

PAVOL ŠTEKAUER / DON CHAPMAN / SLÁVKA TOMAŠČÍKOVÁ / ŠTEFAN FRANKO

**WORD-FORMATION AS CREATIVITY WITHIN PRODUCTIVITY CONSTRAINTS:
SOCIOLINGUISTIC EVIDENCE***

Abstract

Productivity has been one of the central topics in the field of word-formation in recent decades. Heretofore, productivity has been mainly, if not solely, discussed in formal terms, such as which affixes can be used with which stems, the productivity of rival affixes, etc. Such a formal approach leaves out the speakers' needs for creating new words. Accounting for speakers' word-formation needs requires a re-evaluation of the notion of creativity. In our approach to word-formation, this notion emphasizes the active role of language users, reflecting the fact that, in each act of naming, there is more or less significant space for a coiner's individual selection out of the options. Since each individual has unequal experiences, knowledge, intellectual capacity, imagination, education, age, professional interests, and so on, one would expect speakers to bring considerable variation to the naming task. Therefore, this article examines the influence of education, profession, and language-background upon the act of naming and the related word-formation productivity. In addition, we will examine, whether and to what degree these factors exert any influence upon the resolution of the fundamental conflict in word-formation (and language in general), namely that between the explicitness of expression and the economy of expression.

1. Introduction

Productivity has been one of the central topics in the field of word-formation in recent decades. It was especially the 1990s and the turn of the millenium that brought new and comprehensive insights into this field, presented by, *inter alia*, H. Baayen,¹ I. Plag (1999), and L. Bauer (2001). Their excellent and seminal studies may be considered a culmination of a long-term effort by derivational morphologists to identify the nature of productivity in word-formation, especially with regard to the deep-rooted belief that productivity of word-formation processes is much lower than that of syntactic and inflectional processes (cf. Chomsky 1970). Gains in recent decades have shown that word-formation is more productive than first thought, when one is careful about the definition of productivity. But this paper argues that those refinements to the concept of productivity have not gone far enough. Heretofore, productivity has been mainly discussed in formal terms, such as which affixes can be used with which stems, the productivity of rival affixes, etc. Such a formal approach leaves out the speakers' needs for creating terms, and leaving out those needs has been precisely what has skewed evaluation of productivity in word-formation. When those needs are taken into account, word-formation seems to be as productive as syntax. It is that claim that this paper will argue.

Accounting for speakers' word-formation needs requires a re-evaluation of the notion of **creativity**. Traditionally, creativity within word-formation has usually referred to idiosyncrasies and deviations from rules. In contrast, the term **productivity** has been used to apply to regular, or rule-governed patterns. Word-formation theory has largely limited

* The authors would like to thank Ingo Plag and an anonymous reviewer for their comments on an earlier version of this paper.

¹ A series of articles. Cf. some of them in the References.

itself to productivity in this sense, as it has mainly considered productivity within a framework of **rival affixes** (or, patterns of word-formation) used in generative morphology, such as *-ity* vs. *-ness*. In our approach to word-formation, however, the concept of creativity applies more generally to any **act of naming** by individual speakers of a language, whether idiosyncratic or regular. Of course a speaker's choices in the act of naming will be constrained by the speaker's language system (langue), but usually, there is more or less significant space for a coiner's individual selection out of the options available for the act of naming related to a particular object of extra-linguistic reality. Since each individual has unequal experiences, knowledge, intellectual capacity, imagination, education, age, professional interests, and so on, one would expect speakers to bring considerable variation to the naming task. Thus, the notion of **rivalry**, more prevalent in traditional, static views of word-formation, is just a part of a much more comprehensive concept of the act of naming, the concept whose focal point is the **active role of language users**.

An examination of **naming needs**, as opposed to the distribution of formal affixes, is accommodated within the onomasiological theory of word-formation developed by Štekauer (1998) and his subsequent publications. In this paper, we will extend Štekauer's notion of a Word-Formation Type cluster to cover three other levels, namely the Onomasiological Type cluster, Morphological Type cluster (see also Štekauer 2003), and Word-Formation Rule cluster. And if indeed speakers vary in their naming strategies according to their different experiences, we ought to find such variation correlating with social variables, such as education, profession, and language background.

A further aim of this paper is to test that assumption. In particular, the integrated theory of productivity presented in this paper will be tested and illustrated in a questionnaire-based evaluation of the influence of sociolinguistic factors upon acts of word-formation, and, by implication, of the word-formation productivity. In addition, we will examine, whether and to what degree these factors exert any influence upon the resolution of the **fundamental conflict** in word-formation (and language in general), that **between the explicitness of expression and the economy of expression**.²

In short, this paper aims to examine the larger notion of creativity that includes a speaker's naming needs within an onomasiological theory of word-formation, and to demonstrate the usefulness of such an examination with a sociolinguistic study of speakers' word-formation choices.

The purpose of Section 1 is to provide a brief introduction to the topic. For more profound analyses of the state-of-the-art in the field, the reader is referred to Plag (1999), and mainly Bauer (2001).

² We are aware of a subtler classification in cognitive linguistics, such as that proposed by Geeraerts (1983) who – at the level of what we label as 'economy of expression' - distinguishes between conceptual efficiency (metaphor, metonymy) and formal efficiency (ellipsis, folk-etymology, avoidance of homonymic clash), and – at the level of our 'explicitness of expression' - between conceptual expressivity (word formation, borrowing, semantic change) and formal expressivity (creation of specific word-formation patterns). See also Grzegą (2002: 1029ff). However, the opposition economy : explicitness fits the purpose of our analysis.

1.1. At the beginning...

The beginnings were gloomy, and the outlook poor. Noam Chomsky (1970) sentenced word-formation productivity to the inferior position, with the master being the productivity of syntax and inflectional morphology. Chomsky emphasized the idiosyncratic semantic and phonological character of nominals derived from verbs, and concluded that the accidental character of word-formation is more typical of lexical structure. According to Chomsky, word-formation processes, unlike syntactic and inflectional processes, cannot be accounted for with productive transformational rules. Not surprisingly, the transformationalist approach to word-formation (such as Lees 1960) gave way to the lexicalist position which unambiguously separates the issues of word-formation from the issues of syntax based on the recognition that “word structure and sentence structure were not governed by the same set of principles, and that they belonged to different modules of the grammar” (Mohan 1986: 4).

But many of Chomsky’s arguments are open to objection. To Chomsky’s argument that specific affixes do not attach to all possible bases, two possible directions of argumentation can be suggested (Štekauer 1998: 84ff). If we pursue the formal approach we can illustrate that the limitations on productivity operate over syntax as much as morphology (as suggested by Di Sciullo & Williams 1987), and these limitations are of the same nature. It is true, for instance, that the suffix *-ion* does not combine with all verbs. But it is equally true that not all verbs can be used in the sentence structure N – V – Object. The limitation permits only transitive verbs to be inserted. Both limitations (syntactic and morphological) are based on the same principle – they pertain to the **combinability** of structural units. For more examples, see Di Sciullo & Williams (1987). These authors seem to have been inspired by the following observation of S. R. Anderson:

It is true that different verbs take different formations (*describe/description, laugh/laughter, recite/recital, etc.*); but the point is that *some* action nominal formation is available for every verb (subject only to semantic limitations). One cannot really say that the diversity of the forms involved is a limitation on the productivity of the process, any more than the existence of varying conjugation classes constitutes a limitation on the productivity of verbal inflection in languages in which these are found (1982: 585-586).

In this connection, S. L. Strauss also maintains that “we cannot really claim that derivational morphology is any more idiosyncratic than the other structure-generating rules”; in addition, “rules of derivational morphology are as regular, both semantically and phonologically, as other generative rules” (1982: 23, 24).

A second line of argumentation is of pragmatic-generative nature. If we concentrate on the generation aspect, both syntax and word-formation respond to some **demand** of a language community, and they are capable of fully meeting the need. In that respect they are absolutely productive. This also applies to their subsystems. Thus, the systems of Word-Formation Types and Morphological Types are capable of providing a naming unit whenever a new item, such as Agent noun (or, more explicitly, a noun denoting a person performing some activity), is required. Then, the **clusters of Word-Formation Types** and **clusters of Morphological Types** (see below, Section 2.2.4) ‘guarantee’ the coining of a new naming unit of a specific semantics whenever such demand arises.

All in all, with the advancements in the theory of word-formation in recent decades, the view of low WF productivity has been gradually modified. As a result, I. Plag (1999: 2)

could stress that “derivational processes are much more *regular* than previously conceived.”

Our claim that word-formation is absolutely productive because speakers can always provide a new naming unit when required is not simply a trivial evasion of the problems of productivity of word-formation that have engaged morphologists for decades. It is instead a recasting of the issues in terms that should lead to more insight as it brings the analysis of word-formation closer to the analysis of other linguistic phenomena, particularly syntax. The pragmatic level of analysis has always been implicit in syntax. That speakers can generate an infinite number of sentences that they have not previously heard acknowledges that speakers do indeed create sentences. There is no effort to keep track of actual sentences or to note all the sentences that could possibly occur but don't. Syntax is not held to be any less productive because some sentences do not occur. The communication needs of the speakers is taken as a given before the composition of sentences is analyzed.

Furthermore, the analysis of those sentences does not proceed on the word level; sentences are not regarded simply as strings of words. Instead sentences are considered to be made up of phrases, which would be another level of abstraction above words. The occurrence of particular strings of words, then, has little to do with productive patterns in syntax. Instead the issue is whether certain phrase types occur and how they combine together to create a sentence. A noun phrase, for example, must be present as the subject of nearly all sentences, yet the composition of that noun phrase – whether a single pronoun, a determiner and noun, a noun with a complement clause, or something else – is unimportant, so long as the phrase is well-formed. Much less do the individual words constituting the noun phrase matter.

In word-formation studies, however, the pragmatic needs of speakers for new words have largely been ignored. So too has an abstract level corresponding to the notion of a phrase. Instead, individual formants are considered as productive or not. By focusing on the naming needs of the speech community and by acknowledging a functional level of analysis comparable to phrase structures, a theory of word-formation ought to account for productivity in word-formation with more coherence.

It is for that reason that this paper proceeds with an onomasiological approach that accounts for speakers' naming needs.

1.2. Potential Words and Naming Needs

The attractiveness of focusing on the naming needs of speakers also comes up with regard to another sticking point of word-formation theory, namely the role of possible or potential words. Halle (1973) had introduced the notion of **overgeneralization** in his generative account of morphology and had called the non-existence of such words in English an **accidental gap**. We believe that the notion of ‘accidental gap’ is misleading, and is due to the purely formal point of view. If this issue is approached from the point of view of the naming demand of a speech community the non-existence of such words is expediently accountable – they are not needed by the speech community. But what about **potential words**? Should a theory of word-formation account for all words that could be generated or just those that have? Linguists from Allen to Aronoff to Kiparsky have grappled with this question. The most comprehensive analysis of the relevant problems is given in Bauer (2001) where the relations between actual, existing, established, possible, potential, and probable words are discussed in detail. Bauer's ideas may be succinctly summarized as

follows:

The notion of **existing word** raises the fundamental problem of for whom and what such a word exists. We agree with Bauer that an existing word must exist for a **speech community** keeping in mind a number of problems connected with this approach that are pointed out by Bauer (the lower limit of speech community, non-occurrence of all existing words in reference works, the identification of the date when a word comes into existence – the first coining or the establishment of the word?, etc.). Bauer (2001: 36) suggests the following definition:

...a word is an *existing* word from the moment it is first coined...The word may be *item-familiar* to individual speakers, without having become part of the norm of the language. A word is *established* once it becomes part of the norm, that is, once it is item-familiar to a large enough sub-set of the speech community to make it worth listing in reference works.

The notion of potential word is, in Bauer's view, closely related with the notion of lexical gap. Importantly, a coinage only occurs if there is a need, a real or perceived gap in the speaker's lexicon: "Productivity is all about potential. A process is productive if it has the potential to lead to new coinages, or to the extent to which it does lead to new coinages. We are aware of productivity only through the new coinages and the patterns of familiar and unfamiliar words coined by the relevant process" (2001: 41).

Bauer further treats the role of naming with his notion of probable words, which are words that are likely to occur. Bauer suggests that **possible word** be defined in terms of the linguistic system while **probable word** by extra-systemic factors (2001: 42).

The questions of potential words and actual words can be seen to hinge on the role of extra-linguistic reality in word-formation. Ignoring speakers' naming needs in favor of formal analysis of the *langue* gives more importance to potential words. Accounting for those naming needs gives more importance to actual words. In presenting our approach to productivity, we will argue in favour of including actual words in productivity computations.

The theory we propose for accounting for the naming needs of a community is a cognitive onomasiological theory. Its fundamental principles are presented in the next section.

2. A cognitive onomasiological theory of productivity

The following approach to word-formation productivity is based on a series of articles and a monograph chapter on this topic, including Štekauer (1994, 1998, 1999, 2001, 2003, 2005a, 2005b), and attempts at providing a comprehensive theory of productivity within the cognitive onomasiological framework.

2.1. General

It goes without saying that productivity is one of the universal properties of language. It is most clearly manifested at the level of word-formation because the productivity of Word-Formation Types and Rules and Morphological Types makes it possible to generate a new naming unit whenever a speech community needs it. From this it follows that word-formation deals with Word-Formation/Morphological Types and Rules which are productive, that is to say, which, from the synchronic point of view, make it possible to

form new naming units whenever need be. Obviously, productivity implies regularity: this enables language users to understand (in an appropriate context) and use new naming units they have never heard before. In the initial period of existence of a new naming unit, regularity can also be used as a kind of mnemotechnics.

2.2. Main Factors

There are several factors influencing an approach to productivity and the resulting shape of any theory of productivity:

- General theoretical framework. This affects answering the questions like ‘What is the place of word-formation in the system of linguistic disciplines?’, ‘Is word-formation a part of syntax (the transformationalist hypothesis), a part of morphology (‘classical’ structuralist theories), a part of the lexical module (lexicalist approach), or is it an independent module as proposed, for example, in Štekauer (1998)?’, ‘What is the relation between the individual modules?’
- Scope of word-formation: Do *cranberry* words, word series like *receive*, *perceive*, *conceive*, *pertain*, *retain*, clippings, and acronyms fall within the scope of word-formation? Is compounding the matter of word-formation or syntax?
- Productivity – of what? Productivity of affixes, WF rules, WF processes, the whole WF module?
- Attitude to possible, potential, actual and established words, and to the overgenerating morphology.
- Are Word-Formation Types/rules productive and regular or is their regularity much lower than that of syntactic and inflectional rules as assumed by Chomsky (1970)?
- Method of productivity assessment. Should productivity be assessed in abstract terms or is it possible to employ precise mathematical methods? In the latter case, what should the computation be based on? Can we employ absolute numbers or should we relate the computation to certain formal or semantic elements? Should the computation be based on the system level (langue) or speech level (parole)? Or, can these two levels be combined for the sake of productivity computation? Thus, should the calculation be based on dictionaries or corpora?

2.2.1. Theoretical Framework – the Place of Word-Formation within the System of Linguistic Disciplines

The cognitive onomasiological theory of word-formation identifies word-formation as an independent and fully-fledged component as illustrated in Figure 1.

The scheme reflects the relations between the individual linguistic components and within the word-formation component itself. It follows from the scheme that the word-formation component is an independent module on a par with any other linguistic module. The scheme represents the crucial triad of relations: extra-linguistic reality – speech community – word-formation component, thus emphasizing the fact which has been ignored by the vast majority of the mainstream word-formation theories, that is to say, that new words do not come into existence in void (as might follow from purely formal theories). Each act of naming responds to a very real and specific naming need (demand) on the part of a member (members) of a particular speech community.

Second, the scheme indicates a direct connection between the word-formation and the lexical components, and an ‘only’ mediated connection between the word-formation and the syntactic components. This makes this model different from those theories which

consider word-formation as a part of Lexicon or a part of Syntax. The relation between the word-formation and the Lexical components is based on close ‘co-operation’. The Lexicon stores all naming units (monemes and complex words, borrowed words, clippings and acronyms) as well as affixes, and feeds the word-formation component with word-formation bases and affixes in accordance with its needs. On the other hand, the word-formation component supplies the Lexicon with new naming units formed in it.

By implication, no new words are generated either in the Lexicon (however, any semantic and/or formal modification of naming units formed in the word-formation component may only take place in the Lexicon) or Syntax.³

It should be noted that word-formation concerns the formation of **isolated naming units** rather than their use (which is the matter of syntax). Word-formation treats naming units as linguistic signs stored in particular semantically and morpho-syntactically defined paradigms in the lexical component. The process of forming new naming units means that the new naming units can be subsequently retrieved from the lexicon for the purpose of sentence formation.

2.2.2. Productivity – O. K., But of What?

Productivity is a term frequently employed by linguists in general, and – like a number of other linguistic terms – it is quite vague, especially in view of the diversity of its ‘applications’. Bauer (2001) demonstrates the ambiguity of this term when he points out that for some scholars particular affixes (Fleischer) are productive, for others, it is morphological processes (Anderson) that are productive; for yet others, it is rules (Aronoff, Zwanenburg); for a very few it is words (Saussure); for some it is groups of processes (Al and Booij, Anderson). Bauer (1983) discusses the productivity of a complete module of grammar; for yet another group of scholars, productivity is a feature of the language system as a whole. Bauer (2001) himself prefers to define productivity as **a feature of individual morphological processes**.

³ For an account of constructions like *around-and-do-nothing-ish*, *leave-it-where-it-is-er lady-in-waiting*, *pain-in-stomach-gesture*, see Štekauer (2001).

