISSN: 1989-6131

7. THE COMPLEXITY OF POLYSEMY/HOMOPHONY

In Figure 1 above, *HEAD* functions as a random example of quantitative complexity. The example raises the general question of how one can find *EDD* words of this kind. A first step could be the search for Wright's metalinguistic use of the term *meanings* (plural!) in his definitions (Figure 6).

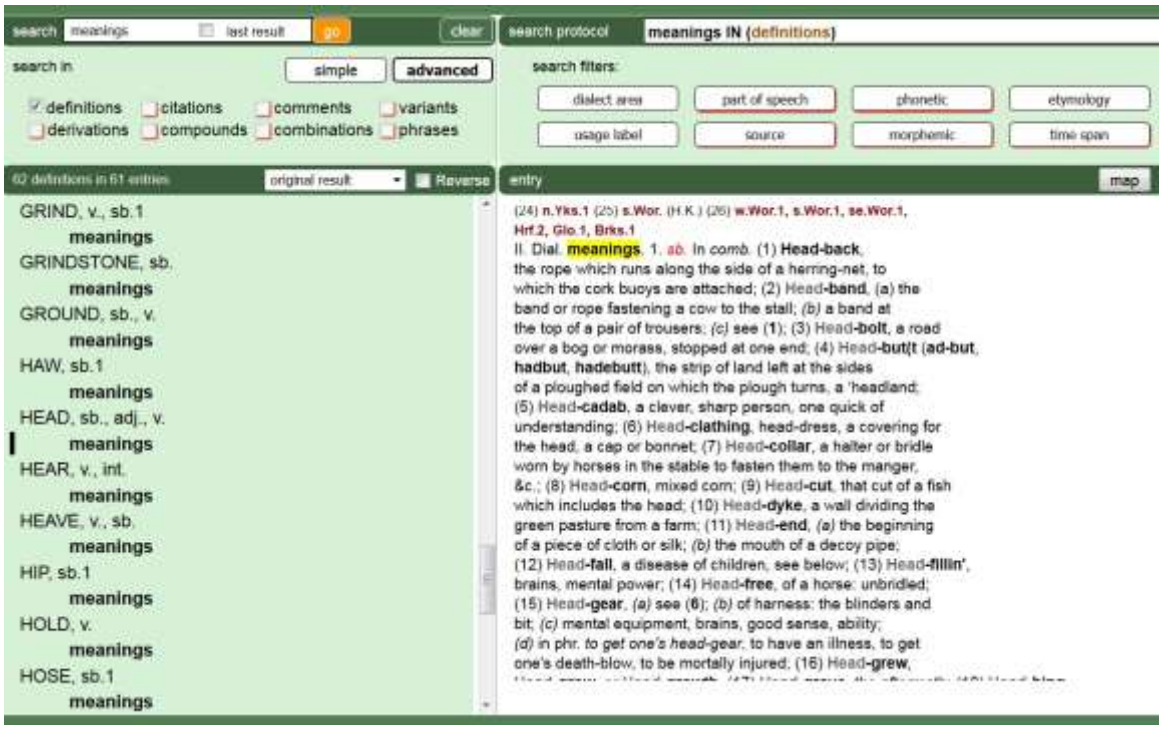

Figure 6: Search for string *meanings* in *definitions.*

Figure 6 shows that the lexeme HEAD is among the entries retrieved by Wright's use of the plural form *meanings*. Therefore, sticking to this example of polysemy or homonymy, we may now go through the 26 meanings of the entry HEAD and notice the fundamental role of figurative expansion (see Figure 1, right half). When we use the term *head* for the hair of the head or for the mouth of a speaker, these are cases of metonymy. Metaphorical use seems even more frequent: the term *head* refers to the upper part of a stem of corn, as well as to the top of grass, beer and milk, each time denoting something different: the ears of grain (corn), the growth at any given time (grass), the froth or foam on ale, and the cream on the surface of milk.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
Print ISSN: 1578-7044; Online ISSN: 1989-6131

Figurative use in the *EDD* is, however, easier to trace with the help of the filter *usage label*, subtype *semantics*. Wright has often tagged words as *fig*. or by some other abbreviation standing for figurative use. Figure 7 shows the result of a query for *foot* (not to use *head* again) in combination with all available semantic usage labels, of which *figurative* is one (as we will see shortly, there is no need to search for the keyword *figurative* in isolation). The user can see all the labels in the memo box above the filters.

