

# Investigating the Notion of Slang in Lexicography

Masaaki OGURA

The University of Tokyo (Japan)

[rrtrnrtrn@ogdict.com](mailto:rrtrnrtrn@ogdict.com)

## REZUMAT: Investigând noțiunea de argou în lexicografie

Dicționarele trec în momentul de față prin multe modificări în alcătuirea lor. Printre acestea se numără descrieri îmbunătățite ale sintagmelor, mărcilor discursive și metodologiei – pentru a enumera doar câteva. Într-o astfel de situație, fiecare dicționar încorporează noi trăsături pentru a se diferenția de altele. Aceste diferențe reflectă atitudinile descriptive față de limbă, lăudându-se cu exemple atestate sau bazate pe corpus. Recent, dicționarele se orientează din ce în ce mai mult spre furnizarea descrierilor care să contribuie la cultivarea competenței comunicative a celor care le studiază. Printre aceste încercări se numără introducerea normelor CECRL în a treia ediție a *Cambridge Advanced Learner's Dictionary*, pentru a indica nivelul vocabularului. *Cadrul European Comun de Referință pentru Limbi* (CECRL), al cărui uz e din ce în ce mai răspândit în zilele noastre, își propune să impună un standard de competență comunicativă. Când aspectul comunicativ e pus în discuție, persoana care studiază vrea să apeleze la etichetele din dicționare, unde fiecare cuvânt e descris ca formal, informal, argou, și așa mai departe. Dintre acestea, argoul ar trebui să fie un concept cheie pentru luarea în calcul a prescriptivismului și descriptivismului în alcătuirea dicționarelor. Deși DUMAS & LIGHTER (1978) au stabilit în mod interesant câteva norme în funcție de care un cuvânt poate fi considerat argotic, nu pare să existe un consens în privința felului în care argoul poate fi definit. Studiul de față investighează felul în care noțiunea de argou e surprinsă în lexicografie, pornind de la comparația între diferite corpusuri, și încearcă să plaseze noțiunea de argou în funcție de continuumul prescriptivism-descriptivism.

CUVINTE-CHEIE: *argou, lexicografie, corpus, prescriptivism, descriptivism*



## ABSTRACT

Dictionaries are undergoing many changes in their compilation. Among these are refined descriptions of collocations, discourse markers, and formality – to list just a few. In such a situation, every dictionary is incorporating new features to differentiate itself from others. These differences reflect descriptive attitudes toward language, boasting of corpus-based or attested examples.

In recent times, dictionaries are becoming increasingly oriented to providing descriptions to contribute to foster the communicative competence of learners. Such efforts include introduction of CEFR norms in *Cambridge Advanced Learner's Dictionary* 3<sup>rd</sup> to show the level of vocabulary. CEFR, which is coming into wide use nowadays, aims to set the standard for communicative competence. When the communicative aspect is discussed, a learner wants to take recourse to labels in dictionaries, where each word is explained as formal, informal, slang, and so on.

Among these, slang should be a key concept to consider prescriptivism and descriptivism in the compilation of dictionaries. Although DUMAS & LIGHTER (1978) interestingly set out some norms for a word to be judged as slang, the consensus on how slang can be defined does not seem to be well reached. The present study investigates how the notion of slang is captured in lexicography, based on the comparison between several corpora, and attempts to locate the notion of slang in relation to the prescriptivism-descriptivism continuum.

**KEYWORDS:** *slang, lexicography, corpora, prescriptivism, descriptivism*



## RÉSUMÉ : Investigation sur la notion d'argot en lexicographie

Les dictionnaires subissent de nombreux changements dans leur compilation. Parmi ceux-ci il y a des descriptions raffinées de collocations, de marqueurs du discours, et de la méthodologie, pour ne citer que quelques-uns. Dans une telle situation, tous les dictionnaires intègrent de nouvelles fonctionnalités pour se différencier des autres. Ces différences reflètent les attitudes descriptives vers la langue, se vantant avec des exemples attestés ou basés sur le corpus.