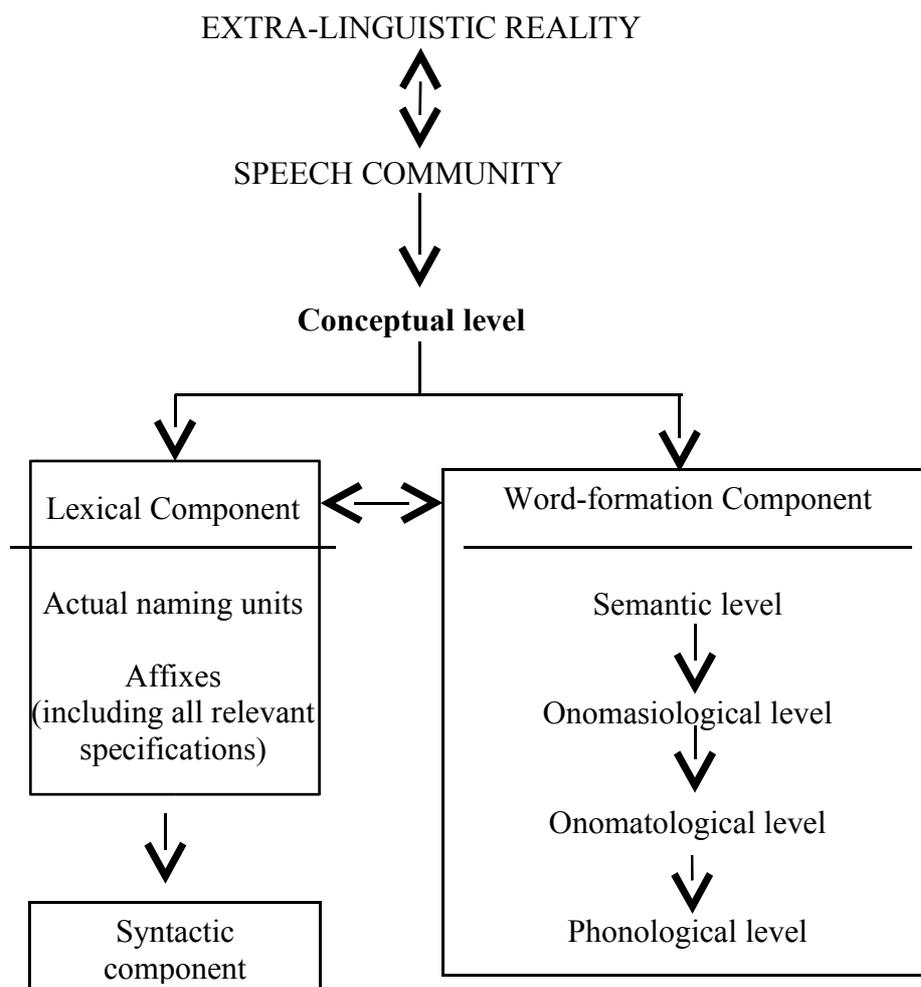


Figure 1: Word-Formation Component and its relation to other components

Dokulil (1962) also presents several possibilities of examining productivity in word-formation:

- (i) The productivity of a word-formation formant (affix). Here he distinguishes between
 - (a) an absolute productivity of a formant, i.e., its applicability in forming new words in general, irrespective of the particular Word-Formation Type it is used in, and
 - (b) relative productivity of a formant, i.e., its applicability in a specific semantic function and/or in a particular Word-Formation Type.
- (ii) The productivity of a Word-Formation Type, in which case a WFT functions as a pattern for forming new words.⁴
- (iii) The productivity of a word-formation base.⁵

⁴ Dokulil (1962: 72) defines Word-Formation Type as a unity of onomasiological structure (Agentive nouns, bearers of Quality, etc.), lexical-grammatical nature of WF base (deverbatives, desubstantives, deadjectives), and formant (words in *-er*).

⁵ Dokulil notes that the productivity of WF base is usually relative: it can mostly be evaluated relative to a particular Word-Formation Type. As such, it is viewed as a condition promoting/reducing the productivity of a particular WF type (1962: 84).

Our approach outlined below discusses productivity at four different, and mutually complementary levels (see Section 2.3).

One of the major disadvantages of various computation methods employed for the evaluation of productivity in word-formation is their limited scope; they are usually restricted to the productivity of affixes. This contradicts the generally accepted scope of word-formation which also includes other word-formation processes. But even if the focus is laid on affixation the existing methods differ in defining the notion of affix, notably in terms of the polysemy – homonymy relation. Both of these facts may significantly distort the results of productivity computation. The prevailing restriction to affixation processes is also reflected in the methodology of computing productivity which seems to be tailor-made to this word-formation process.

It may be proposed (Štekauer 2003) that rather than an **affix-driven productivity** approach a conception is required which, instead of focussing on items (affixes), ranges over all word-formation processes (WFP) (compounding, prefixation, suffixation, conversion, blending, etc.), i.e., one which overcomes the limitations imposed by affixation in particular and by the individual word-formation processes in general. What is therefore needed is a general **WF-Rule-driven theory of productivity** covering the whole stock of complex naming units.

The latter approach forces us into the definition of the notion **Word-Formation Rule** (WFR). Unfortunately, this seems to be another strongly ambiguous term, which, on the one hand, heavily depends on the underlying theoretical background, and, on the other hand, crucially determines the results of productivity computation. Should a WFR be defined in Aronoffian (1976) terms as a combination of a base plus affix, or in accordance with Selkirkian (1982) system based on the maximum level of generalization, such as $X^n \rightarrow Y^n X^a$ for suffixation, Allen's (1978) Primary Compound Formation Rule: $[\#X\#]_N \dots [\#Y\#]_N \rightarrow [[\#X\#][\#Y\#]]$, Kiparsky's (1982) generation of primary compounds by insertion of Y Z into a categorial frame X, i.e., $[Y Z]_X$, or some other formally defined principles? What is the optimum level of generalization in this case? Can a formal definition of WFRs provide a base for covering all word-formation processes? The major trends in research do not favour a positive answer to this question.

Consequently, since the formal base for productivity computation seems to be unable to provide a unified footing for all complex naming units, attention should be, in our view, zeroed in on the **conceptual-semantic** facet. Is such a conceptually and semantically oriented theory of WFRs viable? Should it take the form of separation hypothesis proposed within the framework of Beard's (1995) lexeme-morpheme base theory, or is there any other way of treating WFRs?

2.3. Proposal

The present model departs from a form-based approach to productivity, and proposes to examine productivity within a particular unifying conceptual category (Agent, Patient, Instrument, Negation, Result of Action, Location, Quality, etc.). This approach follows from the onomasiological theory of word-formation: productivity is the matter of **formation** of new words. Each act of naming (as it follows from the scheme in Figure 1) starts at the conceptual level. It is at this level that the 'object' to be named is identified as one falling within the conceptual category of Agent, Patient, Instrument, etc. When the

conceptual category of the named object is identified, the naming process proper starts within which the semantic and morphematic components are identified that will constitute the “naming structure” of the resulting complex word. Irrespective of the numerous variations in the ‘naming structure’, all words denoting ‘persons performing (professionally) some activity’ are words denoting Agents. By implication, we may study the share of the various ‘naming structures’ from different points of view to identify their respective Productivity Rates (PRs). In the onomasiological approach, we can identify four different ‘levels’ of ‘naming structures’, and therefore four **levels of productivity**:

- (i) the productivity at the level of Onomasiological Types (OTs)
- (ii) the productivity at the level of Word-Formation Types (WFTs)
- (iii) the productivity at the level of Morphological Types (MTs)
- (iv) the productivity at the level of Word-Formation Rules (WFRs)

2.3.1. Productivity of Onomasiological Types

The onomasiological model of word-formation (Štekauer 1998, 2001) obliterates the differences between the traditional word-formation processes by proposing a unified basis for the description of word-formation. Such a unified basis makes it possible to objectify the computation of productivity. This **cognitively based model** of word-formation, taking the naming demand of **speech-community** as its starting point, distinguishes five **Onomasiological Types** ranging over the traditional word-formation processes. They are based on the criterion of which constituents of the **onomasiological** (logical-semantic) **structure** are linguistically expressed at the **onomatological (morphematic) level** (see Figure 1). In general, the onomasiological structure includes three basic constituents:

- | | | |
|-----------------------------|--|--|
| (4) Determining constituent | – Determined constituent –
of the onomasiological
mark | Onomasiological base
of the onomasiological
mark |
|-----------------------------|--|--|

where Onomasiological base corresponds to the head of a complex word, and the determined constituent of the Onomasiological mark generally stands for the concept of ACTION. Then, the individual Onomasiological Types can be exemplified as follows:

In **Onomasiological Type 1**, all three onomasiological structure constituents, i.e., the base, the determining and the determined constituents of the mark, are linguistically expressed at the onomatological level by being assigned morphemes with the corresponding meaning. This operation is labeled as the **Meaning-to-Seme-Assignment principle** (MSAP):

- (5) *truck-driver* (A Person (Agent) operates (Action) a vehicle (Object)
Object – Action – Agent
truck drive er
- (6) *house-keeping* (The Process of performing some Action aimed at an Object):
Object – Action – Process
house keep ing
- (7) *signal-generator* (Instrument for an Action producing some Result)
Result – Action – Instrument
Signal generate or

In **Onomasiological Type 2**, the determining constituent of the onomasiological mark is left unexpressed:

- (8) Factitive – Action – Agent
0 write er
- (9) Object – Action – Instrument
0 spinning wheel

In **Onomasiological Type 3**, the determined constituent of onomasiological mark is left unexpressed:

- (10) Result – Action – Agent
novel 0 ist
- (11) Patient – State – Evaluation (Diminutive)
dog 0 ie
- (12) Temporal Stative – State – Patient
summer 0 house

In **Onomasiological Type 4**, the onomasiological mark is simple and unstructured, i.e., it cannot be divided into the determining and the determined constituents.

- (13) Negation – Quality
un happy
- (14) Quality – State
blue-eye ed
- (15) Repetition – Action
re gain

Onomasiological Type 5 (onomasiological recategorization) concerns conversion, and the method of representation of semantic relations between the members of conversion pairs is illustrated by the following examples:

- (16) *bond_N – bond_V*: SUBSTANCE^{Result}ACTION
(in the meaning of a joint)
Interpretation: Substance as a Result of Action
- (17) *switch_N – switch_V*: SUBSTANCE^{Instrument/Result}ACTION_
(in the meaning of a device for completing or breaking an electric circuit)
Interpretation: Substance as an Instrument of Action
- (18) *insert_V – insert_N*: ACTION^{Object}SUBSTANCE
Interpretation: Substance as an Object of Action

- (19) *time_N–time_V*: CIRCUMSTANCE^{Temporal}ACTION
 Interpretation: Action in terms of Temporal dimension
- (20) *clear_A–clear_V*: QUALITY^{Result}ACTION
 Interpretation: Action Resulting in a certain Quality

As indicated above, the present model distinguishes five **Onomasiological Types** ranging over all productive methods of forming new complex words. Since they are based on the criterion of which constituents of the onomasiological structure are linguistically expressed at the onomatological level, the determination of their respective productivities is an important indicator of the **preferences** of language users (or better, coiners) in terms of employing **different cognitive processes** underlying the act of naming, on the one hand, and the **different ways of their linguistic representation**, on the other. The productivity calculation at this level may indicate which of the two universal, contradictory tendencies, i.e., **economy of expression** and **explicitness of expression** (comprehensibility), dominates in a particular language (area). Here we face two gradual oppositions:

- (i) Onomasiological level
- (a) Onomasiological Types 1–3 (complex onomasiological structure)
 - (b) Onomasiological Type 4 (simplified onomasiological structure)
 - (c) Onomasiological Type 5 (absence of onomasiological structure)
- (ii) Onomatological level
- (a) Onomasiological Type 1 (complex morphematic representation of complex onomasiological structure)
 - (b) Onomasiological Types 2 and 3 (economized morphematic representation of a complex onomasiological structure)
 - (c) Onomasiological Type 4 (economy due to onomasiological structure)
 - (d) Onomasiological Type 5 (absolute economy – no morphematic representation).

As indicated above, productivity of the individual Onomasiological Types is given by their respective share of all the complex words that belong to a particular conceptual category (e.g., Agent). From this it follows that the **Onomasiological Type Cluster** is 100% productive with regard to a particular conceptual category as it can ‘produce’ a word belonging to that particular conceptual category whenever a (member of a) speech community needs to give a name to an object belonging to this category.

2.3.2. Productivity of Word-Formation Types

A more specific level is represented by **WF Types**. The computation of productivity of WF Types is also related to a particular **conceptual category**. This makes it possible to include in the computation of the productivity of, for example, Agent names complex words of different onomasiological structures, hence different WF Types (for example, Object – Action – Agent (*woodcutter*); Action – Agent (*writer*); Location – Action – Agent (*street-fighter*); Factitive – Action – Agent (*novel writer*); Instrument – Action – Agent (*anthrax-killer*); Manner – Action – Agent (*slam-dunker*); and a number of other possible WF types).

All of these WF Types may be used to coin new complex words falling within one and the same conceptual category (Agent, in our example), and therefore represent a single **Word-Formation Type Cluster** (WFTC). Any WFTC is – with regard to the particular

conceptual category – 100% productive. Therefore, the productivity of the individual WF Types may be computed internally, within the WFTC, as a share of the individual WF Types of the total number of complex words belonging to the given WFTC.

2.3.3. Productivity of Morphological Types

Any WF Type may have various *morphological representations* (*wood-cutter* (=N+V+er) – *novelist* (N+ist) – *writer* (V+er) – *cheat* (conversion) – *oarsman* (N+s+man) – *transformational grammarian* (A+N+ian) – *bodyguard* (N+N), etc.). All of these different morphological structures represent various **Morphological Types**. Since they are used to coin new complex words falling within one and the same conceptual category (Agent, in our example), they represent a single **Morphological Type Cluster** (MTC). Any MTC is – with regard to the particular conceptual category – 100% productive, and the productivity of the individual Morphological Types may be computed internally, within the particular MTC.

2.3.4. Productivity of Word-Formation Rules

Word-Formation Rules are constituted by the unity of WF Types and Morphological Types. Thus, the conceptual category of Agent category may be exemplified, inter alia, by the following WF Rules:

- (21)
- | | | | |
|----|-------------------------|---------|------------------------|
| a. | Action – Agent | | |
| | Verb | er | (<i>driver</i>) |
| b. | Instrument – Agent | | |
| | Noun | (s) man | (<i>oarsman</i>) |
| c. | Object – Action – Agent | | |
| | Noun | Verb | er |
| | | | (<i>wood-cutter</i>) |

From this it follows that the WFR is constituted by the unity of the onomasiological and onomatological structures.

2.3.5. Justification

The reason for preferring this approach to the calculation of Productivity Rate is that

- (a) it makes it possible to examine productivity from different viewpoints reflecting both linguistic and supralinguistic levels;
- (b) it takes into consideration all new words (not just some WF processes like affixation);
- (c) it restricts itself to actual words (i.e. words coined in response to the needs of a particular speech community) in order to avoid the one-sided formalism of the mainstream discussion on word-formation.

From the previous discussion it follows that productivity is conceived as an **implemented capacity reflecting the naming needs of a particular speech community**. As suggested in Štekauer (1998, 2001), what seems to be crucial is that by coining a new word in response to the specific demand of a speech community the particular language manifests its productive capacity to provide a new, well-formed linguistic sign by employing its productive types/rules whenever need arises. By implication, inclusion in the model of the

extra-linguistic factor (speech community) makes it possible to eliminate the notion of overgeneration.⁶

This approach is in accordance with Bauer who maintains that “[t]he fact remains ... that the production of new words may be the only evidence the observer has of this potential, and the lack of new words appears to deny the potential” (2001: 21) and that “...words are only formed as and when there is a need for them, and such a need cannot be reduced to formal terms” (2001: 143). In principle, the conception of productivity as implemented capacity corresponds with Bauer’s (2001) notion of ‘profitability’.

Obviously, the proposed model of computing the productivity takes **dictionaries** as its basic source of data. This is not viewed as its drawback. It is believed that the method can be advantageously applied to the determination of productivity in selected lexical fields (sciences, sports, culture, etc.) as captured – generally fairly well – in a number of special-purpose dictionaries. It can also be applied to identify the latest trends in coining new naming units thanks to the dictionaries of neologisms and/or lists of new words as published, for example, in *American Speech*. Since productivity changes are not the matter of weeks, months, nay even one or two years, the time lag of covering these trends by dictionaries does not seem to be a relevant objection against this method. Moreover, it may be proposed that studying the general productivity should be subordinated to the determination of the productivity in the individual spheres of life as captured by special-purpose dictionaries. Namely, it may be postulated that the situation and the trends in coining new words in the fields like, for example, medical research versus fashion pursue different trajectories. From this it follows that any generalizations based on unequal amount, structure, and range of data may be fairly misleading.

The model proposed can be illustrated by the results of a case study focussed on the names of INSTRUMENTS (including tools, devices, machines, equipment, appliances, implements, apparatus, etc.) in the *English-Slovak Technical Dictionary* by A. Caforio (1996) under the arbitrarily selected letter “S”. The analysis of 192 naming units indicates that – out of the five Onomasiological Types – the most productive is Onomasiological Type 3, with over 55% Productivity Rate, followed by Onomasiological Type 1 with 28% PR, Onomasiological Type 2 with 12.5% PR, and conversion (almost 5% PR). From this it follows that there is a very strong tendency to morphematic representation of the Actional semantic component of the onomasiological structure (over 80% of all naming units).