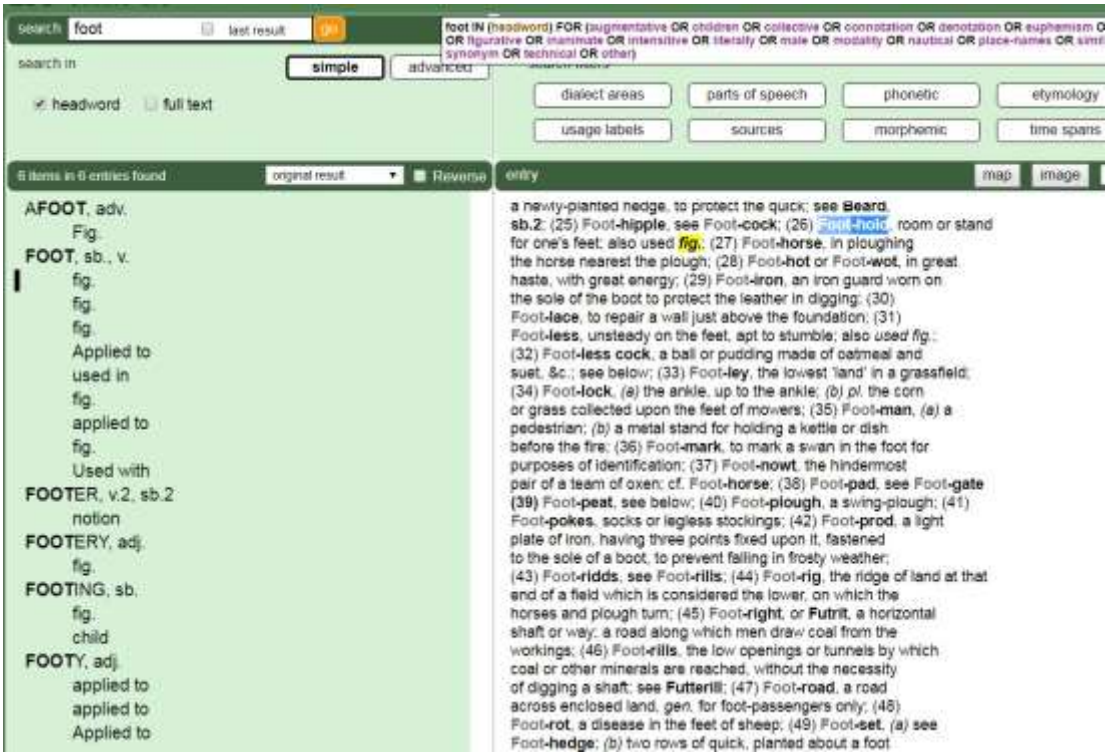

Figure 7: Search for *foot* as a headword, with activated *usage label* filter "semantics".

Figure 7 reveals in its entry window (on the right) that the figurative use selected for illustration does not refer to the lexeme *foot* as a whole, but only to one of its compounds, *foot-hold* (meaning no. 26). In spite of this drawback, given that Wright has marked figurative use in the *EDD* reliably, we still have an excellent tool here for finding both metaphors and metonymies, with a sum total of 3,666 matches of all keyword strings concerned, in 2,739 headwords of 2,737 entries (see Figure 8).

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
Print ISSN: 1578-7044; Online ISSN: 1989-6131

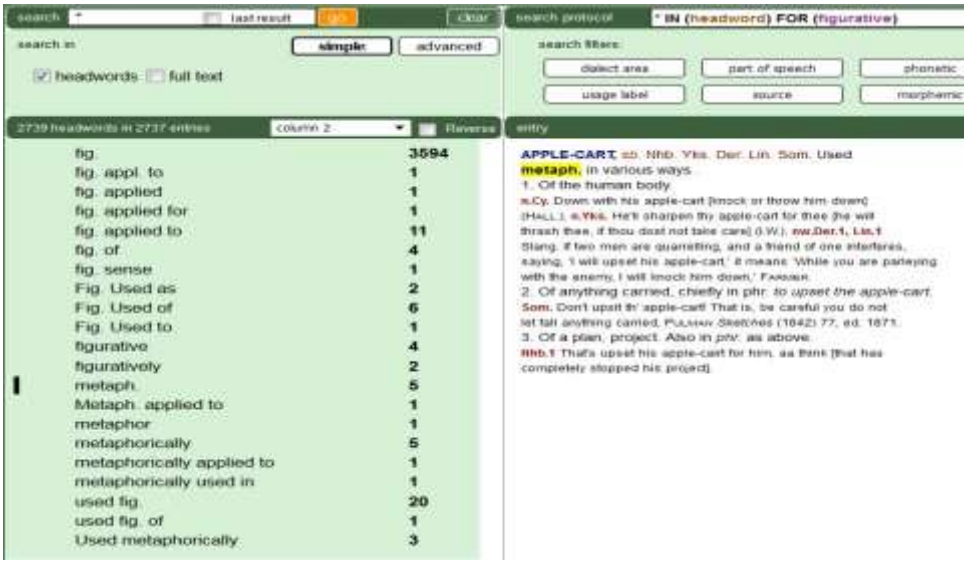

Figure 8: Use of *figurative* keyword strings (in the *column-2* mode of presentation).

Now that the major role of figurative use of dialect words and its contribution to their semantic complexity can clearly be seen, a focus on detailed examples, preferably selected from the domain of everyday life, seems highly desirable. For example, *water* as a string of definition provides 47 items of figurative use, *dog* 34, and *horse* 49.