Ces derniers temps, les dictionnaires sont de plus en plus orientés vers des descriptions contribuant à favoriser la compétence communicative des apprenants. Ces efforts comprennent l'introduction de normes CECR dans la 3<sup>e</sup> édition du *Cambridge Advanced Learner's Dictionary* pour montrer le niveau de vocabulaire. CECR, qui connaît une large utilisation de nos jours, vise à établir la norme pour la compétence communicative. Lorsque l'aspect communicatif est discuté, un apprenant veut avoir recours aux étiquettes des dictionnaires, où chaque mot est expliqué comme formel, informel, argotique, et ainsi de suite.

Parmi ceux-ci, l'argot devrait être un concept clé à considérer le prescriptivisme et le descriptivisme dans la compilation des dictionnaires. Bien que DUMAS & LIGHTER (1978) aient défini quelques normes intéressantes pour qu'un mot soit jugé comme argotique, le consensus sur la façon dont l'argot peut être défini ne semble pas être bien atteint. La présente étude examine comment la notion de l'argot est saisie dans la lexicographie, basée sur la comparaison entre plusieurs corpus, et tente de localiser la notion d'argot par rapport au continuum prescriptivisme-descriptivisme.

**MOTS-CLÉS :** *argot, lexicographie, corpus, prescriptivisme, descriptivisme*

## 1. Introduction



DICTIONARIES PLAY A ROLE in both describing and prescribing words in use. To be prescribed, a word has to be described; to be described, it must first be recorded by lexicographers. Words are always elusive, requiring painstaking work by lexicographers to collect data about their actual usage. This was especially true before the advent of corpus linguistics. Over many years, lexicographers who were sensitive to how language was used, including Samuel Johnson, Noah Webster, and Philip Gove, immersed themselves in this daunting work.

While some earlier dictionaries have been discussed in terms of prescriptive–descriptive issues, this seems untenable as dictionaries are more or less eclectic, and the dichotomy has been revisited by such scholars as BERGENHOLTZ (2003), BALTEIRO (2011), and BEAL (2004). The complexity of this relationship between prescriptivism and descriptivism is informed by various underlying factors, among which is the issue of what constitutes a “standard” or “standard language.” For example, BURKE (2004) distinguishes three types of standard’ language: (1) the public form, (2) the most authoritative form, and (3) the most current and prevailing form of the language. This indicates that the word “standard” is subject to interpretation and remains complicated to define. In addressing this complex concept, one might begin from what is regarded as non-standard, which is likely to include slang. In light of BEAL’s (2004) classification, slang may be seen to fall into BURKE’s (2004) third category, in a community where a particular type of slang is prevalent. However, this does not preclude the need to examine the notion of slang as a label that is often attached to language that some would consider non-standard. The next section explains how slang is defined in some dictionaries, followed by an investigation that uses corpus data in an attempt to capture how slang may be accepted or acquired, taking into account the context and connotations of slang as indicated by the words it collocates.

## 2. Definitions of slang

Slang has been described in several dictionaries for native English speakers, whose purpose is to provide plausible definitions of words. Slang refers to “words that are not considered part of the standard vocabulary of a language and that are used very informally in speech, especially by a particular group of people” (*Merriam-Webster*). Alternatively, slang is “a type of language consisting of words and phrases that are regarded as very informal, are more

common in speech than writing, and are typically restricted to a particular context or group of people" (*Oxford Dictionaries*). These definitions mention "standard," "informally," and "a particular group of people," indicating that slang is confined to informal use within a particular community. Another definition (*American Heritage Dictionary*) offers a deeper explanation by referring not only to a community but to "raciness" and "humor" in attempting to show the social effect of slang:

1. A kind of language occurring chiefly in casual and playful speech, made up typically of coinages and figures of speech that are deliberately used in place of standard terms for added raciness, humor, irreverence, or other effect.
2. Language peculiar to a group; argot or jargon; thieves' slang.

According to the *Collins* definition, "*Vocabulary, idiom, and so on, that are not appropriate to the standard form of a language or to formal contexts may be restricted as to social status or distribution and are characteristically more metaphorical and transitory than standard language.*" This encompasses those aspects of slang as defined in the previous three sources. The *Collins* definition also differs slightly in that it mentions "appropriateness" in relation to standard language, which is worth noting as a value judgment on slang. The three important aspects of these definitions are their reference to "standard," "formality," and a "group or community."