At the level of WF Types, the most productive is the [**Action^{Purpose}Instrument**] type with more than 55% PR, followed by [**Object–Action^{Purpose}Instrument**] with 15% PR. The limited sample indicates the tendency for Instrumental names to leave the determining constituent of the OM unexpressed, thus producing less specialized terms to the benefit of higher-level generalizations, and – by implication – broader applicability of the instrumental naming units.

At the level of Morphological Types, the [**Action^{Purpose}Instrument**] type, for example, is dominated by the [**stem + -er/-or**] MT (e.g. *sensor*, *slipper*, *selector*) the productivity of which amounts to almost 72 %. The remainder is represented by the [**stem – stem**] MT with over 25 % productivity (e.g. *suction funnel*, *search coil*, *summation instrument*), and **conversion** (e.g. *slide*, *rule*). Again, important conclusions can also be drawn at this lowest

⁶ Which means that ‘our’ word-formation component (unlike, for example Halle’s (1973) and Allen’s (1978) does not ‘generate’ possible, but ‘non-existing’ words, i.e., it does no more than is actually needed by a speech community.

productivity level, i.e., the most frequently employed Morphological Type for Instrumental names is one with the *-er/-or* suffix.

This outline indicates that the proposed model makes it possible to draw relevant conclusions by interrelating all word-formation processes at various levels of generalization, depending on the specific needs of analysis.

2.3.6. Word-Formation as Creativity within Productivity Constraints

The terms ‘creativity’ and ‘productivity’ are usually understood as mutually excluding principles in coining new words. While productivity is said to be rule-governed, creativity is conceived of as any deviation from the productive rules. In the present context, creativity is used in a different meaning in which it is complementary with productivity. First, the logical spectrum (conceptual level) does not necessarily lead to one single Onomasiological Structure. For illustration, if we try to form a naming unit for ‘a person who meets space aliens on behalf of the human race’ the logical spectrum may yield various word formation types, such as Theme – Action – Agent, Location/Theme– Action – Agent, Location – Action – Agent, Object/Location – Action – Agent, Object – Action – Agent. Second, these different Word-Formation Types may be assigned various morphological realizations by the MSAP principle, for example,

- | | | | |
|------|----|---|--|
| (22) | a. | Theme – Action – Agent
<i>human race representative</i>
<i>homosapience representative</i> | (Onomasiological Type1)
(Onomasiological Type 1) |
| | b. | Location/Theme – Action – Agent
<i>earth-representative</i>
<i>earth ambassador</i>
<i>world ambassador</i> | (Onomasiological Type 1)
(Onomasiological Type 2)
(Onomasiological Type 2) |
| | c. | Location – Action – Agent
<i>intergalactic diplomat</i>
<i>interstellar diplomat</i> | (Onomasiological Type 2)
(Onomasiological Type 2) |
| | e. | Object/Location – Action – Agent
<i>extra-terrestrial greeter</i>
<i>space alien meeter</i>
<i>outerspace wellcomist</i> | (Onomasiological Type 1)
(Onomasiological Type 1)
(Onomasiological Type 1) |
| | f. | Object – Action – Agent
<i>contactee</i>
<i>greeter</i> | (Onomasiological Type 3)
(Onomasiological Type 3) ⁷ |

Example (22) thus illustrates what can be labeled as **creativity within productivity constraints**. It illustrates, on the one hand, different onomasiological realizations of a particular logical spectrum, and, on the other hand, different onomatological realizations of various onomasiological structures. It is the interaction between the conceptual, onomasiological, and onomatological levels which – within the limits of productive types and rules and the relevant constraints – provides certain space for a **creative** approach to word-formation (as it follows from several options in our example). This meaning of creativity emerges from a cognitive onomasiological approach. The inclusion of speech community in the model and viewing each new naming unit as a result of a very specific and real act of naming by a coiner makes it possible to reflect in the present model

⁷ The examples in (22) were proposed by Native speakers.

individual preferences, the influence of one's age, education, and profession, as well as one's linguistic family background (in a bilingual setting), fashionable trends, etc., i.e., the sociolinguistic factors which may affect the application of the MSAP in those cases that provide more than one option. Thus, it is in this sense of 'creativity within productivity constraints' that the presented onomasiological approach treats word-formation, and in particular, the relation between productivity and creativity. This brings us to an experimental research aimed at the application of the 'multilevel' computation of productivity and at demonstrating the validity of the concept of word-formation as creativity within productivity constraints. For that purpose, we will present sociolinguistically oriented evidence.

3. Sociolinguistic Research into WF Productivity

3.1. General

It is generally accepted that word-formation processes are never totally unrestricted, and even the most productive affixes seem to be subject to certain structural constraints (Plag 1999: 35). In the literature on word-formation, a number of restrictions upon productivity were mentioned. In addition to the 'traditionally' adduced systematic constraints,⁸ including phonological, morphological, lexical, syntactic, and semantic ones, both Bauer (2001) and Plag (1999) list some other, extra-linguistic factors, including

- (a) **Pragmatics** (because of denotation and connotation of some WF patterns they are not in common use, e.g. suffix *-some* in words like *twosome*. In principle, *twenty-five-some* may be possible, but it is not usual because we do not usually operate with groups of 25 people);⁹
- (b) **Aesthetics**, e.g., word-length;
- (c) **'Accidents of cultural history'**, e.g., a person whose job is to sell things happens not to be called *seller* since *salesman/saleswoman* is the established form. In these cases there is no linguistic reason for the current usage, it just so happens that a particular possible form has not become part of the norm;¹⁰
- (d) **Failure of hypostatization**: Coining a new word presupposes that there is an entity to be denoted by the new word. If there is no such entity, there is no need for a word.¹¹

It appears, however, that in spite of abundant literature on productivity constraints, there is at least one factor that has been neglected and that deserves attention of morphologists, in particular, the sociolinguistic factor. We believe that productivity of Onomasiological Types, Word-Formation Rules/Types, and Morphological Types is also affected by sociolinguistic factors which may be divided into two groups:

⁸ For a comprehensive review of various restrictions as well as blocking theories see Plag (1999), Bauer (2001), and Rainer (2005).

⁹ For a review of pragmatic factors (fashionability, demand, attitudinal function, hypostatization, nameability) see Plag (1999).

¹⁰ This is not to say that constraints on productivity of any type are absolute. In the case of 'accidents of cultural history', for example, the blocking principle can be 'overpowered' by a particular Word-Formation Type gaining in productivity (for any reasons, including, inter alia, those concerning voguish use).

¹¹ That the situation can change can be illustrated by **loather*. Bolinger (1975: 109) notes that this word is not an actual word of English not because it cannot be formed, but because "we have no use for it. What retinue of people would it designate?". Bolinger's view is also referred to by L. Bauer (2001: 43). However, Ingo Plag, (personal communication) drew our attention to "numerous nice attestations of this word on the internet (two even in dictionaries)."

- (i) **Horizontal factors**, including the previous linguistic experience. This factor plays its role in multinational countries, such as the USA, Australia, Great Britain, and in fact, a number of other countries due to the growing migration. There are millions of people whose grandparents, parents, or they themselves were born and have lived in a linguistically different environment.
Interestingly, while the factor of linguistic interference has been a topic of many treatises focussed on grammar, pronunciation, etc., the issues of interference in word-formation has not been – to our knowledge – studied yet.
- (ii) **Vertical factors**, including various social strata, education levels, professions, etc. It goes without saying that these factors affect the extent of actively and passively mastered vocabulary of a speaker, and hence influence his/her linguistic behaviour, which cannot – in our view - remain without effects upon the formation of new naming units.

Given these postulates, it may be proposed that any act of word-formation is a kind of intersection of three factors:

- (i) the pressure of the productivity of individual Onomasiological/Word-Formation/Morphological Types and Word-Formation Rules within the respective conceptual-semantic clusters;
- (ii) the extent of experience (including no experience) with a native language other than English;
- (iii) vertical sociolinguistic factors.

As it follows from experimental data, the latter two factors have their say at the onomasiological and the onomatological levels of the word-formation model (Figure 1), that is, at the level of conceptually identified logical-semantic structure establishing the basis for the act of naming, and at the level of its linguistic expression (assignment of WF bases and affixes to semes). It is these two levels that provide – as we believe – sufficient space for the operation of extralinguistic factors. In other words, it is at these levels that one's naming preferences may be implemented as the above-mentioned sociolinguistic factors may affect a coiner's selection (influenced by his/her former mother language word-formation patterns, education, extent of his active vocabulary, the register used in his/her social stratum and occupation, etc.) of one or the other affixation type, a verbal compound type, a non-verbal (primary) compound type, a conversion, blending (to use traditional terminology), etc.

3.2. Experimental research

3.2.1. General

Our experimental research was aimed at identifying the validity of our hypothesis concerning the influence of sociolinguistic factors upon the productivity in word-formation, in particular, the role played by linguistic background, education, and profession.

For the sake of our experiment, a questionnaire was developed (see Appendix 1). The basic task of the informants was to give names to 'objects' for which there did not exist any corresponding names in English at the time of our experiment. To avoid inconsistency, all

the objects to be named were conceived of as Agents. Our decision to concentrate on Agents was motivated by a relatively large number of different rules that make it possible to coin Agent names.

To avoid any distortion of results due to one-sided/inappropriate formulation of the experimental task, we decided to provide our informants with a questionnaire consisting of two basic parts, with the first part including three differently formulated naming tasks. The first was a selection task. Each object to be named was briefly characterized, e.g., ‘a person who frequently interrupts other people when they are talking’. The characterization of the object of naming was followed by a set of options. In this particular case, they included *interrupter*, *interruptist*, *butt-in*, *butter-inner*, *cutter-in*, *cutman*, *interposer*, and a few others. In addition, the final option in each set was a blank line which could be filled in if an informant did not find any of the options offered to be a suitable way of naming the object.

Task 2 differed from Task 1 in not containing any options. The informants had to propose their own naming units based on a brief specification of the object to be named, for instance, ‘Suppose that space aliens were about to land on Earth for the first time. What would you call a person who was supposed to meet them as a representative of the human race?’

Task 3 replaced wording by a drawing of a situation in which an object performs some unusual activity, for example:

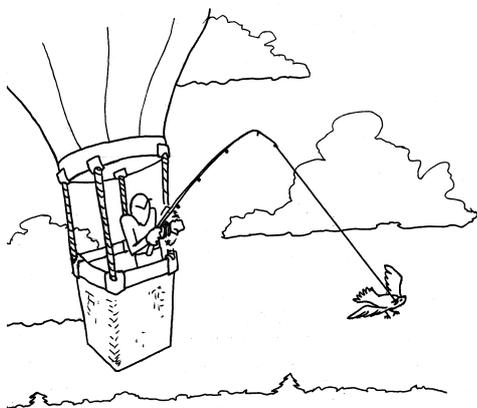


Figure 2

Part 2 was of a different nature. The purpose was to identify any possible differences in linguistic (naming) behaviour of various groups of language users with respect to unproductively coined naming units, i.e., naming units which were formed in defiance of relevant productivity constraints. For this purpose, five naming units were formed: *engroupment*, *thinnen*, *swimmee*, *sleepable*, and *satisfactority*. These naming units violate different productivity constraints.

The views of the suffix *-ment*, by means of which *engroupment* was formed, differ. While Bauer (1983: 49) maintains that this suffix does not seem to be productive any more, others like Plag (1999: 72-75) and Adams (2001: 28) demonstrate that *-ment* is low productive. It follows from Plag's analysis that the best candidates for *-ment* derivation are verbs ending in the suffix *-en*, having a disyllabic bases with stress on the second syllable, and with a

prefix like *be-*, *en-*. As Plag notes, “[t]he preference for prefixed stems is especially obvious with stems containing the prefixes *eN-* and *be-*, which seem to take *-ment* obligatorily” (1999: 73). Thus, *engroupment* is possible as it meets the specified restrictions. On the other hand, the specific WF pattern is low productive. In addition, *engroupment* violates the ‘Avoid Synonymy Principle’ (Kiparsky 1982) because its place in the system has already been filled in with *group* and *grouping*.

For *thinnen* the constraint is different: the inchoative suffix *-en* only attaches to monosyllabic stems if and only if they end in an obstruent, optionally preceded by a sonorant (Halle 1973). Furthermore, the suffix *-en* does not seem to be productive any more (Lieber 2004: 76).

The selection of *swimmee* based on the $[V+ -ee]_N$ pattern requires a more detailed explanation. As suggested by Barker (1998: 708), the suffix *-ee* can be viewed as a counterpart of *-er*, and “it is possible to entertain the hypothesis that the conditions for use of *-ee* are defined negatively, in contrast to those for *-er*: *-er* picks out subject participants, and *-ee* covers everything else.” As he, however, notes this hypothesis faces the problem of the existence of a considerable number of *-ee* nouns referring to subject participants. This is confirmed by Lieber (2004, 2005), who points out that while *-er* nouns “most often form personal agent nouns, and *-ee* most often forms patient/theme nouns, not infrequently we find precisely the opposite situation, where *-er* and its cohort form patient nouns and *-ee* agent or at least subject-oriented nouns” (2005: 404).

Moreover there are instances of both *-ee* and *-er* attached to the same WF base having synonymous meaning (*escapee/escaper*, *absentee/absenter*, *arrivee/arriver*, etc.) (Barker 1998: 709). Based on the analysis of a large corpus, Barker arrives at a conclusion that there are at least three types of *-ee* derivations that are productive: direct object, indirect object, and subject. Our *swimmee* is the subject type. This possible naming unit safely meets two of three of Barker’s (1998) semantic conditions imposed on productive *-ee* derivation. First, it meets the condition of sentience¹² of the referent referred to by *swimmee*. Second, it meets the semantic constraint of ‘episodic linking’, according to which “the referent of a noun phrase headed by an *-ee* noun must have participated in an event of the type corresponding to the stem verb” (1998: 711). In this particular case, *swimmee* participates in a swimming event. Problematic is the third semantic constraint, defined as “a lack of volitional control on the part of its referent either over the occurrence or the duration of the qualifying event itself or (given a punctual qualifying event) over its immediate direct consequences” (1998: 717). Being a subject type *swimmee* refers to Agent, but the category of Agency implies volition. Thus, for a *swimmee* to preserve the ‘lack-of-volition’ constraint it would have to mean the action into which a swimming person is forced somehow – in contrast to *swimmer*, who, in principle, does his/her activity voluntarily, fully based on his will.¹³

It follows from the above discussion that *swimmee* is a possible naming unit that can be produced by a productive WF rule. What made us include this word in the ‘unproductivity’ test is the much more productive competitor, the *-er*-based pattern that underlies the existing and well-established Agent noun *swimmer*. Thus, while the *-ee* Agent noun is possible (other meanings of *-ee* nouns, such as Patient or Theme can hardly be expected as was also manifested by our subsequent experimental research) it is blocked on a general

¹² Which means reference to an Animate entity.

¹³ Even this is not quite so: a professional *swimmer* training is a hard drill under the control of a coach, and not always in accordance with the will of the *swimmer*.

Agentive level by a much more productive WF rule that has already produced a firmly established (institutionalized) naming unit *swimmer*. On a fine-grained semantic level, the blocking is eliminated by the ‘volition – lack of volition’ opposition. The question behind the inclusion in the experimental ‘unproductivity’ research of *swimmer* was whether the informants (native speakers) will perceive this kind of semantic distinction.

As it will follow from the results of our experiment, while almost each of our native speaker informants was able to propose a sentence in which they used *swimmee*, the vast majority of them find *swimmee* to be ‘extremely unlikely’ or ‘somewhat unlikely’ just because the uses proposed were in the absolute majority of cases connected with the Agent-based interpretation that did not distinguish the ‘volitional’ constraint. In fact, **none** of our native speaker informants referred to the volitional aspect of the swimming action. As such the meaning of *swimmee* was to the vast majority of the informants blocked by *swimmer*.¹⁴

The suffix *-able*, which occurs in our naming unit *sleepable*, does not meet the traditionally adduced restrictions, summarised in Anderson (1992: 186):

- (23) WFR: $[X]_V \rightarrow [X\text{əbl}]_{\text{Adj}}$
 Condition: $[X]_V$ is transitive (i.e., $[+ _NP]$)
 Syntax: ‘Object’ argument of $[X]_V$ corresponds to ‘Subject’ of $[X\text{əbl}]_{\text{Adj}}$
 Semantics: ‘(VERB)’ \rightarrow ‘capable of being VERBed’

Sleepable is intransitive and there is hardly any acceptable reading that would meet the syntactic condition. From the semantic point of view, it rather features a ‘property meaning’, to use Plag’s (2004) term. Importantly, as noted by Plag (ibid), “the forms exhibiting the property meaning are in a clear minority. In fact, this pattern has ceased to be productive as early as the 17th century...” All these facts imply that *sleepable* is a good candidate for our unproductivity test.

Finally, *satisfactority* violates the constraint according to which *-ity* is only productive (is potentiated – Williams 1981) in combination with the productive *-able* function, which maps transitive verbs to adjectives. By implication, the domain of the *-ity* function f_{ity} is the function f_{able} , and its range is the composed function $f_{ability}$ (Raffelsiefen 1992).

This task was thus aimed at recognizing the ‘sensitivity’ of different groups of speakers to productivity constraints and the ‘inappropriately’ coined words. The informants were given a five-degree scale, including the options of ‘extremely unlikely’, ‘somewhat unlikely’, ‘likely’, ‘very likely’, and ‘extremely likely’. In addition, they were asked to give an example of a sentence, including an ‘unproductively’ coined word.