8. THE MEANING OF (BOUND) MORPHEMES

The semantics of dialects should also be concerned with their morphemes. A lemma search for hyphenated suffixes in *EDD Online* provides only three examples: -AZ, -EN, and – WARD(S. The Dictionary, however, contains many more suffixes (and prefixes), which are worth studying semantically. These are accessible via the *morphemic* filter, which allows queries for 47 suffixes, for example *–able* (see Figure 9).

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
Print ISSN: 1578-7044; Online ISSN: 1989-6131

Figure 9: Result list of search for suffix *-able* (automatically in headwords and derivations) via filter *morphemic.*

The computer does not really search for morphemes as such, but only for strings that "look like" morphemes, so that the result list may include some invalid findings, such as CABLE in the present case. Nevertheless, users will appreciate the list of nine common prefixes (such as *a-*) and 47 selected suffixes (such as **dom*) provided in the *morphemic* filter. Of course, users may complement these lists by typing into the search box any morphemic strings of their own. In any case, the inclusion of derivations subordinated in an entry is an important factor of the complexity of word formation in dialects – as the above analysis of the common keywords *head*, *foot*, etc. suggested. A simple search for all *derivations* in *EDD Online* provides no less than 11,636 matches. The corresponding figure in the *OED* (online version) would certainly be much higher if we could retrieve it, but there is no separate search routine for derivations in the *OED*.

The limits of a quantitative comparison of the *EDD* with the *OED* are obvious here, not only for the reason of differences in the policy of lemmatising in the dictionaries just mentioned. The same holds true for comparisons with other dictionaries, such as DARE (*Dictionary of American Regional English*).[15] Of course, the general productivity of derivations as a pattern of word formation cannot be questioned for any specific time in the history of English, whether in the Standard variety or in dialect. The question is when derivational affixes were particularly productive, and which ones were concerned by this productivity.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

The output figures of the *OED* are a bit less overwhelming if we limit the parameter "Date of entry" to the time span covered by the *EDD*, 1700 to 1904. However, even for suffixes inherited from Old English, such as *-dom*, *-hood* and *-ship*, the *OED* still provides roughly ten times as many headwords as the *EDD*.[16] However, if one focuses on suffixes that are known to have lost ground in the history of English, such as the diminutive suffixes (cf. Markus 2010), the comparison is worthwhile: *-el* 1,151 (vs. 500 in the *OED*), *-et* 1,190 (vs. c. 1,000 in the *OED*), *-ie* 2,097 (vs. c. 500 in the *OED*) and *–ling* 933 (vs. 600 in the *OED*). A promising query could also be carried out on the suffix *–in/-in'* (for Standard *–ing*).[17] Another fascinating issue would be a comparison of different nominal suffixes expressing abstractness, such as *–ment*, *-ness*, *-dom*, *-tion*, *-age*, *-ity* and the like. *–ment*, for example, was apparently more productive in dialects towards the end of the 19th century than Bauer, on the basis of the *OED*, has put forward for English in general (Bauer, 2001: 9).[18] Such examples may give scholars food for further thought on derivational complexity in dialects.

9. PRAGMATIC COMPLEXITY

Fortunately, the *EDD* reveals Joseph Wright's strong sense of typical usage contexts of dialect words and phrases. Many of the usage labels belong to what we would nowadays call pragmatics. The list of pragmatic keywords offered in *EDD Online* is based on a normalisation of the rather variable terms used in the Dictionary. While this procedure of normalisation has decreased the number of keywords considerably, the list is still remarkably long, thus testifying to the important role of pragmatic features in the everyday language of dialects. Figure 10, with the beginning of the keyword list, is suggestive of the many pragmatic questions that we may now raise. The whole list comprises 141 keywords, which, when all selected at the same time, provide access to 4,571 headwords. To demonstrate a focused query, the search of Figure 10 is based on three synonymous keywords (*abusive*, *curse* and *epithet*).

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
Print ISSN: 1578-7044; Online ISSN: 1989-6131

Figure 10: Pragmatic labels, with a specific selection of the three filter keywords *abusive*, *curse* and [logically: OR] *epithet.*

Figure 10, in the entry window of the left half, displays part of the 379 headwords in the *EDD* with the pragmatic feature "swearing" – not a negligible amount. The list of features, to the extent that it is visible in Figure 10, confirms the general observation that pragmatic keywords are very mixed in kind, being either speaker-/hearer-conditioned or referring to speech acts or text types, including literary genres, such as *ballad*. Alternatively, from the viewpoint of modern linguistics, they would be attributable to English for Specific Purposes (ESP), as in the case of the keyword *army*.