One question that arises, then, concerns what is or is not standard. This issue is highly controversial because, as already noted, words and their meaning are elusive, as is the notion of what is "standard." This elusiveness is what defies any "accurate" definition, and it is therefore helpful to refer to DUMAS & LIGHTER (1978), who state that true slang satisfies at least two features from among the following four.

1. Its presence will markedly lower, at least for the moment, the dignity of formal or serious speech or writing.
2. Its use implies the user's special familiarity either with the referent or with that less statusful or less responsible class of people who have such special familiarity and who use the term.
3. It is a tabooed term in ordinary discourse with persons of higher social status or greater responsibility.
4. It is used in place of the well-known conventional synonym, especially in order (a) to protect the user from the discomfort caused by the conventional

item or (b) to protect the user from the discomfort or annoyance of further elaboration.

(DUMAS & LIGHTER 1978: 14–15)

What is particular to this definition, unlike those of the cited dictionaries, is the phrase “protect the user from the discomfort or annoyance of further elaboration.” This is consistent with the analysis presented in the present study, which explores the kind of “discomforting” elaboration used in slang as a form of euphemism.

### **3. Analysis of the term “slang” based on the Corpus of Contemporary American English**

This section investigates how the notion of slang can be captured, based on corpus-based linguistic data and using the Corpus of Contemporary American English (COCA), which is a balanced corpus of American English (Davis 2009). A first reason for using COCA is to conduct an analysis that is as objective and evidence-based as possible by exploiting its extensive data in an attempt to characterize the nature of slang. A second reason is to cover the wider span of time during which the word “slang” has been used. When researchers use corpora, there are other options such as the BNC (British National Corpus). The choice of COCA for the present analysis reflects its comprehensive time scope and its extensive data. A third reason is to analyze how slang has been captured in the context of American English, as a key interest that motivates this research is Webster’s third controversy, in which the role of dictionaries was fiercely argued (see GREEN 1996), and which is rooted in notions of language in America.

#### **3.1. Genre**

The data set comprised five genres: academic, fiction, magazine, newspaper, and spoken. The occurrence of slang in each section was in the frequency band of 1–1000 [1].

To obtain an overview, the difference in frequency of slang between spoken and written English was measured. Here, *written* refers to the sets of academic, fiction, magazine, and newspaper genres, while *spoken* pertains to the spoken genre. To determine any difference in the occurrence [2] of slang between these two meta-genres, the number of occurrences was calculated and then processed using a log-likelihood test. The formula used was  $G^2 = 2 * \{ \sum A * (\log_e A - \log_e B) \}$ , where A denotes the actual frequency of slang, and B represents the expected frequency of slang. The raw frequency was converted per 1000000

words. Based on the critical value (3.841), there was a significant difference ( $G^2 = 17.8$ ,  $df = 1$ ,  $p < .05$ ). These results indicate that the term “slang” is more frequently used in written than in spoken English, which possibly shows that slang is often a target of discourse about language.

### 3.2. Collocation

Next, for an overview of the context and connotations of slang, collocations consisting of five words before and after a slang word were examined. Generally, four words provide the collocation measurement specified by STUBBS (2002), but to more broadly capture the words collocated with a slang word, a benchmark of five words was adopted. To begin, I collected those words that occurred more than once (880 words) and then eliminated “stop words” that most frequently appear in corpora. In this way, 660 words were obtained, from which the most frequent 100 words were chosen for consideration. Each word was labeled as (1) pertaining to *culture or society*; (2) a *value judgment*; (3) *talking about language*; and (4) *general or other*. Table 1 shows the breakdown of how many times each word occurred. The number of words belonging to each group is (1) 360, (2) 69, (3) 308, and (4) 398. The present study confines its scope to (1) and (2), as the intention is to explain how slang is accepted in terms of its cultural aspects and people’s view of slang.

#### 3.2.1. Words pertaining to culture or society

First, those words related to (1) culture or society were grouped into four subcategories: (a) national or racial, (b) trend or fashion, (c) gender or age, and (d) other. The recorded frequencies were 246 for (a), 40 for (b), 12 for (c), and 12 for (d). It should be noted that (a) was the most frequent of the subcategories, with 246 times of occurrence of words collocating with slang. Among these collocations, the top four words were *American*, *street*, *English*, and *black*. From these findings, it can be concluded that the word slang is frequently used in contexts where people talk about a particular group. It is also worth noting that in the COCA, the word *America* occurs more frequently than words pertaining to English-speaking communities outside the United States, such as British, Canadian, and Australian. It follows that the term *slang* may especially be used when people talk about a particular group to which they belong.