The questionnaires were collected (and the informants were approached) in various ways, in particular through personal contacts, through our students and friends in English-speaking countries, through the Internet *LinguistList* service, and finally, through a special-purpose www.page. It follows that it was fairly difficult to meet the initial goal of having the individual subgroups per profession, occupation, and different linguistic background evenly distributed. In any case, we believe that the extent of the sample made it possible to

¹⁴ Instead of the volition-related constraint, proposed by Barker, some of the informants distinguished between *swimmee* and *swimmer* in terms of ‘swimming skill’, mostly in favour of *swimmer*. The unequal skill, ability, and capacity *per se* (in any activity) do not, however, seem to be a sufficient justification for a productive WF process. This would lead to an extremely high number of naming units and a considerable overload of a language user’s memory.

accomplish the basic objectives of our experimental research and to draw relevant conclusions. The sample of informants was divided into two groups, the native speakers whose parents were born in an English-speaking country (language proficiency A in the questionnaire), i.e., those who were not influenced at home by immediate contact with a different language; and native speakers whose parents were not born in an English-speaking country plus non-native speakers living in an English-speaking country (language proficiency B and less in the questionnaire), i.e., those whose English competence had to cope with the influence of another language. The former group (speakers unconnected with another native language) has been subdivided accordingly into various subgroups by occupation (students, educators, ‘other’ professions), and by education (high school, college, graduate). The latter group (speakers connected with another language) has been divided into groups based on the morphology of noun, namely, synthetic/agglutinative, synthetic/fusional, analytic/isolating, and polysynthetic. Due to very low numbers of informants (three), the polysynthetic group was not taken into consideration.

3.2.2. General Analysis

3.2.2.1. Native Speakers

The experiment encompassed 145 native speakers from various English-speaking countries, mostly from the USA. The total number of ‘responses’ amounts to 4 tasks times 5 subtasks per each, which gives 20 responses per informant, which, ideally, adds up to 2,900 ‘responses’ in total. However, the actual number of responses is smaller (1,531) for two major reasons:

1. Not all of the informants completed all sub-tasks.
2. Some informants did not specify all relevant data within the demographic information section of the questionnaire, and therefore their replies could not be taken into account in all parts of our analysis;

The number of 1531 responses was further reduced down to 1300 relevant responses that became an object of our analysis. The difference of 231 responses that were eventually eliminated from consideration follows from the fact that our research was focused on the productivity in word-formation, and therefore all those naming units proposed by the informants which resulted from sources other than productive word-formation were eliminated from the scope of analysis:

- (a) They were mostly proposals based on **semantic shift** of an already existing word – in which case no new naming unit comes into existence.
- (b) In addition, only those naming units were taken into account that indicated the meaning specified by the descriptive wording or drawing. Therefore, we disregarded proposals like *ambassador*, *welcomer*, *ET*, and *President of the USA* for ‘a person meeting space visitors’; *risk-taker*, *show-off*, *crazy*, *mad*, *retard*, *daredevil*, *weirdie* for ‘a person riding on car-body top’; *comedian*, *platinum record*, *idiot*, *cruel*, and *joker* for ‘a person frequently joking about blondes’, *weird*, to denote ‘a person who dials a telephone number with a feather’, *zoologist* for ‘someone who does research about spider webs’; *perfectionist*, *fussy*, *meticulous gardener*, *biologist*, and *frowny face* for ‘a person cutting grass with a knife’; *show-off*, *Michael Jordan* and *monkey* for ‘a basketball player who always hangs onto the rim after a slam-dunk’, *macho* and *acrobat* for ‘a person lifting weights on a crane’; *time-killer* for ‘a person tying shoelaces to customers’, etc. They are either ‘mere’

extensions of the original meanings, or are too general to say anything relevant about the actual mission/activity of the individual objects to be named.

- (c) We also eliminated those naming units that do not correspond with the productive WF types in English or are ungrammatical in any other way, for example, *arachologue*, *slam duckle*, *pedlacier*, *person flying over car*, *believer in miracles*, *researcher on spider webs*, etc.
- (d) Finally, the following analysis does not take into consideration the names of Patients¹⁵ that also occurred in the experiment. This leaves us with 1,300 responses.

Importantly, since the focus of our research is on WF productivity rather than on individual naming units the following analysis concentrates on types and rules. Brief comments on some interesting cases of individual naming units are given in 3.2.6.

3.2.2.2. Productivity of Onomasiological Types

As already suggested above, there are two contradictory tendencies in language, the tendency to the **economy of expression** and the **tendency to the explicitness of expression** (clarity of communication). If we analyse the results in view of scale (i) specified in Section 2.2.4.1, that is to say, in view of the complexity of onomasiological structure, we find out that the total number of responses for the onomasiologically ‘explicit’ types 1, 2 and 3 is 1,272 and that for the ‘non-explicit’ type 5 is 28 (Table 1). No naming units were based on Type 4. From the point of view of scale (ii), in particular, the explicitness of the onomatological level, it may be concluded that the number of explicit types (Type 1) roughly corresponds with the number of ‘economic’ types (51.54% : 48.46%).

Moreover, the central role is played by those Onomasiological Types (Types 1 and 2) whose determined constituent (i.e., the Action-representing constituent) is explicitly represented by a morpheme. In total, they represent 75% of all naming units. This result is not surprising because it is this constituent that is vital to the understanding of new naming units. The Actional constituent namely relates the onomasiological base with the determining constituent of the mark in Onomasiological Type 1 thus significantly contributing to the interpretability of such naming units. Also in type 2, the determined constituent clearly indicates the ‘Action’ of the Agent represented by the onomasiological base. Thus, for example, the determined constituent of the mark (*surf*) in *roof-surfer* clearly and unambiguously identifies the relation between the polar members of the onomasiological structure, i.e., *roof* and *-er*, and makes the interpretation of this naming unit easy. On the other hand, a naming unit falling within the scope of type 3, i.e., *roofer*, makes the process of meaning interpretation pretty demanding just because there is no Actional constituent that would appropriately identify the relation between the polar members of the onomasiological structure represented by the morphemes *roof* and *-er*. As a result, the number of possible interpretations of *roofer* is considerably high.¹⁶

The Onomasiological Type 2 variant of the same ‘object’, i.e., *surfer* is more valuable in terms of meaning predictability than the type 3 variant because it identifies the actual Action of the Agent.

¹⁵ Patient is here defined as ‘Bearer of State’. Examples from our experiment include *car-topped guy*, *obsessionist*, *clone*, etc.

¹⁶ For a theory of meaning predictability of naming units coming into existence by word-formation processes see Štekauer (2005a).

There is no Agentive naming unit of Type 4. In general, however, this Onomasiological Type is highly valuable in terms of easy interpretation because the direct connection between the unstructured mark and the base at the onomasiological level gives no chances for multiple interpretations. This can be exemplified by a Patient name that occurred in our research, *sub-clone*, where the mark *sub-* directly specifies the Quality of Patient *clone*.

Given our results, however, type 4 does not seem to be a productive type for Agent names for the simple reason that it usually specifies the Quality rather the Action performed by Agent.

	No. of responses	PR (%)
OT1	670	51.54
OT2	299	23.00
OT3	303	23.31
OT4	0	0.00
OT5	28	2.15
Total	1300	

Table 1: Predictability Rate of Onomasiological Types (native speakers)

3.2.2.3. Productivity of Word-Formation Types

Since the experiment examined the naming preferences of English speakers in the field Agents (1300 responses), i.e., persons performing some Action, the dominant position of WFTs [**Object–Action–Agent**], [**Action–Agent**] and [**Theme–Action–Agent**] is not surprising: since Agents are human beings performing some Action, the presence of the determined constituent of the onomasiological mark (which, as we already know, stands for Action in general) is expected.

The most productive types in our research indicate two basic tendencies in the naming ‘behaviour’ of native language users. First, they select a more general naming unit because they either wish to increase its extension, for example, to avoid the exclusion of some unpredictable special-purpose cases, or, because the scope of Agent’s Action is vaguely defined. In our experiment, one such reason which contributes to the productivity of the [**Action–Agent**] WFT to the detriment of a more explicit type was the fact that some of our informants were not quite sure about the specific nature of the broadly conceived activity to be named – for example, the ‘grass-cutting’ and the ‘shoe-lacing’ drawings (cf. Appendix 1).

Second, in the majority of cases, there is an effort of native language users to be more specific (circumstances-permitting) and express those categories which are inherently related to Action, such as Object of Action, Instrument of Action, Theme of Action, Location of Action, Time of Action, etc. Consequently, there is no wonder that the [**Object–Action–Agent**] WFT is the most productive in our sample, with 416 responses, yielding the Productivity Rate of 32.0%. The PR of the [**Theme–Action–Agent**] WFT is 17.3%, [**Instrument–Action–Agent**] 10.4%, and [**Location–Action–Agent**] 9.1%. The PR of the above commented, more general WFT [**Action–Agent**] is 23.4%. These five WFTs represent about 92% of all Agentive naming units in the native-speaker group of informants, which clearly indicates their high productivity, on the one hand, and a minor role played by the remaining WFTs.

	PR (%)
Object–Action–Agent	32.00
Action–Agent	23.38
Theme–Action–Agent	17.31
Instrument–Action–Agent	10.38
Location–Action–Agent	9.08

Table 2: Predictability Rate of Word-Formation Types (native speakers)

3.2.2.4. Productivity of Morphological Types¹⁷

The number of options for Morphological Types is not large as it is limited by the combinability of stems and affixes, subdivided (in English) into prefixes and suffixes. Since Agentive functions are primarily expressed by suffixes in English, those Morphological Types are more productive which combine stems (S) with suffixes.¹⁸ While the most productive Morphological Type **[S+S+suffix]** (PR=47%) corresponds with the expectations stipulated in 3.2.2.1.2 above, i.e., that a more explicit structure is preferred, the distribution of MTs in terms of implicit and explicit structures is roughly balanced, with the two most productive MTs being two-constituent structures **[S+suffix]** (37%) and **[S+S]** (11%), respectively. On the other hand, the MTs with a suffix in the role of onomasiological base clearly prevail with 85% PR.

	(PR%)
S + S + suffix	46.67
S + suffix	37.41
S + S	11.29

Table 3: Predictability Rate of Morphological Types (native speakers)

3.2.2.5. Productivity of Word-Formation Rules

Word-Formation Rules result from the operation of the Morpheme-to-Seme-Assignment Principle, which means that they reflect the interrelation between the onomasiological and the onomatological levels. As such, they should reflect the basic tendencies in the domain of Word-Formation Types (onomasiological level) and Morphological Types (onomatological level). Therefore, since the most productive WFT is **[Object–Action–Agent]**, since the most productive MT is **[S+S+suffix]**, since Objects of Action are usually expressed by nouns, and, finally, since it is generally known that the *-er* suffix is the most productive Agentive suffix in English (much more productive than its competitors, like *-ist*, *-ant*, *-ee*, *-ian*) one may expect the dominating position of the following WFR

$$(24) \quad \begin{array}{ccc} \text{Object–Action–Agent} & & \\ \text{N} & \text{V} & \text{-er} \end{array}$$

And actually, the results bear out this postulate, as it follows from Table 4. The prominent position of the ‘*-er-for-Agent*’ structures is strengthened by the fact that the four most productive WFRs are of this sort, with their share of the total number of Agentive

¹⁷ The cases of OT5 are not included for obvious reasons. Thus, the number of responses taken into consideration for Agents is 1 275, those for Patients is 165.

¹⁸ The naming units with *man* in the position of onomasiological base are classified as stem-based units in spite of the fact that a number of authors treat this element as semiaffix.

WFRs exceeding 50%. In addition, it can be concluded from Table 4 that WFRs with stems in the function of an onomasiological base are far from being high-productive in English.

	PR (%)
Object–Action–Agent (N V -er)	20.92
Action–Agent (V -er)	17.85
Theme–Action–Agent (N V -er)	10.62
Location–Action–Agent (N V -er)	7.08
Object–Action–Agent (N 0 -ist)	4.38
Instrument–Action–Agent (N 0 -ist)	4.23
Action–Agent (V N)	2.69
Instrument–Action–Agent (N V -er)	2.23

Table 4: Productivity Rate of Word-Formation Rules (native speakers) – Agents

3.2.3. Influence of Occupation

3.2.3.1. Analysis of the Experimental Data

Taking the general picture, discussed in Section 3.2.2, as a reference point, we can proceed to the comparison of the data obtained for the individual groups of informants, based on their occupation. The available sample of informants necessitated their division into three groups, in particular, students, teachers, and ‘other’ professions. The sample included 60 students, 35 educators, and 50 ‘other’ professions who produced 1531 (1300 for Agents and 231 for Patients) responses in total. The latter group of occupations was originally subdivided into those of civil servants, natural scientists and engineers, managers, manual workers, and medical doctors, but the data of all these sub-groups had to be cumulated into one because of insufficient number of questionnaires per subgroups. As a result, we obtained three basic occupational groups of comparable sample size.

3.2.3.1.1. Onomasiological Types

The data offered in Table 5 indicate that there are differences between the groups of students and teachers on one hand, and the ‘other’ professions, on the other. They mainly concern Onomasiological Type 1 where the respective Productivity Rates are 53.86% and 57.30%, for the first two informant groups, and much lower in the ‘other’ group (43.25%). This is, naturally, projected onto the situation in Onomasiological Types 2 and 3 where the PRs in the ‘other’ professions are the highest of all. Since the main difference seems to be between those who are in education professions and those who are not, the students and educators have been grouped together in the statistical tests. (OT4 has been left out of the chi-square test because it was zero for all groups.)

The tendency emerging from the data outlined is that while the language speakers belonging in the education-oriented professions, including education-related major activity (study), tend to form more comprehensive naming units, aimed at maximum explicitness and accuracy of ‘labeling’ the objects of naming, the speakers belonging in the ‘other’ professions prefer morphematically reduced ways of expression (economy of expression) (Type 2), and/or vaguer naming units with broader extension, the meaning of which is more difficult to predict (Type 3). Thus, in this particular case, we witness a different treatment by the representatives of different groups of professions of the above-indicated conflict between the explicitness of expression and the economy of expression.

	Education Professions	Non-education Professions
No. of responses	900	400
OT1	497 (55.2%)	173 (43.3%)
OT2	193 (21.4%)	106 (26.5%)
OT3	193 (21.4%)	110 (27.5%)
OT4	0	0 (0%)
OT5	17 (1.9%)	11 (2.8%)

Chi-square = 16.089 p = .001 df = 3

Table 5: PR (%) of Onomasiological Types by occupation (native speakers)

3.2.3.1.2. Word-Formation Types

The five most productive WFTs follow the tendencies from the general discussion of Onomasiological Types. Also here, the PR values of more explicit Word-Formation Types in the education-oriented groups are higher than those in the non-education group, even if the differences are not significant by a chi-square test, and are distributed among the individual WFTs to give the indicated cumulative effect – the PRs of three-constituent WFTs are generally higher in the education-oriented occupations than in the third group of informants. The tendency observed for Onomasiological Types gets the most persuasive support from the **[Action-Agent]** type whose PR in the ‘non-education group’ (25.75%) clearly outscores those for the other two groups (22.61% and 21.35%, respectively).

The largest number of different WFTs has been found in the group of students (16). The other two groups proposed naming units belonging in 14 different WFTs. This high number of different WFTs used and the differences between the individual occupational groups provide unequivocal evidence of the validity of the concept of word-formation as ‘creativity within productivity constraints’.

	Education Professions	Non-education professions
Object–Action–Agent	283 (34.4%)	130 (32.5%)
Action–Agent	199 (24.2%)	103 (25.8%)
Theme–Action–Agent	158 (19.2%)	67 (16.8%)
Instrument–Action–Agent	97 (11.8%)	38 (9.5%)
Location–Action–Agent	86 (10.5%)	32 (8.0%)

chi-square = 2.907 p = 0.5735 df = 4

Table 6: PR (%) of the top five Word-Formation Types by occupations (native speakers)

3.2.3.1.3. Morphological Types

The explicitness-economy conflict and its occupation-based solution at the level of Onomasiological Types is unambiguously acknowledged at the level of Morphological Types by similar PR differences: while a three-constituent structure **[S+S+suffix]**, in which all three constituents of the onomasiological structure are morphematically expressed, dominates the education-related groups (47.47% and 50.57%, respectively, versus 41.09% for the ‘other’ group), the highest PR in the ‘other’ professions is achieved by the **[S+suffix]** structure (42.12% versus 36.59% for students and 33.71% for educators), and

the PR of another relatively productive two-constituent structure, [S+S], is also higher in this group of informants (13.18% versus 11.07% for students and 10.29% for educators).

	Education professions	Non-education professions
S + S + suffix	438 (51.3%)	164 (41.1%)
S + suffix	319 (37.4%)	168 (42.1%)
S + S	97 (11.4%)	53 (13.2%)
Other	46 (5.1%)	15 (3.6%)

chi-square = 9.252 p = 0.0261 df = 3

Table 7: PR (%) of the most productive Morphological Types by occupations (native speakers)

3.2.3.1.4. Word-Formation Rules

The above-mentioned results are weakly supported by the data of the domain of Word-Formation Rules. In principle, they detail the general results obtained for Onomasiological Types, and therefore the results cannot differ significantly. By implication, the most productive WFR for education-related professions in the field examined is (25):

(25) Object–Action–Agent
 N V -er

i.e. an explicit, three constituent structure both at the onomasiological and the onomatological levels. For non–education professions, this WFR is surpassed by a two constituent WFR

(26) Action–Agent
 V -er

by one response. In the education fields, a slightly higher percentage of the WFRs are based on Onomasiological Type 1, while the WFRs without morphematic expression of the determining constituent of onomasiological structure play a more important role among the ‘other’ professions.