Pragmatics has never been at the forefront in historical English dialectology.[19] Given then that the state of the art is in need of improvement, the easy access to pragmatic features in *EDD Online* deserves general scholarly attention. However, within the line of thought of this paper, one could perhaps go a step further, postulating that the language of dialects, naturally a spoken medium, uses the pragmatic factors of speech more intensely than does the Standard variety of English. It does not seem a coincidence that Wright had to use so many pragmatic features to explain the "meanings" of his dialect words. A general search for all (i.e. 111) pragmatic markers at the same time provides no less than 7,148 items in 7,146 entries, with 9,017 tokens of the keywords concerned. Figure 11 presents the end of the retrieval list in *the column-2-counted* mode, i.e. in an alphabetical arrangement of the pragmatic keywords as types and tokens.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
Print ISSN: 1578-7044; Online ISSN: 1989-6131

Figure 11: Search for all pragmatic labels available in *EDD Online.*

When viewing the long list, a great many fascinating topics come to mind. The short extract visible in Figure 11 alone suggests the study of contemptuous expressions, of markers of affirmation/asseveration and of emotional or exclamatory expressions of various kinds. Unlike the approach in most previous studies on pragmatics in historical English dialects (which has predominantly been semasiological, cf the MLA bibliography), *EDD Online* allows for an onomasiological method on a large scale, namely by paying attention to plethoras of forms (in different dialects) that accomplish similar pragmatic purposes.

10. THE COMPLEXITY OF VARIATION

This paper's approach to, and interest in, English dialects as a linguistic "genre" implies that different forms – lexical, phonological, morpho-syntactic – are, up to a point, used for the same aims at different places as well as by different people, and also, if we include the diachronic axis, at the same places at different times. Variation is the natural outcome of dialectal specification.

The *EDD* abounds in itemised variants (53,970 altogether). As mentioned earlier, they are occasionally lemmatised, but usually listed near their headwords. Figure 12, after a search for *a**-words, shows a selection of the entry ANATOMY, with its variants and their dialect areas added.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
Print ISSN: 1578-7044; Online ISSN: 1989-6131

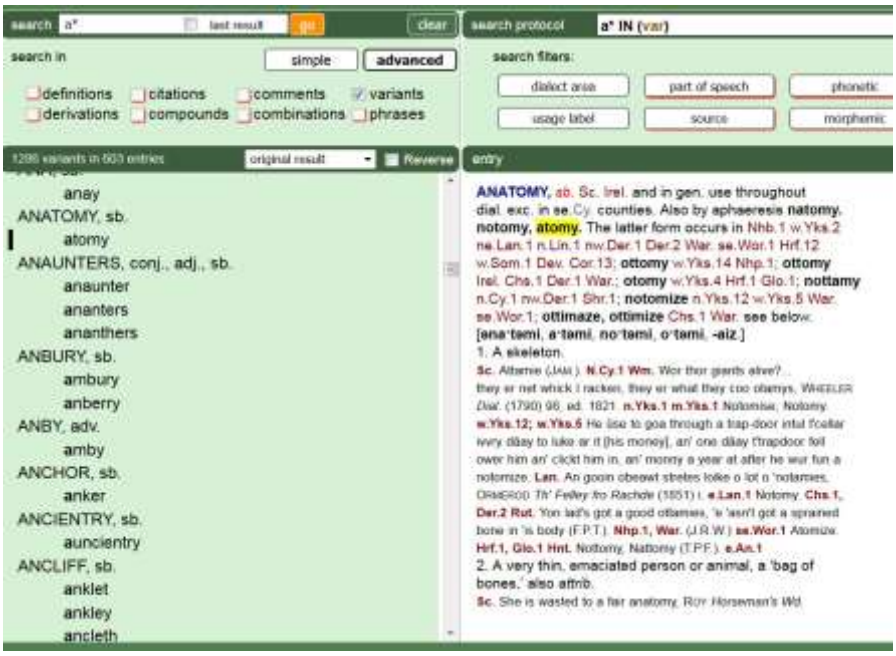

Figure 12: Search for variants of words with initial *a*, arranged according to headwords.

Since certain variants refer to certain objects of meaning in certain dialect areas and at a certain time, the variant distribution indirectly adds up to the semantic complexity of dialect words. For example, in the form *ottomy*, *anatomy* in western Ireland meant "A pigmy, diminutive person" (see sense 3). The information that *ottomy* is a variant of *anatomy* may, of course, be of interest to present-day linguists, but average Irish speakers would hardly have cared about such etymological roots. In other words, as meanings ramify into different sub- meanings, with a certain distance from the original or prototypical meaning involved, so word forms in dialect have more or less diverted from the original form or pronunciation. A close-at-hand question is whether form and meaning may correlate as regards their distance from the "core", that is, the standard form and the prototypical/original meaning. We may assume such a correlation but cannot now prove it. A first precondition for answering this kind of questions would be the precise attribution of variant forms to dialect areas, and this is exactly what *EDD Online* can achieve. In Figure 13, the query refers to all variants found in Worcestershire. The sorting mode *column 2 (with 3)* provides the option of the variants arranged not according to their headwords, but alphabetically and with the dialect areas added.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
Print ISSN: 1578-7044; Online ISSN: 1989-6131