#### 3.2.2. Words pertaining to value judgment

The breakdown of words under the *value judgment* category takes account of whether a word represents a positive or a negative meaning. There are

24 positive words and 45 negative words, which seems to match the general nature of slang. What matters, then, is what kinds of words occur most frequently among these negative words. The three most common types of word among negative words collocating with the term *slang* are *sex* or *sexual*, *crude*, and *profanity*. The words *crude* (meaning offensive or rude, especially in a sexual way) and *profanity* (offensive or religious words used in a way that suggests a lack of respect for God or for holy things) show that slang is used with words that have sensitive connotations related to sex or religion.

### 3.2.3. Words pertaining to language

Among the subcategories for *value judgment*, words about language also occurred with slang. The words most frequent 100 words referring to language are shown in *Table 1*.

*Table 1.* The 100 most frequent words pertaining to language

Words	Counts
term(s)	76
word(s)	62
language	26
dictionary	18
slang	18
phrase(s)	14
jargon	12
rhyming	12
speak	11
speech	10
expression	9
idioms	7
meaning	6
means	6
talking	6
accents	5
dialogue	5
glossary	5

This set of words indicates that slang occurs in contexts where the slang itself is discussed or mentioned metalinguistically as a target of critique, which seems consistent with data indicating that slang inhabits the written rather than the spoken context.

## 4. Conclusion

As discussed in the preceding sections, three tentative conclusions can be drawn from these features of the word *slang*. First, it is used to describe the community to which the user of the term *slang* belongs. Second, it arises in negative contexts concerned with sensitive issues such as sex or religion. Third, the term is used in situations where the users discuss language in a metalinguistic way. Objectives to be addressed in future studies should include subcategorizing words under the category of *other* and carefully examining (in a more qualitative way) the context in which slang is used so as to assess the validity of tentative conclusions drawn from corpus-based study. This process of gaining a concise understanding of slang has merit because it may lead to a better sense of what is “standard,” a controversial but critical issue that remains to be properly addressed.

## NOTES

- [1] Because of restrictions on the use of data, exact frequencies are not revealed here.  
 [2] Because of restrictions on the use of data, exact frequencies are not revealed here.

## BIBLIOGRAPHY

- American Heritage Dictionary*. URL: <<https://www.ahdictionary.com/>>. Accessed 30/11/2015.
- BALTEIRO, I. (2011). “Prescriptivism and descriptivism in the treatment of Anglicisms in a series of Spanish-English dictionaries.” *International Journal of Lexicography*, 24, 3, 277–305.
- BEAL, J. (2004). *English in Modern Times*. Routledge.
- BERGENHOLTZ, H. (2003). “User-oriented understandings of descriptive, prescriptive and prescriptive lexicography.” *Lexicos*, 13, 1.
- BURKE, P. (2004). *Languages and communities in early modern Europe*. Cambridge: Cambridge University Press.
- Collins Dictionaries*. URL: <<http://www.collinsdictionary.com/>>. Accessed 30/11/2015.
- DAVIS, M. (2009). “The 385+ million word Corpus of Contemporary American English (1990–2008+): Design, architecture, and linguistic insights.” *International Journal of Corpus Linguistics*, 14, 2, 159–190.
- DUMAS, B.K., & J. LIGHTER. (1978). “Is slang a word for linguists?” *American Speech*, 53, 5–17.
- GREEN, J. (1996). *Chasing the Sun: Dictionary-Makers and the Dictionaries They Made*. Jonathan Cape.



*Merriam-Webster*. URL: <<http://www.merriam-webster.com/>>. Accessed 30/11/2015.

*Oxford Dictionaries*. URL: <<http://www.oxforddictionaries.com/>>. Accessed 30/11/2015.

STUBBS, M. (2002). *Words and phrases: Corpus studies of lexical semantics*. Malden, MA: Blackwell.