	Education Professions	Other professions
Object–Action–Agent (N V –er)	198 (37.5%)	74 (18.5%)
Action–Agent (V -er)	157 (29.7%)	75 (18.8%)
Theme–Action–Agent (N V –er)	106 (20.1%)	36 (9.0%)
Location–Action–Agent (N V -er)	67 (12.7%)	25 (6.25%)

Chi-square = 2.663 p = .4465 df = 3

Table 8 PR (%) of the most productive Word-Formation Rules by occupations (native speakers)

3.2.3.2. Summary

In summarizing the observations based on the experimental data, the following may be concluded:

There is an obvious tendency indicating different strategies in the naming acts in two different groups of language users. While the education-process-related English language users incline to those Onomasiological, Word-Formation, and Morphological Types and Word-Formation Rules that are more explicit, thus capturing the objects to be named in a more comprehensive way, the ‘other’ professions prefer brevity of expression, i.e., they favour economy of expression, simpler, more general, and therefore, less definite naming units. The first tendency is interpretation-friendly, because the meaning of a more explicit naming structure is more easily interpretable and predictable. The latter tendency favours the opposite universal feature of language, i.e., the effort for the maximum possible economy of speech to the detriment of clarity of expression.

3.2.3.3. Perception of ‘Unproductivity’

The data indicate that the perception of ‘unproductivity’ among native speakers in general is fairly strong. While almost all informants gave relevant examples of use of unproductive coinages in sentences they prevalingly reject these words as extremely unlikely. In particular, out of 708 responses, 397 (56.1%) fall within the ‘extremely unlikely’ class of answers, and 176 responses (24.9%) in the class of ‘somewhat unlikely’. Thus, the sample words are considered to be unlikely to over 80%. Yet, there are some differences among the individual naming units, with the greatest number of ‘likely-oriented’ responses being for *sleepable* – the only naming unit in this sample, for which there is more ‘somewhat unlikely’ votes than ‘extremely unlikely’ ones. In addition, the number of ‘likelys’ is fairly high. The great majority of the ‘likelys’ are connected with the meaning of ‘apt for sleeping’, mostly with the Location argument, in some cases also with the Temporal argument, for instance, ‘That bed looks very sleepable’, ‘This noise maked the room far from sleepable’, ‘The bears about to go to hibernation could be considered in a sleepable state’, etc.

The data indicate that the constraint, in particular, the subcategorization restriction permitting the suffix *-able* to combine with transitive verbs only, does not seem to be so strongly anchored in the minds of language users as the other restrictions covered in our experiment.

The differences among the individual occupation groups in terms of their respective perception of such naming units are not significant, with the exception of *sleepable*, in which case the ‘extremely unlikely’ votes are distributed with steps by about 10 per cent: 18.3% for students, 28.6% for educators, and 39.6% for other professions. With this naming unit, the percentage of ‘likely’ responses among the students is extraordinarily high – as much as 28.3%.

In general, the number of ‘very likely’ and especially ‘extremely likely’ responses approaches zero in the majority of cases, with the exceptions apparently being related to individual, idiosyncratic, usually stylistically motivated evaluation of a particular sample naming unit (as suggested by three of the informants who avoided classifying *swimmee* and *sleepable* as ‘extremely unlikely’ but emphasized that they could imagine the use of such words in ‘jocular’ context only).

The dominating prevalence of the ‘extremely unlikely’ and ‘somewhat unlikely’ responses suggests that the informants, irrespective of their occupation, have a strong awareness of ‘grammaticality’, hence of the relevant productivity constraints. On the other hand, the existence of a relatively high number of the ‘likely’ responses acknowledges their feeling

for a creative approach to their language. This mainly applies to the group of students who most readily accept unconventional naming units and break the existing rules. This does not seem to be a surprise, and might be accounted for psychologically by the dynamism of the young generation compared to the more conservative generation of their parents.

Legend: EU – extremely unlikely
 SU – somewhat unlikely
 L – likely
 VL – very likely
 EL – extremely likely

	Students	Educators	Other professions
<i>Engroupment</i>			
EU	31 (52.5%)	18 (52.9%)	33 (67.3%)
SU	19 (32.3%)	13 (38.2%)	11 (22.4%)
L	8 (13.6%)	1 (2.9%)	5 (10.2%)
VL	1 (1.7%)	1 (2.9%)	0 (0%)
EL	0 (0%)	1 (2.9%)	0 (0%)

Chi-square = 3.910 p = 0.4183 df = 4 (L, VL, and EL were combined)

<i>Thinnen</i>			
EU	30 (50.8%)	26 (76.5%)	35 (72.9%)
SU	20 (33.9%)	5 (14.7%)	5 (10.4%)
L	8 (13.6%)	1 (2.9%)	4 (8.3%)
VL	1 (1.7%)	2 (5.9%)	3 (6.3%)
EL	0 (0%)	0 (0%)	1 (2.1%)

Chi-square = 11.699 p = 0.0197 df = 4 (L, VL, and EL were combined)

<i>Swimmee</i>			
EU	32 (53.3%)	24 (70.6%)	35 (74.5%)
SU	18 (30.0%)	7 (20.6%)	6 (12.8%)
L	9 (15.0%)	2 (5.9%)	4 (8.5%)
VL	1 (1.7%)	1 (2.9%)	2 (4.3%)
EL	0 (0%)	0 (0%)	0 (0%)

Chi-square = 7.192 p = 0.1260 df = 4 (L, VL, and EL were combined)

<i>Sleepable</i>			
EU	11 (18.3%)	10 (28.6%)	19 (39.6%)
SU	17 (28.3%)	18 (51.4%)	13 (27.1%)
L	17 (28.3%)	5 (14.3%)	14 (29.2%)
VL	12 (20.0%)	2 (5.7%)	2 (4.2%)
EL	3 (5.0%)	0 (0%)	0 (0%)

Chi-square = 15.608 p = 0.0035 df = 4 (L, VL, and EL were combined)
 Chi-square = 21.629 p = 0.0014 df = 6 (VL and EL were combined)

Satisfactority

EU	35 (58.3%)	26 (78.8%)	32 (66.7%)
SU	11 (18.3%)	4 (12.1%)	9 (18.8%)
L	8 (13.3%)	2 (6.1%)	3 (6.3%)
VL	4 (6.7%)	1 (3.0%)	3 (6.3%)
EL	2 (3.3%)	0 (0%)	1 (2.1%)

Chi-square = 4.758 p = 0.3130 df = 4 (L, VL, and EL were combined)

Table 9: Perception of ‘unproductivity’ by native speakers¹⁹**3.2.4. Influence of Education**

In reference to various incorrect interpretations of compounds, such as *house-bird glass*, Gleitman/Gleitman (1970) relate their misinterpretation to the educational level of language users. Their informants fell within three different educational groups: (a) graduate students and PhD’s in various fields; (b) undergraduates and college graduates; and (c) secretaries with high school degrees. In many cases, their informants from the group of secretaries proposed various ‘unacceptable’ readings which corresponded to the compound *glass house-bird*, *glass bird-house*, or a paraphrase like *a house-bird made of glass* (in contrast to PhD informants who avoided such errors). The analysis of their research results made Gleitman & Gleitman conclude that there were “very large and consistent differences among these subjects of differing educational background” (1970: 117) and that “[t]he less educated groups make more errors, and to a significant extent make different errors than the most-educated group” (ibid. 128). While the research of the Gleitmans concerns the predictability of meaning, i.e., the interpreter’s pole rather than the coiner’s pole, it indicates that the level of education may play a role in word-formation, in general, and in the productivity of word-formation, in particular. No wonder, productivity of Word-Formation Rules appears to be one of the factors influencing the predictability of novel complex words (cf. Štekauer 2005).

In analyzing the questionnaires, our native speaker informants was divided into three groups, including those with high school, college, and graduate education. The total number of responses taken into consideration in evaluating the research data was 1,276 for the category ‘Agents’ nouns.

3.2.4.1. Analysis of the Experimental Data**3.2.4.1.1. Onomasiological Types**

The data for the Onomasiological Type 1 show a rising curve in the direction towards higher education level, though a chi-square test cannot establish significance. The Onomasiological Type 1 PR of the graduate group is noticeably higher than that of the high-school informants. This is compensated for by the higher PRs of the high-school and college informants for the other three Onomasiological Types. The highest PR in the Onomasiological Type 3 (one without the morphematic expression of the Actional constituent) is attributable to the informants with the lowest education-level, with the PR curve falling down towards the higher-educated speakers. The data indicate the preference of higher educated people for explicit way of expression, and the preference for more general way of expression in the lower educated language users.

¹⁹ Note: Not all informants provided answers to all individual tasks. Hence the numbers may differ.

	High school	College	Graduate
No. of responses	245	715	316
OT1	119 (48.6%)	358 (50.1)	176 (55.7%)
OT2	57 (23.3%)	171 (23.9)	69 (21.8%)
OT3	63 (25.7%)	169 (23.6)	68 (21.5%)
OT4	0 (0%)	0 (0%)	0 (0%)
OT5	6 (2.5%)	17 (2.37%)	3 (1.0%)

Chi-square = 5.655 p = 0.4629 df = 6 (OT4 was excluded from calculations)

Table 10: PR (%) of Onomasiological Types by education (native speakers)

3.2.4.1.2. Word-Formation Types

Given the previous data, the downward-leading PR curve in the direction towards higher education for the [**Action-Agent**] WFT was expected. As with the majority of other tables, the ranking of the individual WFTs is the same for all three groups of informants.

	High school	College	Graduate
No. of responses	245	715	316
Object–Action–Agent	77 (31.5%)	212 (29.7%)	102 (32.3%)
Action–Agent	70 (28.5%)	188 (26.3%)	71 (22.4%)
Theme–Action–Agent	47 (19.2%)	125 (17.6%)	49 (15.4%)
Instrument–Action–Agent	24 (9.6%)	75 (10.5%)	33 (10.6%)
Location–Action–Agent	16 (6.5%)	44 (8.1%)	33 (10.6%)
Other	11 (4.6%)	57 (7.9%)	28 (8.8%)

Chi-square = 13.378 p = 0.2032 df = 10

Table 11: PR (%) of top 5 Word-Formation Types by education (native speakers)

3.2.4.1.3. Morphological Types

The data for the [**S + S + Suffix**] structure in Table 12 acknowledge the growing importance of a more complex morphematic representation of complex onomasiological structure, i.e., more precise expression with the growing education of language users (even though the differences in PR between the high-school and college informants are minimal).

	High school	College	Graduate
No. of responses	245	715	316
S + S + suffix	107 (43.7%)	317 (44.4%)	151 (47.9%)
S + suffix	105 (42.9%)	290 (40.5%)	119 (37.8%)
S + S	24 (9.8%)	81 (11.2%)	25 (7.9%)

Chi-square = 8.690 p = 0.1917 df = 6

Table 12: PR (%) of the most productive Morphological Types by education (native speakers)

3.2.4.1.4. Word-Formation Rules

The level of Word-Formation Rules used for the coining of Agent names seemingly does not bear out the different naming strategies of the speakers of different education levels;

this bears on the data obtained for the top PR WFRs. It goes without saying that the different strategies revealed at the levels of Word-Formation Types and Morphological Types must find their mapping also in the domain of WFRs. The differences, however, are not so conspicuous, because they are scattered among the numerous low PR WFRs. An indicator of such low PR range differences is the last WFT given in Table 13, showing a PR gap of about 3% between the graduate speakers, on one hand, and the other two groups, on the other. The PR gap of 3% in the low predictability level range is striking.

	High school	College	Graduate
Number of responses	245	715	316
Object–Action–Agent			
N V -er	56 (22.7%)	145 (20.3%)	71 (22.4%)
Action–Agent			
V -er	50 (20.4%)	146 (20.4%)	63 (19.9%)
Theme–Action–Agent			
N V -er	29 (11.9%)	78 (10.8%)	27 (8.5%)
Location–Action–Agent			
N V -er	15 (6.2%)	44 (6.2%)	29 (9.1%)

Chi-square = 6.3676 p = 0.6061 df = 8

Table 13: PR (%) of most productive Word-Formation Rules by education

3.2.4.2. Summary

Tables 10 - 13 suggest, albeit weakly, that education seems to exert influence upon the approach to word-formation. There is a noticeable inclination of higher educated people to label objects of extra-linguistic reality as precisely as possible and, for this purpose, to employ more extensive naming structures. Lower educated informants demonstrated their preference for more ‘economic’ expressions to the detriment of clarity and precision of new naming units.

3.2.4.3. Perception of ‘Unproductivity’

	High school	College	Graduate
Number of responses	34	92	34
<i>Engroupment</i>			
EU	21 (61.8%)	52 (56.5%)	24 (70.6%)
SU	11 (32.4%)	28 (30.4%)	8 (23.5%)
L	2 (5.9%)	9 (9.8%)	2 (5.9%)
VL	0 (0%)	2 (2.2%)	0 (0%)
EL	0 (0%)	1 (1.1%)	0 (0%)

Chi-square = 3.358 p = 0.4998 df = 4 (L, VL, and EL were combined)

Thinnen

EU	19 (55.9%)	25 (63.0%)	28 (82.4%)
SU	10 (29.4%)	22 (23.9%)	3 (8.8%)
L	2 (5.9%)	12 (13.0%)	1 (2.9%)
VL	3 (8.8%)	0 (0%)	2 (5.9%)
EL	0 (0%)	0 (0%)	0 (0%)

Chi-square = 5.7480 p = 0.2187 df = 4 (L, VL, and EL were combined)

Swimmee

EU	21 (61.8%)	57 (62.0%)	26 (76.5%)
SU	6 (17.7%)	23 (25.0%)	5 (14.7%)
L	6 (17.7%)	9 (9.8%)	1 (2.9%)
VL	1 (2.9%)	2 (2.2%)	2 (5.9%)
EL	0 (0%)	1 (1.1%)	0 (0%)

Chi-square = 4.1656 p = 0.3840 df = 4 (L, VL, and EL were combined)

Sleepable

EU	9 (26.5%)	25 (27.2%)	15 (44.1%)
SU	10 (29.4%)	27 (29.4%)	12 (35.3%)
L	9 (26.5%)	27 (29.4%)	5 (14.7%)
VL	5 (14.7%)	11 (12.0%)	1 (3.0%)
EL	1 (3.0%)	2 (2.2%)	1 (3.0%)

Chi-square = 6.556 p = 0.1613 df = 4 (L, VL, and EL were combined)

Satisfactority

EU	20 (58.8%)	59 (64.1%)	30 (88.2%)
SU	6 (17.7%)	14 (15.2%)	3 (8.8%)
L	5 (14.7%)	9 (9.8%)	0 (0%)
VL	3 (8.8%)	6 (6.5%)	1 (2.9%)
EL	0 (0.0%)	4 (4.4%)	0 (0%)

Chi-square = 9.040 p = 0.0601 df = 4 (L, VL, and EL were combined)

Table 14: Perception of ‘unproductivity’ by native speakers

While none of these distributions can be shown to be significant by a chi-square test (though *satisfactority* comes close), the data still show some suggestive trends. Consistently, the ‘extremely unlikely’ assessment is higher for the ‘graduate’ group. This holds true of all five ‘unproductively’ formed sample naming units. These results suggest that people with more education make stronger judgments of grammaticality. People with more education could very likely be more committed to notions of correctness, including for Word-Formation Rules, and thus are more reluctant to accept words that appear ‘ungrammatical.’ The differences between college and graduate informants follow the same trend, though the differences are smaller. In general, the negative attitude to the ungrammaticality of coinages grows with the education of native speakers, with the major leap in this attitude characterizes the graduate group of the informants.

To conclude, the unproductivity experiment data provide us with another piece of evidence of education-related differences in the naming strategies.

3.2.5. Influence of Other Languages

3.2.5.1. General

The sample of informants encompasses 109 speakers²⁰ presently living in an English speaking country, but born to parents coming from non-English speaking countries. The data acquired from questionnaires indicate that while their parents are fluent in their

²⁰ The actual number of informants was 112. However, the group of polysynthetic language speakers was too small (3 informants). By implication, these informants were not included in our analysis.

mother tongue none of them can speak English with proficiency corresponding to a native speaker.

The expected total number of questionnaire responses (4 tasks with 5 sub-tasks each accounts for 20 responses per informant) provided by 109 informants is 2180. In fact, they provided 1012 relevant responses for the category of Agent nouns. The reasons why some of the responses have had to be excluded from the analysis are analogical to those in the native-speaker group of informants.

The sample includes some sociolinguistically complicated cases. Thus, for example, although an informant was born in Holland, his/her native language is Vietnamese and his/her parents were born in China and Vietnam. Another case is an informant born in Moldova with Romanian as a native language. One of his/her parents was born in Germany with Russian as a native language, and the other parent was born in Moldova with Romanian as a native language. Since the language most frequently spoken at home is also Romanian he is analyzed in the group of analytic/isolating languages. The same criterion is applied to an informant born in Switzerland one of whose parents was born in France with French as a native language.