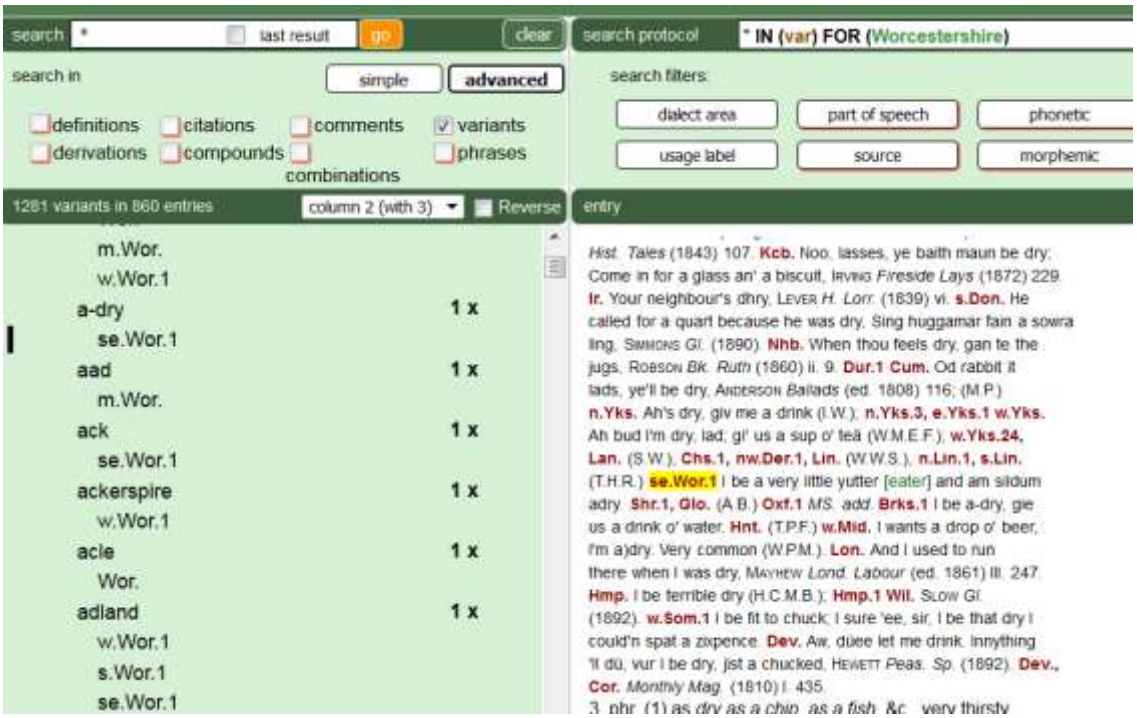

Figure 13: 1,281 variants from Worcestershire rearranged alphabetically, with English counties (here Worcestershire) added.

One can easily see that the form *a-dry* (for 'thirsty'), this typically old form going back to *on-dry*[20], is a variant from southeast Worcestershire. Such findings may give rise to further questions, for example, on new evidence of Worcestershire's linguistic conservativism not only in Middle English (which is well- known[21]), but perhaps also later.

11. CONCLUSION AND OUTLOOK

While "semantic complexity" is a topic with far-reaching theoretical implications, this pilot study has focused on its practical lexicographic and lexicological repercussions, discussing a selection of lexical items of semantic relevance and the way to retrieve them in *EDD Online*. The discussion of these items, unlike investigations in previous studies based on *EDD Online* (e.g. Markus, 2014b; 2018; Krapf, 2017; Ruano-García, 2018) was not meant to be exhaustive, but served the methodological purpose of demonstrating the justification of this paper's topic. The results have supported the author's hypothesis of the greater semantic complexity of the lexis of traditional dialects vs that of the Standard variety of English, but there has been no definite proof as regards the comparison. While it is unquestionable that dialect speakers, compared to Standard

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
Print ISSN: 1578-7044; Online ISSN: 1989-6131

speakers, have less command of the full lexical inventory of a language in its standard version, there is still the question of how and to what extent they may make up for this deficit in language competence. It seems fair to assume that ordinary people generally did not favour the Latinate lexical heritage of the English Standard and substituted home-made words and idioms for them. This paper has only provided indicators suggesting that dialects, though less schoolmasterly settled, less linguistically organised (i.e. full of inconsistencies, signs of "corruption" and folk-etymology, etc.), and less based on education, are yet ready linguistically to answer people's daily needs.[22]

Of course, the proof of such a far-reaching hypothesis can only be provided in future detailed studies on the semantic items addressed in this paper. On the other hand, "English dialects", unlike the normed language system of the Standard, are a conglomeration of systems, naturally complex when seen as a whole.

Looking at things from the point of view of modern semantics and its theoretical bias, traditional dialectology, whose method has always been characterised by its positivist down-to-earth attitude, seems diametrically opposed. This paper takes a middle-of-the-road attitude of cooperation and recommends a willingness to compromise. In semantics, there have been many approaches and constructs in the face of the complexity of "meaning". Classifications of all kinds of lexemes, -nyms, and -types (prototypes, archetypes) have been suggested. Semantics has been slow and hesitant to accept its pragmatic and communicative foundation. Given the data of a comprehensive dialect dictionary like Wright's *EDD*, the analysis of meaning and its separation from the variance of form seem particularly difficult. The reason is that dialect words have widely grown "naturally" and are particularly far away from the linguistic norms prescribed by authorities of states and by individual members of the norm-abiding middle-to-upper classes of society.[23]

Dialect speech, however, has followed norms and rules of its own. Given that traditional dialectology has mostly concerned itself with the areal distribution of words and sounds, it seems high time seriously to analyse these implicit norms of dialect words on a large scale. They should be traceable if we study individual dialect words not in isolation, but in the contexts of their semantic fields and of textual, situational, local, and historical conditions. Lexical gaps and their opposite, semantic density, are an obvious part of this picture.