The informants were divided into four groups based on the morphological typology of languages. It is generally known that there are hardly any morphologically pure languages. Given the focus of our experimental research on Agentive nouns, in classifying the languages the most important criterion was the prevailing morphological features of nouns. As a result we obtained the following groups

SYNTHETIC/AGLUTINATING (19 informants) – Korean, Japanese, Finnish, Hungarian, Estonian, Armenian, Swedish, Norwegian, Tagalog, Tonga

SYNTHETIC/FUSIONAL (17 informants) – German, Slovak, Russian, Polish, Croatian, Czech, Ukrainian, Arabic, Urdu

ANALYTIC/ISOLATING (73 informants) – French, Portuguese, Romanian, Italian, Spanish, Dutch, Chinese, Bulgarian, Bangla, Samoan, Creol, Afrikaans, Mandarin, Chinese, Vietnamese, Cantonese

POLYSYNTHETIC (3 informants) – Indonesian, Laotian, Hmong

3.2.5.2. Analysis of the Experimental Data

3.2.5.2.1. Productivity of Onomasiological Types

The most noticeable difference between native speakers and non-native speakers as demonstrated in Table 15 is the respective roles played by Onomasiological Types 1 and 3 in these two groups of informants. With the other three Onomasiological Types featuring almost identical productivity, the PR for the Onomasiological Type 1 is higher by about 4 % in the non-native group, and the PR of the Onomasiological Type 3 is lower by the same value in the same group of informants.

We may surmise that one of the reasons for this difference is as follows: since the informants, falling within the non-native group, do not master English as fluently as native speakers (levels B and lower in the questionnaire) their linguistic uncertainty makes them try very hard in the naming act to make their ‘products’ as comprehensible as possible, and therefore, most explicit. For this reason, they prefer Onomasiological Type 1. Obviously, this is a possible psychological motivation behind this preference. The second reason may

be connected with the structural characteristics of the non-English languages that were shaping the linguistic behaviour of the informants in the past.

	Native speakers	Non-native speakers
No. of responses	1300	1012
OT1	670 (51.5%)	561 (55.4%)
OT2	299 (23.0%)	236 (23.3.1%)
OT3	303 (23.3%)	193 (19.1%)
OT4	0 (0%)	0 (0%)
OT5	28 (2.2%)	22 (2.2%)
Chi-square 6.4094 p = 0.0933 df = 3 (OT4 was excluded from calculations)		

Table 15: PR (%) of Onomasiological Types (non-native speakers)

3.2.5.2.2. Productivity of Word-Formation Types

The differences discussed in the previous section cannot be, for obvious reasons, manifested at the level of WFTs. In spite of this fact, the agreement of the respective PRs is surprisingly high. While the top WFT ranking agreement was expected the PR differences are extraordinarily small (for the top five WFTs in succession: 0.19; 0.26; 0.61; 1.37; and 1.77 %, respectively).

	Native speakers	Non-native speakers
No. of responses	1300	1012
Object–Action–Agent	416 (32.0%)	322 (31.8%)
Action–Agent	304 (23.4%)	234 (23.1%)
Theme–Action–Agent	225 (17.3%)	169 (16.7%)
Instrument–Action–Agent	135 (10.4%)	119 (11.8%)
Location–Action–Agent	118 (9.1%)	74 (7.3%)

Chi-square = 4.6546 p = 0.4594 df = 5

Table 16 PR (%) of the top five Word-Formation Types (non-native speakers)

3.2.5.2.3. Productivity of Morphological Types

The suffix-based Morphological Types of [S+S+suffix] and [S + Suffix] for Agent names is understandable with respect to the large number of Agentive suffixes in English. When the central suffix-based Agent types are added up, they represent 84.08% for the native group and 77.67% for the non-native group. What makes the two groups of informants different is the much stronger role of the [S + suffix] MT in the native group of speakers (37.41% compared to 28.66% in the non-native group), and, on the other hand, a slightly higher PR of the [S+S+suffix] MT in the non-native group. These data correspond with the observations concerning the productivity of Onomasiological Types.

	Native speakers	Non-native speakers
No. of Responses	1300	1012
S + S + suffix	607 (46.7%)	496 (49.01%)
S + suffix	486 (37.4%)	290 (28.66%)
S + S	147 (11.3%)	163 (16.11%)

Chi-square = 26.1042 $p < .0001$ $df = 3$

Table 17: PR (%) of Morphological Types (non-native speakers)

3.2.5.2.4. Productivity of Word-Formation Rules

The results for the Onomasiological Types are also mapped onto the level of Word-Formation Rules. First, while in the group of native speakers there is one WFR in the ‘top 5 chart’ in which the determined constituent of the onomasiological mark is not expressed, there is no such WFR among the top five in the non-native group. Second, the strong position of OT2 is supported by two WFRs among the top 5, in which the determining constituent is not expressed (ranks 2 and 5) as opposed to only one such WFR in the native speaker group.

	Native speakers	Non-native speakers
No. of Responses	1300	1012
Object–Action–Agent		
N V -er	272 (20.9%)	222 (21.9%)
Action–Agent		
V -er	232 (17.9%)	159 (15.7%)
Theme–Action–Agent		
N V -er	138 (10.6%)	121 (11.96%)
Location–Action–Agent		
N V -er	92 (7.1%)	65 (6.4%)
Action–Agent		
V N	35 (2.7%)	57 (5.6%)

Chi-square = 16.3394 $p = 0.0059$ $df = 5$

Table 18: PR (%) of Word-Formation Rules (non-native speakers)

3.2.5.2.5. Summary

The comparison of the naming behaviour of the native and the non-native informants has shown considerable and significant differences. Their naming strategies appear to differ primarily in the non-native group of speakers laying much greater emphasis on the ‘accuracy of naming’, which implies explicitness especially in relation to the determined constituent of onomasiological mark.

3.2.5.3. Comparison of Three Cohorts of Influencing Languages

Given the non-existence of a word-formation typology of languages, the ‘background’ languages were divided into three groups, based on the morphological typology of noun, that is, the synthetic/agglutinating, synthetic/fusional, and analytic/isolating types. The

fourth group, the polysynthetic type, was represented by only three informants, representing three languages (Indonesian, Laotian and Hmong), and therefore, it was not included in the analysis.²¹

3.2.5.3.1. Onomasiological Types

An overview of the results for Onomasiological Types is given in Table 19.

No. of responses	Native speakers	Non-native linguistic background		
		Agglutinative	Fusional	Analytic
	1300	180	161	671
OT1	670 (51.5%)	109 (60.56%)	94 (58.39%)	358 (53.35%)
OT2	299 (23.0%)	38 (21.10%)	38 (23.60%)	160 (23.85%)
OT3	303 (23.3%)	28 (15.56%)	28 (17.40%)	137 (20.42%)
OT4	0 (0%)	0 (0%)	0 (0%)	0 (0%)
OT5	28 (2.2%)	5 (2.78%)	1 (0.62%)	16 (2.38%)

Chi-square = 12.6218 p = 0.1804 df = 9 (native v. all)

Chi-square = 7.21461 p = 0.0653 df = 3 (native v. agglutinative)

Chi-square = 5.21562 p = 0.1566 df = 3 (native v. fusional)

Chi-square = 2.17689 p = 0.5365 df = 3 (native v. analytic)

Table 19: PR (%) of Onomasiological Types by language background

A crucial observation following from Table 19 is that the hypothesis of the influence of the influence of language background seems to have been confirmed. English is predominantly an analytic language and therefore the results obtained from native speakers should be closest to those obtained from the isolating/analytic group of background languages. The data seem to suggest this, though the differences don't rise to statistical significance. There is a striking similarity between the naming tendencies in these two groups of informants, while the agglutinative and the fusional background languages deviate from the 'native' data in a noticeable way, as reflected in their lower p-values in the chi-square test. This primarily concerns the role played by Onomasiological Types 1 and 3. Furthermore, it is no surprise that the agglutinative group's PR for OT1 is the highest of all. This may be explained – in addition to the psychological reasons relevant to all three groups of non-native language background – by the morphological characteristics of agglutinative languages, aiming at expressing complex morphological meanings within one word.

3.2.5.3.2. Word-Formation Types

An overview of the results for Word-Formation Types is given in Table 20.

²¹ The classification of the languages in terms of the morphology of noun was based on Krupa/Genzor/Drozdiak (1983), Comrie (1981), and the Internet sources, <http://www.paul-raedle.de/vtrain/db-xx-info.htm>, <http://www.geocities.com/indoeurop/tree/balk/armenian.html>, <http://www.cambridge.org/uk/catalogue/catalogue.asp?isbn=0521573785&ss=exc>, <http://www.linguistics.emory.edu/POLYGLOT/morphology.html>, and on personal communication. Therefore, we wish to express our gratitude for help to Jan Don, A. Olofsson, S. Valera, and M. Volpe.

No. of responses	Native speakers	Non-native linguistic background		
		Agglutinative	Fusional	Analytic
	1300	180	161	671
Object-Action-Agent	416 (32.0%)	62 (34.4%)	50 (31.1%)	210 (31.3%)
Action-Agent	304 (23.4%)	38 (21.1%)	38 (23.6%)	158 (23.6%)
Theme-Action-Agent	225 (17.3%)	26 (14.4%)	27 (16.8%)	116 (17.3%)
Instrument-Action-Agent	135 (10.4%)	22 (12.22)	20 (12.42)	77 (11.48)
Location-Action-Agent	118 (9.1%)	13 (7.22)	11 (6.83)	50 (7.46)

Chi-square = 11.201 p = 0.7382 df = 15 (native v. all)

Chi-square = 2.6524 p = 0.7533 df = 5 (native v. agglutinative)

Chi-square = 3.5797 p = 0.6113 df = 5 (native v. fusional)

Chi-square = 6.8404 p = 0.2327 df = 5 (native v. analytic)

Table 20: PR (%) of Word-Formation Types by language background

Table 20 gives support to the observations given in 3.2.5.3.1. In each of the top five Word-Formation Types the PRs of native speakers and the ‘analytic’ language background group of informants are closer to each other than the results obtained from the other two groups, even if the differences between the PRs are small in general. Nonetheless, the ‘native-analytic’ comparison features extraordinarily small differences: 0.7; 0.17; 0.03; 1.10; and 1.62 respectively, for the first five WF Types.

3.2.5.3.3. Morphological Types

An overview of the results for Morphological Types is given in Table 21.

No. of Responses	Native speakers	Non-native linguistic background		
	Native	Agglutinative	Fusional	Analytic
	1300	180	161	671
S+S+suffix	607 (46.7%)	99 (55.0%)	85 (52.8%)	312 (46.5%)
S+suffix	486 (37.4%)	47 (26.1%)	41 (25.5%)	202 (30.1%)
S+S	147 (11.3%)	22 (12.2%)	29 (18.0%)	112 (16.7%)
Other	60 (4.6%)	4 (2.2%)	5 (3.1%)	25 (3.7%)

Chi-square = 32.5464 p = 0.0001 df = 9 (agglutinative v. all)

Chi-square = 10.107 p = 0.0176 df = 3 (agglutinative v. native)

Chi-square = 2.1152 p = 0.5488 df = 3 (agglutinative v. fusional)

Chi-square = 5.6579 p = 0.1294 df = 3 (agglutinative v. analytic)

Table 21: PR (%) of Morphological Types by language background

Given the prevailing word-formation tendencies in the languages under evaluation, one might, in general, expect major differences in Morphological Types and Word-Formation Rules. This follows from the purely formal nature of the traditional classification of word-formation processes. Thus, we might expect that the share of the suffix-based types and rules in agglutinative languages will be higher than that in the native group of speakers and

in the other two groups of background language. These expectations have been confirmed to a considerable degree, especially with regard to the [S+S+suffix] Morphological Type where the dominance of the agglutinative background is dominant, especially with regard to the native speaker and the isolating background groups of informants. The only unexpected outcome is an even lower PR of the [S+S] MT in the native speaker group than the PR of the same type in the agglutinative group. A remarkable parallel between the Productivity Rates of the suffixed MTs in the native and the ‘isolating background’ groups can also be traced here, with the exception of the [S+S] type.

3.2.5.3.4. Word-Formation Rules

An overview of the results for Word-Formation Rules is given in Table 22.

No. of Responses	Native speakers	Non-native linguistic background		
		Agglutinative	Fusional	Analytic
Object–Action–Agent N V -er	1300	180	161	671
Action–Agent V -er	272 (20.9%)	49 (27.2%)	36 (22.4%)	137 (20.4%)
Theme–Action–Agent N V -er	232 (17.9%)	26 (14.4%)	24 (14.9%)	109 (16.3%)
Location–Action–Agent N V -er	138 (10.6%)	19 (10.6%)	22 (13.7%)	80 (11.9%)
Action–Agent V N	92 (7.1%)	12 (6.7%)	11 (6.8%)	42 (6.3%)
	35 (2.7%)	8 (4.4%)	12 (7.4%)	37 (5.5%)

Chi-square = 23.5232 p = 0.0736 df = 15 (native v. all)

Chi-square = 6.3357 p = 0.2749 df = 5 (native v. agglutinative)

Chi-square = 13.4806 p = 0.0192 df = 5 (native v. fusional)

Chi-square = 11.6355 p = 0.0401 df = 5 (native vs. analytic)

Table 22: PR (%) of Word-Formation Rules by language background

Table 22 also demonstrates a coincidence between the native speakers and those with the analytic language background. The PRs of the most productive WF Rule in the two groups of informants are almost identical, significantly differing from the agglutinating background PR, and the same situation may be observed for the second most productive WF Rule. For other WFRs the differences between the individual groups of informants are minimal.

3.2.5.4. Comparison of the ‘Unproductivity’ Results

3.2.5.4.1. Native vs. Non-native Speakers (as a Whole)

If we concluded in Sections 3.2.3.3 that the perception of unproductivity in native speakers is very strong Table 23 shows us that a similar statement is applicable to the non-native informants. In spite of this general conclusion, there **are** certain differences between the two groups of speakers of English. While the share of the ‘extremely unlikely’ responses in

the native group is 56.1%, in the group of non-native speakers it is less (50.0%) which indicates that the pressure of productive WF rules is perceived by native speakers a little stronger. This tendency has been borne out in three of the five ‘unproductively’ coined naming units. Two gaps are significant: almost 16% for *satisfactority* and over 12% for *engroupment*. This difference is mostly compensated at the next lower level, the level of ‘somewhat unlikely’ answers. In one case (*thinnen*) we might speak of a draw because the percentages were almost identical (64.54% vs. 64.76%). The non-native speakers manifested about 3% higher distaste for *sleepable*.

The non-native speakers feature higher percentages at the medium assessment level, i.e., at the level of the ‘likely’ answers. The biggest assessment gap at this level is observed for *satisfactority* (10%); in two other cases, *swimmee* and *sleepable*, the gap is about 5% in favour of the non-native speakers.

The results are not very conclusive in one or the other direction. What may be assumed based on them is that native speakers are slightly more ‘aware’ of the productive WF processes. On the other hand, the differences are not significant. In both groups of respondents, we can observe certain will to creative ‘experimentation’ which depends on the nature of the constraint violated. Table 23 gives a comparison of the two groups of informants for the individual ‘non-words’.

	Native speakers	Non-native speakers
Number of responses	142	109
<i>Engroupment</i>		
EU	82 (57.8%)	51 (46.8%)
SU	43 (30.3%)	41 (37.6%)
L	14 (9.9%)	9 (8.3%)
VL	2 (1.4%)	6 (5.5%)
EL	0 (0.0%)	2 (1.8%)
Chi-square = 3.2609 p = 0.20 df = 2		
<i>Thinnen</i>		
EU	92 (64.5%)	72 (66.1%)
SU	30 (21.3%)	26 (23.9%)
L	13 (9.2%)	9 (8.3%)
VL	6 (4.3%)	1 (0.9%)
EL	1 (0.7%)	1 (0.9%)
Chi-square = 1.0166 p = 0.60 df = 2		
<i>Swimmee</i>		
EU	91 (64.50%)	59 (54.1%)
SU	31 (22.00%)	32 (29.4%)
L	15 (10.60%)	14 (12.8%)
VL	4 (2.80%)	4 (3.7%)
EL	0 (0%)	0 (0.0%)
Chi-square = 2.8198 p = 0.24 df = 2		

Sleepable

EU	40 (28.00%)	36 (33.0%)
SU	48 (33.60%)	35 (32.1%)
L	36 (25.20%)	30 (27.5%)
VL	16 (11.20%)	7 (6.4%)
EL	3 (2.10%)	1 (0.9%)

Chi-square = 0.7811 p = 0.68 df = 2

Satisfactoriness

EU	93 (66.00%)	58 (53.2%)
SU	24 (17.00%)	23 (21.1%)
L	13 (9.20%)	17 (15.6%)
VL	8 (5.70%)	7 (6.4%)
EL	3 (2.10%)	4 (3.7%)

Chi-square = 4.4179 p = 0.11 df = 2

Table 23: Perception of 'unproductivity': Native vs. non-native informants

3.2.5.4.2. Non-Native Speakers (Individual Types)

As for the internal structure of the non-native informants, a clearly highest resistance to unproductively coined naming units is exercised by those with a fusional language background, much higher than the other two groups of informants. With the exception of *satisfactoriness*, the differences between the individual groups of informants are very high. For example, *thinnen*, the difference between agglutinative and the fusional groups is over 40%.

We have no explanation for these results. By all accounts, however, the acceptability/non-acceptability of a naming unit coined by violating a restriction on productivity is not influenced by the type of a background language.

	Native	Non-native		
		Agglutinative	Fusional	Analytic
Number of responses	142	(19)	(17)	(73)

Engroupment

EU	82 (57.8%)	9 (47.4%)	10 (58.8%)	32 (43.8%)
SU	43 (30.3%)	7 (36.8%)	5 (29.4%)	29 (39.7%)
L	14 (9.9%)	2 (10.5%)	1 (5.9%)	6 (8.2%)
VL	2 (1.4%)	1 (5.3%)	0 (0%)	5 (6.9%)
EL	0 (0%)	0 (0%)	1 (5.9%)	1 (1.4%)

Chi-square = 4.5140 p = .61 df = 6 (native vs. all types; L, VL, and EL combined)

Chi-square = 0.8366 p = .66 df = 2 (native vs. agglutinative; ; L, VL, and EL combined)

Chi-square = 0.0094 p = .99 df = 2 (native vs. fusional; ; L, VL, and EL combined)

Chi-square = 4.0221 p = .13 df = 2 (native vs. analytic; ; L, VL, and EL combined)

Thinnen

EU	92 (64.5%)	9 (47.4%)	15 (88.2%)	48 (65.8%)
SU	30 (21.3%)	7 (36.8%)	1 (5.9%)	18 (24.7%)
L	13 (9.2%)	3 (15.8%)	1 (5.9%)	5 (6.9%)
VL	6 (4.3%)	0 (0%)	0 (0%)	1 (1.4%)
EL	1 (0.7%)	0 (0%)	0 (0%)	1 (1.4%)

Chi-square = 7.9900 p = .24 df = 6 (native vs. all types; L, VL, and EL combined)

Chi-square = 2.6456 p = .27 df = 2 (native vs. agglutinative; ; L, VL, and EL combined)

Chi-square = 3.8233 p = .15 df = 2 (native vs. fusional; ; L, VL, and EL combined)

Chi-square = 1.0520 p = .59 df = 2 (native vs. analytic; ; L, VL, and EL combined)

Swimmee

EU	91 (64.5%)	11 (57.9%)	14 (82.4%)	34 (46.6%)
SU	31 (22.0%)	3 (15.8%)	3 (17.7%)	26 (35.6%)
L	15 (10.6%)	5 (26.3%)	0 (0%)	9 (12.3%)
VL	4 (2.8%)	0 (0%)	0 (0%)	4 (5.5%)
EL	0 (0%)	0 (0%)	0 (0%)	0 (0%)

Chi-square = 13.7123 p = .03 df = 6 (native vs. all types; L, VL, and EL combined)

Chi-square = 7.12163 p = .03 df = 2 (native vs. agglutinative; ; L, VL, and EL combined)

Chi-square = 21.0597 p < .0001 df = 2 (native vs. fusional; ; L, VL, and EL combined)

Chi-square = 11.4897 p = .0031 df = 2 (native vs. analytic; ; L, VL, and EL combined)

Sleepable

EU	40 (28.0%)	7 (36.8%)	7 (41.2%)	22 (30.1%)
SU	48 (33.6%)	5 (26.3%)	7 (41.2%)	23 (31.5%)
L	36 (25.2%)	6 (31.6%)	3 (17.7%)	21 (28.8%)
VL	16 (11.2%)	0 (0%)	0 (0%)	7 (9.6%)
EL	3 (2.1%)	1 (5.3%)	0 (0%)	0 (0%)

Chi-square = 3.7022 p = .7169 df = 6 (native vs. all types; L, VL, and EL combined)

Chi-square = 0.7358 p = .6921 df = 2 (native vs. agglutinative; ; L, VL, and EL combined)

Chi-square = 2.9737 p = .2260 df = 2 (native vs. fusional; ; L, VL, and EL combined)

Chi-square = 0.1414 p = .9317 df = 2 (native vs. analytic; ; L, VL, and EL combined)

Satisfactority

EU	93 (66.0%)	10 (52.6%)	9 (52.9%)	39 (53.4%)
SU	24 (17.0%)	5 (26.3%)	5 (29.4%)	13 (17.8%)
L	13 (9.2%)	4 (21.1%)	2 (11.8%)	11 (15.1%)
VL	8 (5.7%)	0 (0%)	0 (0%)	7 (9.6%)
EL	3 (2.1%)	0 (0%)	1 (5.9%)	3 (4.1%)

Chi-square = 6.7976 p = .3399 df = 6 (native vs. all types; L, VL, and EL combined)

Chi-square = 1.4154 p = .4927 df = 2 (native vs. agglutinative; ; L, VL, and EL combined)

Chi-square = 1.8914 p = .3884 df = 2 (native vs. fusional; ; L, VL, and EL combined)

Chi-square = 4.3977 p = .1109 df = 2 (native vs. analytic; ; L, VL, and EL combined)

Table 24: Perception of 'unproductivity' (%): Native vs. Non-native informants

3.2.6. Additional Remarks

3.2.6.1. Double Formal Indication of Agent – Redundancy in Word-Formation

Strangely, the questions of word-formation redundancy have been paid little attention in the relevant literature.²² The redundancy phenomenon in word-formation is closely related to one of the central points of discussion of our research, in particular, the conflict between the explicitness of expression and the economy of expression.

Double indication of a single conceptual category runs counter to the very notion of the economy of expression, and counter to a linguistic notion that there should be one to one correspondence of conceptual and formal categories in word-formation. Certainly, the state of isomorphy is an ideal one in morphology in general. In practice, there are a number of cases where a single conceptual category is represented in a language by a number of allomorphs/synonymous morphemes (Agent nouns themselves are a case in point). What is rare in the English language, however, is the doubling of the same formal means within one naming unit, which introduces redundancy. This is captured by Lieber's **Redundancy Restriction** (2004: 161):

- (27) The Redundancy Principle
Affixed do not add semantic content that is already available within a base word (simplex or derived).

Therefore, it may be surprising to find relatively numerous cases of this sort in our research, including *butter-inner*, *hanger-onner*, *butter-innist*, *weberer*, *shoe-tier-upper*, *grass-cutter-upper*, on the one hand, and *bird-fisherman*, *shoe-tierman*, *hangerman*, on the other. The former type, characteristic of nonstandard and informal language and casual, perhaps jocular, speech, may be accounted for by language users feeling uncomfortable putting Agentive suffix on the particle of phrasal verbs. Yet, at the same time, they recognize that agentive suffixes go at the end, so they end up putting a suffix on the particle as well as the verb. With reference to the phrasal verb *butt in*, the American member of the evaluative team finds *butter inner* more 'natural' than *butt inner* or even *butter in*. This assumption has been experimentally acknowledged: the proportion between the occurrences of *butter inner* and *butter in* in our research is 17 to 8, with zero occurrence of *butt inner*. As for the internal structure of *butter inner*, students selected it most of all occupational categories – nine times, which is more than 50%. On the other hand, the situation with *hanger onner* is quite opposite. Its three occurrences represent just 50% of the occurrences of *hanger-on*. *Butter-innist* only occurred once, and its 'author' is a female manager from Great Britain.

Weberer is a different case because *to web* is not a phrasal verb, and is difficult to explain. We suspect that it is related to double comparatives that show up – somehow speakers do not recognize the first suffix and end up putting another suffix on. Let us recall the fairly recent movie called *Dumb and Dumberer* which is a sequel to *Dumb and Dumber*. This naming unit only occurred once, and was proposed by a female informant in the category of 'Management'. The informant lives in Great Britain.²³

²² Important exceptions to this rule are Plag (1999) and Lieber (2004).

²³ All in all, the above-given cases are not mere experimental oddities, which is borne out by the existence of established words. Lieber (2004: 164) refers to OED citing a number of similar examples, such as *checkerist*, *consumerist*, *collegianer*, *musicianer*, etc. In addition, the Agent-related redundancy is not the only type of redundancy in English. A much more frequent type are the *-ic-al* adjectives. For further

3.2.6.2. Suffix *-sky* (*-insky*)

This suffix occurred in the naming unit *buttinsky*, apparently attached to the verbal base *butt-in* in the meaning ‘someone who has something to do with V’. It is certainly not a common suffix across varieties of English; rather, it seems to be an Eastern U.S. regionalism. There is definitely something playful or slangish about it. Its origin may be supposed to be in the *-sky* formative encountered in Polish names. Its connotations seem to be slightly pejorative, perhaps suggesting someone who is boorish in connection with an item or an action. It may be more popular among working class, and indeed, in our experimental results, three male and one female informants and all self-identified in the ‘Manual Work’ category chose this option. These four informants represent a third of all those identified in the ‘Manual Work’ occupation.

3.2.6.3. Blends

Blends appear to be quite popular with coiners, supporting the economy of expression at the expense of meaning clarity. The experiment came up with several interesting blends:

Persniskigardener – a blend of *gardener* and *persnickety* ‘fastidious, overly attentive to details, excessively demanding’. Its single occurrence is related to a male manual worker from the USA.

Blondorionious – there are two possible interpretations of this naming unit. Either it is a blend combining *blond* and *erroneous*, which gives the meaning ‘someone who is erroneous about blonds’. Another possible interpretation is one based on a pretentious (and therefore playful) suffix.

Blonde-ogynist – a blend of *blonde* + *misogynist*, proposed by an American female teacher. This is a fascinating formation, since it suggests that the blonds who are ridiculed are female. It also leaves out the part of *mysoginist* that explicitly marks ‘hatred’ and reinterprets the last part for that.

Laceanomist – a blend of *lace* + *-onomist*. *-onomist* shows up on a number of words indicating an expert at a (usually academic or professional) field, such as ‘economist’ or ‘agronomist’. This blend appears to be used to attach some prestige, or at least the notion of a profession, to someone who ties shoes.

3.2.6.4. Other Interesting Cases

Car-top boogieborder – this naming unit, proposed by an American male teacher, is used to denote a person depicted in the picture on car-top, and makes use of figurative expression, i.e., someone who rides on the car top as they would a boogie board (a small surf board made to be ridden in the prone position, as in the picture).

Anal-lawn maintenance worker; *anal grass snipper* and *anal-retentive* – the basic term in this group, *anal-retentive*, comes from Freud’s notions of child development. Apparently, in Freud’s thought, the stages of toilet-training can lead some to become too preoccupied with structure and order and detail, and this is the general meaning of *anal*. Thus, for

discussion see Plag (1999) and Lieber (2004).

instance, when students think that some English usage rules are too picky, they might say something like ‘that rule is just anal’. The first two terms were used by American female teachers, the last one by an American natural scientist.

Representor, race representor – the creative aspect is manifested here very clearly; while there is a word *representative* in the meaning ‘a person duly authorized to act or speak for another or others’, the coiners (an American male young unemployed informant and an American female teacher, respectively) apparently wanted to emphasize the new role of a person who represents the whole mankind by having recourse to a fully grammatical coinage using the suffix *-er*.

4. Conclusions

1. The research has confirmed the concept of **word-formation conceived as creativity within productivity constraints**. While the effectiveness of ‘productivity constraints’ are manifested by the types and rules with high Productivity Rates and by the extensive coincidence of their ranking in the various experimental groups, the word-formation ‘creativity’ is borne out by the diversity of the types and rules fulfilling the same function within a particular conceptually defined cluster. The present research gives ample evidence in favour of this approach to productivity, and shows that rather than excluding each other (as traditionally believed) **productivity and creativity co-exist**.

2. The proposed method of productivity calculation proved to be a feasible tool for an objective evaluation of the role of the individual types and rules without any unjustified preference for any particular word-formation process (as opposed to the mainstream affixation-oriented approaches). This method makes it possible to evaluate the productivity at different levels of generalization, to reflect its different aspects, including the most general onomasiological level; onomasiological structure (logico-semantic relations); onomatological structure (formal realization of coinages); and the interrelation of the onomasiological and the onomatological structures (established by the Morpheme-to-Seme Assignment principle). Importantly, each of these levels of productivity calculation encompasses any and all of the traditional formally defined word-formation processes. Furthermore, this method makes it possible to avoid the classification problems so characteristic of the generative approach to word-formation (compounding vs. affixation, bracketing paradoxes) thanks to the fact that all word-formation processes are treated in a consistent onomasiological manner, and therefore, defined on the basis of a single, unifying principle.

As far as the specific targets of our experimental research are concerned the following conclusions may be drawn:

(a) The conflict between the explicitness of expression and the economy of expression in the field of Agent names favours the **explicitness tendency**. Language users tend to make use of the types and rules which employ the crucial Actional constituent of the onomasiological, and mainly, onomatological structure. It is for this reason that the most productive Onomasiological Type is OT1, the most productive Word-Formation Types are **[Object–Action–Agent]** and **[Action–Agent]**, the most productive Morphological Types are **[S+S+suffix]** and **[S+suffix]**, and the most productive Word-Formation Rule is (28)

(28) Object–Action–Agent
 N V -er

(b) The research has borne out the hypothesis of **sociolinguistic conditioning of the individual acts of word-formation**. The analysis of the results by occupation has shown that there is a tendency indicating **different strategies taken by education-related and ‘other’ professions in the implementation of naming acts**. While the former group have a stronger preference for the explicit types and rules, the latter group more frequently favours the more ‘economic’ solutions. Furthermore, the level of education appears to have similar effects: while **native speakers with university education prefer more precise names, lower educated speakers are more frequently driven by the principle of economy of expression**.

The influence of language-background seems to be equally important. The preference for the ‘Action-expressed’ Onomasiological Types among non-native speakers is even much stronger than with native speakers, especially the role of Onomasiological Type 2 is extremely strong. This is, logically, projected onto the high Productivity Rate of the Morphological Type **[S + suffix]** in this group of speakers, and the absence among the top five Word-Formation Rules of a rule in which the determined constituent is not expressed. In general, the naming strategies of the two basic groups of speakers seem to differ because **non-native speakers seem to lay even greater emphasis on the explicitness of expression than native speakers**.

The influence of **linguistic background plays its role in the naming strategies of non-naming speakers**. Although the limited sample of informants with ‘Germanic linguistic background’ does not enable us to draw any indisputable conclusions, the agreement of the results between them and the native English speakers in terms of almost identical preference for affixal types is remarkable.

(c) It has been shown that any assessment of the influence of any of the above-discussed sociolinguistic factors must be related to the specific **conceptually determined category** of the cluster (Agent, Patient, Instrument, etc.).

(d) Finally, the research into ‘unproductivity’ has demonstrated that the perception of ‘unproductivity’ among both native and non-native speakers in general is strong. On the other hand, the existence of a relatively high number of the ‘likely’ responses acknowledges their feeling for a **creative approach to naming**. This **mainly applies to the group of students** who most readily accept unconventional naming units and break the existing rules. This does not seem to be a surprise, and might be accounted for psychologically by the dynamism of the young generation compared to the more conservative generation of their parents. Moreover, the share of ‘extremely unlikely’ answers in the university-educated informants is much higher than in the lower-educated groups, which suggests that **the awareness of grammaticality of higher-educated speakers is stronger**. The tolerance to ‘creativity’ (even the creativity that trespasses grammaticality) characteristic of the young generation is also typical of speakers with lower education, even if the reasons underlying this fact may partly differ in these two groups of speakers.

The native–non-native comparison shows that while the both groups demonstrate the awareness of unproductivity, there are some differences between the two groups: the pressure of productive Word-Formation Rules is perceived by native speakers a little stronger.

Pavol Štekauer
Slávka Tomaščíková
Štefan Franko
Department of British and American
Studies Faculty of Arts
Prešov University
17. novembra 1
08078 Prešov, Slovakia
e-mail: stekpal@unipo.sk

Don Chapman
Dept. of Linguistics and English Language
Brigham Young University
Provo, UT 84602, USA
e-mail: don_chapman@byu.edu

References

- Adams, Valerie (2001), *Complex Words in English*, Harlow: Longman.
- Allen, Margaret R. (1978), *Morphological Investigations*, Ph.D. Univ. Connecticut.
- Anderson, Stephen R. (1982), "Where's Morphology", *Linguistic Inquiry* 13: 571-612.
- Anderson, Stephen R. (1992), *A-Morphous Morphology*, Cambridge: Cambridge University Press.
- Anshen, Frank / Aronoff, Mark (1981), "Morphological Productivity and Phonological Transparency", *Canadian Journal of Linguistics* 1: 63-72.
- Aronoff, Mark (1976), *Word Formation in Generative Grammar*, [Linguistic Inquiry Monograph 1], Cambridge, MA: MIT Press.
- Aronoff, Mark / Schvaneveldt, Roger W. (1978), "Testing Morphological Productivity", *Annals of the New York Academy of Sciences* 318: 106-114.
- Aronoff, Mark (1983), "Potential Words, Actual Words, Productivity and Frequency", in: *Preprints of the Plenary Session Papers of the XIIIth International Congress of Linguists*, 163-170, Tokyo.
- Baayen, Harald (1989), *A Corpus-Based Approach to Morphological Productivity: Statistical Analysis and Psycholinguistic Interpretation*, Ph.D. Diss. Free University of Amsterdam.
- Baayen, Harald (1992), "A Quantitative Approach to Morphological Productivity", in: Booij, Geert / Marle, Jaap van (eds.), *Yearbook of Morphology 1991*, 109-149, Dordrecht: Kluwer.
- Baayen, Harald (1993), "On Frequency, Transparency, and Productivity", in: Booij, Geert / Marle, Jaap van (eds.), *Yearbook of Morphology 1992*, 181-208, Dordrecht: Kluwer.
- Baayen, Harald (1994a), "Derivational Productivity and Text Typology", *Journal of Quantitative Linguistics* 1: 16-34.
- Baayen, Harald (1994b), "Productivity in Language Production", *Language and Cognitive Processes* 9: 447-469.
- Baayen, Harald (1996), "The Effects of Lexical Specialization on the Growth Curve of Vocabulary", *Computational Linguistics* 22: 455-480.
- Baayen, Harald (1997), "Markedness and Productivity", in: Dressler, Wolfgang, U. / Prinzhorn, Martin / Rennison, John R. (eds.), *Advances in Morphology*, 189-200, Berlin: Mouton de Gruyter.
- Baayen, Harald / Lieber, Rochelle (1991), "Productivity and English Derivation: A Corpus Based Study", *Linguistics* 29: 801-843.
- Baayen, Harald / Renouf, Antoinette (1996). "Chronicling *The Times*: Productive Lexical Innovations in an English Newspaper", *Language* 72: 69-96.
- Barker, Chris (1998), "Episodic *-ee* in English: A Thematic Role Constraint on New Word Formation", *Language* 4: 695-727.
- Bauer, Laurie (1983), *English Word-Formation*, Cambridge: Cambridge University Press.
- Bauer, Laurie (2001), *Morphological Productivity*, [Cambridge Studies in Linguistics 95], Cambridge: Cambridge University Press.
- Beard, Robert E. (1995), *Lexeme-Morpheme Base Morphology. A General Theory of Inflection and Word Formation*, State University of New York Press.
- Bolinger, Dwight (1975), *Aspects of Language*, New York: Harcourt Brace Jovanovich.
- Botha, Rudolf P. (1984), *Morphological Mechanisms: Lexicalist Analysis of Synthetic Compounding*, Oxford: Pergamon Press.
- Chomsky, Noam (1970), "Remarks on Nominalization", in: Jacobs, Roderick A. / Rosenbaum, Peter S. (eds.), *Readings in English Transformational Grammar*, 184-221, Waltham, MA: Ginn.
- Comrie, Bernard (1983), *Language Universals and Linguistic Typology*, Oxford: Basil Blackwell.
- Crystal, David (1997), *Cambridge Encyclopaedia of the English Language*, Cambridge: Cambridge University Press.