Considering these different types of contexts, we may redefine the relationship between form and meaning, as well as the simplistic dividing line between competence and performance. While traditional dialects – not only in the United Kingdom – have proved to be a rather rough and controversial field of study and have generally been of moderate interest for Chomskyan linguistics in the past, we may, with the help of the digital corpus *EDD Online*, come to "terms" with them in the future.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
Print ISSN: 1578-7044; Online ISSN: 1989-6131

ENDNOTES

[1] In the context of the *EDD*, COMB is lemmatised five more times, but the argument at issue can do without these other lemmas.

[2] *EDD Online 3.0* is the product of an Innsbruck-based project carried out under the directorship of Manfred Markus from 2016 to April 2019. It is based on previous projects of the Innsbruck team, the first of which started in 2006.

[3] The two papers in Miestamo et al. (2008) that refer to semantic complexity are mainly concerned with language typology, with Fenk-Oczlon & Fenk (2008) using pidgins and creoles for demonstration and Gil (2008) discussing complexity in isolating languages (pp. 55-60, 62- 63, and 109-131). The papers in Sampson, Gil & Trudgill eds. (2009) are equally remote from a discussion of lexical semantics. One of the papers (Deutscher, 2009) clearly questions the concept (initiated by McWhorter, 2001) of an "overall complexity". Trudgill's paper (2009: 98-109) is one of the few to provide concrete examples of complexification from the history of English.

[4] However, this paper draws on the position held by Maas (2009) (in Sampson, Gil & Trudgill eds., 2009) that orality (vs literacy) is an important factor for different sub-types of language complexity. Dialects manifest themselves mainly in spoken language.

[5] "Corpus linguistics" has considerably widened its meaning and changed its methods over the last decades (see Rayson, 2008). As a result, historical dialect corpora have also been used. However, as the survey paper by Grieve (2008: 364-367) confirms, British LModE dialects have hardly ever been in the focus of interest.

[6] Also cf. Trudgill (2009: 104-109), on several English traditional dialect features illustrating complexification. In his paper, Trudgill also discusses interesting sociolinguistic reasons for complexification in traditional dialects (such as sticking to old norms).

[7] A description of *EDD Online*'s potential and of the innovations in version 3.0 is provided in the short *Guide* of the interface (*eddonline-proj.uibk.ac.at*, see *References* below) and in Markus 2019, respectively. In a more detailed way, the possible query options of the interface are reflected in Markus 2020 (in press).

[8] Truncation automatically works at both ends of the string. To exclude it, the user may put the string in quotation marks ("head"). An asterisk (*) causes initial or final truncation.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
Print ISSN: 1578-7044; Online ISSN: 1989-6131

[9] The concept of *productivity* applied here is a very basic one, in line with Saussure (1973: 228: "mots productifs"), with words, rather than other language units or features, being "productive". For various other concepts, see Bauer (2001: 12).

[10] The figures are based on "Advanced search results" in lemmas. This mode includes phrases and the types of word formations that have been kept apart in the *EDD*. The *OED* does not allow for a separate selection of compounds and derivations.

[11] In the *OED*, there is the option of focusing on parts of speech ("combining forms" and "phrases" are on offer) **after** *full text* searches. However, this query routine is, in my view, misleading because it does not systematically refer to the keywords previously at issue (e.g. *head*, *foot*, etc.), but to the whole of an entry. As a result, the conclusion proves justified that the keywords are mostly part of definitions or quotations.

[12] According to the statistics presented by Langenscheidt-Longmans, English has a lexicon of at least 1.25 million words, plus some 9 million words in English for Specific Purposes (Voigt, 1981: 26). The *OED* has lemmatised 620,000 words (January 2020).

[13] Cf Saussure (1973: 42): "Il y a dans certaines vallées retirées des patois qui n'ont … jamais admit un seul terme artificiel venu du dehors". ('In certain isolated valleys there are dialects that have never taken a single artificial term from the outside.')

[14] The combination of two or more defining terms, such as *woman* and *loose*, in the *last-result* mode works in such a way that the entries of the retrieval of the first search (for *woman*) are then used as a starting point for the second search, with the drawback that there is not necessarily a relationship of contiguity between the two terms at issue. There is, however, at least a remarkable likelihood that the one syntactically refers to the other.

[15] The comparison with both dictionaries was strongly suggested by one of the anonymous reviewers of this paper.

[16] The exact figures are: *EDD* 29; 15; 36. *OED* 219; c. 1,000; 300. The team of the 3rd version of the *OED*, in the years since 2007, has had access to the Innsbruck machine-readable data of the *EDD*. One can derive some evidence on the role of the *EDD* for the *OED* from the *OED*'s menu feature "Browse," sub-menu "Sources" (search for *English Dialect Dictionary*). The total number of quotations taken over from the *EDD* adds up to 1,016. There is no evidence on the number of "second-hand" borrowings, triggered by *EDD Online* and then used by the *OED* directly, without reference to Wright or *EDD Online*.

[17] Without its apostrophe, the suffix (or rather the string *-in*) occurs over 100,000 times in a full-text query (*EDD*).

[18] A search for **ment* combined with the *time-spans* filter provides several matches. In the right box of the filter, users can type in the truncated year *1** to retrieve all possible publication years for sources between 1000 and 1904. Then they may re-arrange the retrieval list in the *column-2-counted* mode. There are 1,264 references to dates, most of them from the second half of the 19th century.

[19] For *pragmatics* AND *historical* AND *English dialectology*, the MLA bibliography lists only

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
Print ISSN: 1578-7044; Online ISSN: 1989-6131

nine scholarly (peer-reviewed) papers, most of which refer to very recent developments in English varieties outside England, thus being "diachronic", but irrelevant for the history of LModE. The three papers that do refer to traditional English dialects (mainly the North), have a very narrow focus of subject matter (*thou* vs. *you*; definite article; the adverb *then*).

[20] For the significant role of this construction, see Markus (2014a).

[21] Cf Markus (1990: 393, fn. 11); Fulk (2017: 150). Worcester Cathedral Library was a refuge for Old English manuscripts in the 12th century. Collier (2000: 207), in connection with the "Tremulous Hand", refers to Worcestershire as "a part of England where a comparatively conservative form of the English language was still used and understood".

[22] For a basic comparison of dialect ('idiome local') and the standard variety of a language ('langue littéraire'), see Saussure (1973: 267-8).

[23] This is a reference to the "three hundred years of prescriptivism" characterising the LModE period; see Beal and the other three introductory papers in Tieken-Boon van Ostade and van der Wurff eds. (2009), entitled "Prescriptive and normative concerns". Also, cf Saussure's (1973: 40-3) more basic remarks on the difference between "la langue littéraire" (i.e. the standard variety) and "la langue parlée" (i.e. the spoken language of dialect).

REFERENCES

Bauer, L. (2001). *Morphological productivity*. Cambridge: Cambridge University Press.

Beal, J. C. (2009). Three hundred years of prescriptivism (and counting). In I. Tieken-Boon van Ostade/W. van der Wurff (Eds.). *Current issues in Late Modern English*. Linguistic Insights 77 (pp. 35-55). Bern etc.: Peter Lang.

Carroll, L. (n.d.) (orig. 1871). *Alice in Wonderland and Through the Looking-Glass*. New York, NY: A.L. Burt Company.

Collier, W. (2000). The Tremulous Worcester Hand and Gregory's Pastoral Care. In M. Swan & E.M. Traherne (Eds.), *Rewriting Old English in the Twelfth Century* (pp.195-208). Cambridge: University Press.

Deutscher, G. (2009). 'Overall complexity': a wild goose chase? In G. Sampson, D. Gil & P. Trudgill (Eds.), *Language Complexity as an Evolving Variable* (pp. 243-251). Oxford: Oxford University Press.

EDD Online 3.0: eddonline-proj.uibk.ac.at (accessed 4 January 2020)

Fenk-Oczlon, G. & Fenk, A. (2008). Complexity trade-offs between subsystems of language. In M. Miestamo, K. Sinnemäki & F. Karlsson (Eds.), *Language Complexity: Typology, Contact, Change* (pp. 43- 65). Amsterdam, Philadelphia: John Benjamins.

Fischer, A. (2000). Lexical gaps, cognition and linguistic change. In J. Coleman & Ch. J. Kay (Eds.), *Lexicology, semantics and lexicography. Selected papers from the fourth G.L. Brooks Symposium, Manchester, August 1998* (pp. 1-18). Amsterdam etc.: John Benjamins Publishing Company.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
Print ISSN: 1578-7044; Online ISSN: 1989-6131

Fulk, R.D. (2012) (repr. with corr. 2017). *An introduction to Middle English: grammar; texts*. London: broadview press.

Gil, D. (2008). How complex are isolating languages? In M. Miestamo, K. Sinnemäki & F. Karlsson (Eds.), *Language Complexity: Typology, Contact, Change* (pp. 109-131). Amsterdam, Philadelphia: John Benjamins.

Grieve, J. (2008). Dialect variation. In M. Miestamo, K. Sinnemäki & F. Karlsson (Eds.), *Language Complexity: Typology, Contact, Change* (pp. 362-380). Amsterdam, Philadelphia: John Benjamins.

Hickey, R. (2017). The pragmatics of *grand* in Irish English. *Journal of Historical Pragmatics*, *18*, 82-103.

Kortmann, B. & Szmrecsanyi, B. (2009). World Englishes between simplification and complexification. In T. Hoffmann & L. Siebers (Eds.), *World Englishes–Problems, properties and prospects: selected papers from the 13th IAWE conference* (pp. 265-85). Amsterdam, etc.: John Benjamins Publishing Company.