- Di Sciullo, Anna-Maria / Williams, Edwin (1987), *On the Definition of Word*, [Linguistic Inquiry Monograph 14], Cambridge, MA: MIT Press.
- Dokulil, Miloš (1962), *Tvoření slov v češtině I. Teorie odvozování slov*, Praha: Nakladatelství Československé akademie věd.
- Dressler, Wolfgang U. / Ladányi, Mária (2000), "Productivity in Word Formation (WF): A Morphological Approach", *Acta Linguistica Hungarica* 47: 103-144.
- Fabb, Nigel (1988), "English Suffixation Is Constrained Only by Selectional Restrictions", *Natural Language and Linguistic Theory* 6: 527-539.
- Geeraerts, Dirk (1983), "Reclassifying semantic change". *Quaderni di semantica* 4: 217-240.
- Gleitman, Lila R. / Gleitman, Henry (1970), *Phrase and Paraphrase: Some Innovative Uses of Language*, New York: W. W. Norton and Co.
- Grzega, Joachim (2002), "Some Aspects of Modern Diachronic Onomasiology". *Linguistics* 40: 1021-1045.
- Halle, Morris (1973), "Prolegomena to a Theory of Word Formation", *Linguistic Inquiry* 1: 3-16.
- Jackendoff, Ray S. (1975), "Morphological and Semantic Regularities in the Lexicon", *Language* 51: 639-671.
- Kiparsky, Paul (1982), "Lexical Morphology and Phonology", in: Yang, I. S. (ed.), *Linguistics in the Morning Calm: Selected Papers from SICOL-1981*, 3-91, Seoul: Hanshin Publishing Company.
- Krupa, Viktor / Genzor, Jozef / Drozdík, Ladislav (1983), *Jazyky sveta*, Bratislava: Obzor
- Lees, Robert B. (1960), *The Grammar of English Nominalizations*, Bloomington (IN): Indiana University Press.
- Lieber, Rochelle (2004), *Morphology and Lexical Semantics*, Cambridge: Cambridge University Press.
- Lieber, Rochelle (2005), "English Word-Formation Processes. Observations, Issues, and Thoughts on Future Research", in: Štekauer, Pavol / Lieber, Rochelle (eds.), *Handbook of Word-Formation*, 375-427, Dordrecht: Springer.
- Marchand, Hans (1960), *The Categories and Types of Present-Day English Word-Formation*, Wiesbaden: Otto Harrassowitz.
- Marle, Jaap van (1985), *On the Paradigmatic Dimension of Morphological Creativity*. Dordrecht: Foris.
- Marle, Jaap van (1986), "The Domain Hypothesis: The Study of Rival Morphological Processes", *Linguistics* 24: 601-627.
- Mohanan, Karuvannur Puthanveetil (1986), *The Theory of Lexical Phonology*, Dordrecht: Reidel.
- Plag, Ingo (1996), "Selectional Restrictions in English Suffixation Revisited: A Reply to Fabb (1988)", *Linguistics* 34: 769-798.
- Plag, Ingo (1999), *Morphological Productivity: Structural Constraints in English Derivation*, Berlin/New York: Mouton de Gruyter.
- Plag, Ingo (2004), "Syntactic Category Information and the Semantics of Derivational Morphological Rules", *Folia Linguistica* 38: 193-225.
- Raffelsiefen, Renate E. (1992), "A Nonconfigurational Approach to Morphology", in: Aronoff, Mark (ed.), *Morphology Now*, 133-162, Albany: State University of New York Press.
- Rainer, Franz (1988), "Towards a Theory of Blocking", in: Booij, Geert / Marle Jaap, van (eds.), *Yearbook of Morphology 1988*, 155-185, Dordrecht: Foris.
- Rainer, Franz (2005), "Constraints on Productivity", in: Štekauer, Pavol / Lieber, Rochelle (eds.), *Handbook of Word-formation*, 335-352, Dordrecht: Springer.
- Roeper, Thomas and Siegel, M. E. A. (1978), "A Lexical Transformation for Verbal Compounds", *Linguistic Inquiry* 9: 199-260.
- Scalise, Sergio (1984), *Generative Morphology*, Dordrecht: Foris.
- Selkirk, Elisabeth O. (1982), *The Syntax of Words*, [Linguistic Inquiry Monograph 7], Cambridge (MA): MIT Press.
- Štekauer, Pavol (1994), "On Productivity in Word-Formation", *Linguistica Pragensia* 2: 67-82.
- Štekauer, Pavol (1998), *An Onomasiological Theory of English Word-Formation*, Amsterdam/Philadelphia: John Benjamins.
- Štekauer, Pavol (1999), "K teórii produktivity v anglickej slovtvorbe", *Časopis pro moderní filologii* 2: 68-82
- Štekauer, Pavol (2001). "Fundamental Principles of an Onomasiological Theory of English Word-Formation", *Onomasiology Online* 2: 1-42 [www.onomasiology.de].
- Štekauer, Pavol (2003), "On Some Issues of an Onomasiological Approach to Productivity", *Ranam* 36. [ESSE 6 – Strasbourg 2002, 2 – linguistics]: 93-99.
- Štekauer, Pavol (2005a), *Meaning Predictability in Word Formation*, Amsterdam/Philadelphia: John Benjamins.
- Štekauer, Pavol (2005b), "Onomasiological Approach to Word-formation", in: Štekauer, Pavol / Lieber, Rochelle (eds.), *Handbook of Word-Formation*, 207-232, Dordrecht: Springer.

- Strauss, Steven L. (1982), *Lexicalist Phonology of English and German*, Dordrecht/Cinnaminson: Foris Publications.
- Williams, Edwin (1981), "On the Notions 'Lexically Related' and 'Head of a Word'", *Linguistic Inquiry* 2: 245-274.
- Zimmer, Karl E. (1964), *Affixal Negation in English and Other Languages: An Investigation of Restricted Productivity*, [Supplement to *Word*, Monograph 5], New York.

first version received 24 Januar 2005
revised version received 17 July 2005

Appendix

Word Choices Survey

We are trying to learn more about the words people use for new or unusual situations. We would appreciate your taking a few minutes to fill out this questionnaire. Your participation is entirely voluntary.

INSTRUCTIONS: This questionnaire asks your *opinions* about words. This is NOT a test, and there are no "right" answers. We don't care whether you make up new words for the answers or whether you choose words that already exist in English. We just want to see what words you think will work best for a few situations.

Task 1. Choose the word that you think is the most suitable for the person described in the question.

1. A person whose smiling face is used for billboard advertisements:

- | | |
|-------------|-----------------|
| a. smiler | e. smile-person |
| b. smilist | f. smile |
| c. smilant | g. other: _____ |
| d. smileman | |

2. A person who dials telephone numbers with a feather:

- | | |
|---------------|-----------------|
| a. featherer | d. featherman |
| b. featherist | e. other: _____ |
| c. featherant | |

3. A person who frequently interrupts other people when they are talking:

- | | | | |
|-----------------|-----------------|-----------------|-----------------|
| a. interrupter | f. butter-inner | i. cut-in | n. interposer |
| b. interruptist | g. butt-innist | j. cutter-in | o. interposist |
| c. interruptant | h. butt-insky | k. cutter-inner | p. other: _____ |
| d. butt-in | | l. cutt-innist | |
| e. butter-in | | m. cutman | |

4. A person who believes in miracles:

- | | |
|----------------------------|---------------------|
| a. miraclist or miraculist | e. miracle-believer |
| b. miracler | f. miracle-hoper |
| c. miraclant or miraculant | g. miracle-hopist |
| d. miracle-man | h. miracle-hope |
| | i. other: _____ |

5. A person who is obsessed by something:

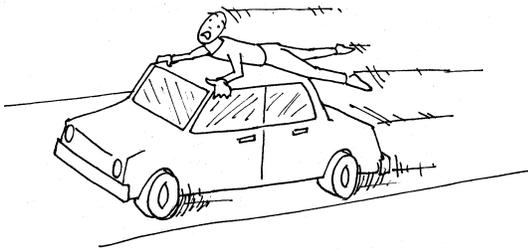
- | | |
|-----------------|---------------------|
| a. an obsessee | e. an obsess |
| b. an obsessor | f. an obsession-man |
| c. an obsessant | g. an obsessive |
| d. an obsessist | h. other: _____ |

Task 2. Each question describes a person in an unusual situation. If you had to come up with a name or title for the person, what would it be? You may make up a word or choose a word that already exists in English.

1. Suppose that space aliens were about to land on Earth for the first time. What would you call a person who was supposed to meet them as a representative of the human race?
2. What would you call someone who does research about spider webs?
3. What name or title would you use for someone who always tells blond jokes?
4. What name or title would you give a basketball player who always hangs onto the rim after a slam-dunk?
5. Suppose that a woman has a clone made of herself. Then suppose that a man has a clone made of himself. Now suppose that the two clones marry each other and have a child. What would you call the child?

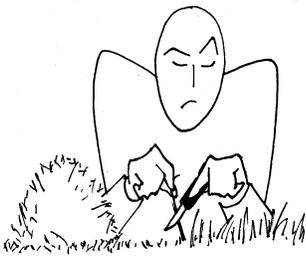
Task 3. Each picture below shows a person performing an unusual action. If you had to come up with a name or title for the person in each picture, what would it be? You may invent a word or choose a word that already exists in English.

1)



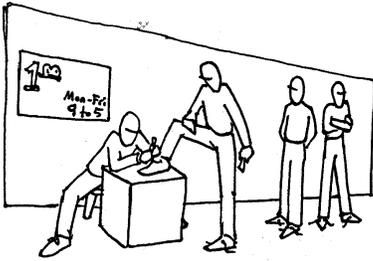
What name or title would you give to this person? _____

2)



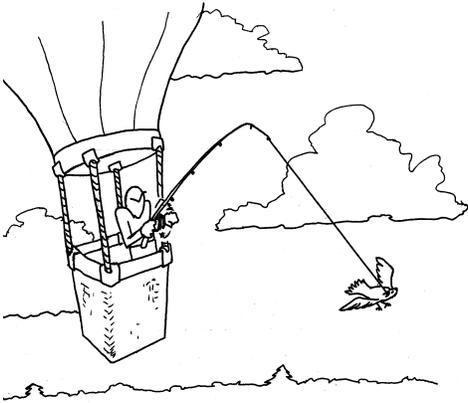
What name or title would you give to this person? _____

3



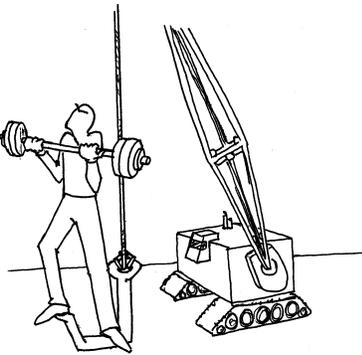
What name or title would you give to this person? _____

4)



What name or title would you give this person? _____

5)



What name or title would you give to this person? _____

Task 4. In this task, there are five words. Use each word in a sentence, even if you think it isn't an English word. Then rate how likely you and other English speakers would be to use the word.

1. *engroupment*

a. Use this word in a sentence:

b. How likely would you use this word? Circle one.

Extremely Unlikely	Somewhat Unlikely	Likely	Very Likely	Extremely Likely
-----------------------	----------------------	--------	----------------	---------------------

c. How likely would others use this word? Circle one.

Extremely Unlikely	Somewhat Unlikely	Likely	Very Likely	Extremely Likely
-----------------------	----------------------	--------	----------------	---------------------

2. *thin*

a. Use this word in a sentence:

b. How likely would you use this word? Circle one.

Extremely Unlikely	Somewhat Unlikely	Likely	Very Likely	Extremely Likely
-----------------------	----------------------	--------	----------------	---------------------

c. How likely would others use this word? Circle one.

Extremely Unlikely	Somewhat Unlikely	Likely	Very Likely	Extremely Likely
-----------------------	----------------------	--------	----------------	---------------------

3. *swimmer*

a. Use this word in a sentence:

b. How likely would you use this word? Circle one.

Extremely Unlikely	Somewhat Unlikely	Likely	Very Likely	Extremely Likely
-----------------------	----------------------	--------	----------------	---------------------

c. How likely would others use this word? Circle one.

Extremely Unlikely	Somewhat Unlikely	Likely	Very Likely	Extremely Likely
-----------------------	----------------------	--------	----------------	---------------------

4. *sleepable*

a. Use this word in a sentence:

b. How likely would you use this word? Circle one.

Extremely Unlikely	Somewhat Unlikely	Likely	Very Likely	Extremely Likely
-----------------------	----------------------	--------	----------------	---------------------

c. How likely would others use this word? Circle one.

Extremely Unlikely	Somewhat Unlikely	Likely	Very Likely	Extremely Likely
-----------------------	----------------------	--------	----------------	---------------------

5. *satisfactoriness*

a. Use this word in a sentence:

b. How likely would you use this word? Circle one.

Extremely Unlikely	Somewhat Unlikely	Likely	Very Likely	Extremely Likely
-----------------------	----------------------	--------	----------------	---------------------

c. How likely would others use this word? Circle one.

Extremely Unlikely	Somewhat Unlikely	Likely	Very Likely	Extremely Likely
-----------------------	----------------------	--------	----------------	---------------------

Demographic Information

This information will be used for statistics only; it won't be used to identify any individual. You don't have to finish this questionnaire if you don't want to, but the information is important for our study. If you don't want to participate, please just keep the questionnaire. If you don't mind participating, please give answers that are as complete as possible and return your questionnaire.

A. PERSONAL INFORMATION

Age:

Sex:

Where born (state or country):

Occupation:

Spouse's Occupation:

Father's Occupation:

Where born (state or country):

Mother's Occupation:

Where born (state or country):

Your Education (circle highest level that applies):

Some High School	High School Graduate	Some College	College Graduate	Graduate School
---------------------	-------------------------	--------------	------------------	-----------------

B. LANGUAGE BACKGROUND

Please list the languages that you speak and rate your ability according to the following scale:

- A. I am a native speaker.
- B. I am not a native speaker, but native speakers usually think that I am. (near-native)
- C. I speak the language fluently, but I have an accent or sometimes say things that do not sound natural to native speakers.
- D. I can speak the language fairly well, but sometimes I have to hesitate to think of words or grammatical constructions.
- E. I know a little bit, but I have a hard time conversing normally in the language.

Language

Ability (A, B, C, D, or E)

1. English
- 2.
- 3.

C. YOUR FATHER'S LANGUAGE BACKGROUND

Please list the languages that your father speaks and rate his ability according to the following scale:

- A. He is a native speaker.
- B. He is not a native speaker, but native speakers usually think that he is. (Near-native)
- C. He speaks the language fluently, but he has an accent or sometimes says things that do not sound natural to native speakers.
- D. He can speak the language fairly well, but sometimes he has to hesitate to think of words or grammatical constructions.
- E. He knows a little bit, but he has a hard time conversing normally in the language.

LanguageAbility (A, B, C, D, or E)

1. English
- 2.
- 3.

D. YOUR MOTHER'S LANGUAGE BACKGROUND

Please list the languages that your mother speaks and rate her ability according to the following scale:

- A. She is a native speaker.
- B. She is not a native speaker, but native speakers usually think that she is. (Near-native)
- C. She speaks the language fluently, but she has an accent or sometimes says things that do not sound natural to native speakers.
- D. She can speak the language fairly well, but sometimes she has to hesitate to think of words or grammatical constructions.
- E. She knows a little bit, but she has a hard time conversing normally in the language.

LanguageAbility (A, B, C, D, or E)

1. English
- 2.
- 3.

E. LANGUAGES SPOKEN IN YOUR HOME

1. _____
a. always b. frequently (daily or nearly so) c. occasionally d. almost never e. never
2. _____
a. always b. frequently (daily or nearly so) c. occasionally d. almost never e. never
3. _____
a. always b. frequently (daily or nearly so) c. occasionally d. almost never e. never

May we contact you for help in locating other people who might be willing to complete this survey? Y / N

Your contact information (name, address, phone, e-mail):