Krapf, A. (2017). *Wenn man vom Teufel spricht*…Benennungen des Teufels in englischen Dialekten. In H. Bichlmeier & H.-D. Pohl (Eds.), *Akten des XXXI Namenkundlichen Symposiums in Kals am Großglockner (09.-12. Juni 2016)* (pp. 101-122). Hamburg: Baar-Verlag.

Maas, U. (2009). Orality versus literacy as a dimension of complexity. In G. Sampson, D. Gil, & P. Trudgill (Eds.), *Language complexity as an evolving variable* (pp. 164-177). Oxford: Oxford University Press.

Markus, M. (1990). *Mittelenglisches Studienbuch*. Tübingen: Francke Verlag.

Markus, M. (2010). Diminutives in English standard and dialects: A survey based on Wright's English Dialect Dictionary. In M. Markus, C. Upton & R. Heuberger (Eds), *Joseph Wright's English Dialect Dictionary and beyond: Studies in Late Modern English dialectology* (pp.111-129). Frankfurt am Main etc.: Peter Lang.

Markus, M. (2014a). The pattern to be a-hunting from Middle to Late Modern English. Towards extrapolating from Wright's English Dialect Dictionary. In K. Davidse, C. Gentens, L. Ghesquière & L. Vandelanotte (Eds.), *Corpus Interrogation and Grammatical Patterns.* (pp. 57–80). Amsterdam: John Benjamins Publishing Company.

Markus, M. (2014b). Spoken features of interjections in English dialect (based on Joseph Wright's *English Dialect Dictionary*). In I. Taavitsainen, M. Kytö, C. Claridge & J. Smith (Eds.), *Developments in English. Expanding Electronic Evidence* (pp. 116-134). Cambridge: Cambridge University Press.

Markus, M. (2017). The Survival of Shakespeare's Language in English Dialects (on the Basis of *EDD Online*). *English Studies*, *98*, 881-896, also published online 8 September 2017: http://www.tandfonline.com/doi/full/10.1080/0013838X.2017.1365558.

Markus, M. (2018). Corporal Punishment in Late Modern English Dialects (an analysis based on *EDD Online*): How beating has been reflected in 'the language of the people'. *English Today, 34*, 1-10.

Markus, M. (2019). *EDD Online*: What is new in its latest version 3.0. *Dialectologia et Geolinguistica 27*, 103-121.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
Print ISSN: 1578-7044; Online ISSN: 1989-6131

Markus, M. (2020). *English Dialect Dictionary Online*: *A New Departure in English Dialectology*. Cambridge: Cambridge University Press (in press).

McArthur, T, (Ed.). (1992). *The Oxford companion to the English language*. Oxford, New York, NY: University Press.

McMahon, A.M.S. (1994). *Understanding language change*. Cambridge: Cambridge University Press.

McWhorter, J.H. (2001). The world's simplest grammars are creole grammars. *Linguistic Typology, 6*, 125-66.

OED: *Oxford English Dictionary*, 3rd ed., online version: public.OED.com (accessed 6 January 2020).

Rayson, P. (2008). Computational tools and methods for corpus compilation and analysis. In D. Biber & R. Reppen (Eds.), *The Cambridge Handbook of English Corpus Linguistics* (pp. 32-49). Cambridge: Cambridge University Press.

Ruano-García, J. (2018). On the Colonial Element in Joseph Wright's *English Dialect Dictionary*. *International Journal of Lexicography, 32*, 38-57.

Sampson, G., Gil, D. & Trudgill, P. (Eds.). (2009). *Language complexity as an evolving variable*. Oxford: UP.

Saussure, F. de. (1973). *Cours des linguistique générale.* Publié par Charles Bally et Albert Sechehaye (1915), ed. Tullio de Mauro. Paris: Payot.

Tieken-Boon van Ostade, I. & van der Wurff, W. (Eds.). (2009). Current issues in Late Modern English. *Linguistic Insights* 77. Bern etc.: Peter Lang.

Trudgill, P. (2001). Contact and simplification: historical baggage and directionality in linguistic change. *Linguistic Typology*, *5*, 371-373.

Trudgill, P. (2009). Sociolinguistic typology and complexification. In G. Sampson, D. Gil & P. Trudgill (Eds.). (2009) (pp. 98-109).

Voigt, W. (1981). Wörterbuch, Wörterbuchmacher, Wörterbuchprobleme. Ein Werkstattgespräch. Wort und Sprache. In *Beiträge zu Problemen der Lexikographie und Sprachpraxis, veröffentlicht zum 125jährigen Bestehen des Langenscheidt-Verlags* (pp. 24-33). Berlin, etc.: Langenscheidt.

Wittgenstein, L. (2001). *Philosophische Untersuchungen. Kritisch-genetische Edition.* Herausgegeben von Joachim Schulte. Frankfurt: Wissenschaftliche Buchgesellschaft.

Wright, J. (1898-1905). *The English Dialect Dictionary*. 6 vols. Oxford: Henry Frowde.

Yule, G. (2006). *The study of language*. 3rd Edition. Cambridge: University Press.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
Print ISSN: 1578-7044; Online ISSN: 1989-6131