

1 Introduction

The theory of onomasiology branches away from formalist approaches that concentrate on factors of linguistic form and centers on a cognitive/conceptual approach to compound (naming unit) interpretation (Grzenga, 2009). Dokulil (1962), who first conceived of the onomasiological theory, suggested a distinction between WORD-FORMATION and WORD-FORMEDNESS as a move away from the more analytical word-formation theories that concentrated on the byproduct of word-formation processes rather than the processes themselves (Horecký, 2007). The word-formation process, according to Dokulil, is categorized by the ONOMASIOLOGICAL BASE and MARK. The onomasiological base is the general category of objects of extra-linguistic reality from which a referent is chosen to receive a name. The specifying element in the creation of a new word became known in Dokulil's theory as the onomasiological mark. The mark narrows the general category (base) of the referent with a specifying morpheme. For example, in *steamboat*, *steam* (the mark) limits *boat* (the base) to only those boats that run on steam.

Onomasiology focuses on studying the PROCESS of word-formation as much as the PRODUCTS of the word-formation processes. Extra-linguistic knowledge, therefore, plays a crucial role in determining the relationships between constituents (mark and base) of a naming unit in this theory. It is from a speaker's understanding of the world and the need to name an object of reality (the meaning of OBJECT not being limited to tangible objects) that morphemes are combined into meaningful units for the creation of a name. Word-formation is inextricably linked with word-interpretation. The challenge, then, is that both coiner and interpreter must home in on the same meaning in order to reach a mutual understanding of a new naming unit. A coiner can assume that hearers of the coined naming unit share a similar knowledge of the world and that the hearers' will take that knowledge into account when interpreting.

The special nature of news headlines illuminates the relationship between the coiner and interpreter of novel, context-free naming units. Writers of news headlines must assume their readers will understand the topic of an article based on the condensing of much information into a few words. The generative tradition has given the label COMPOUND to these naming units; however, in an onomasiological theory, the combining of any morphemes—be they free or bound—is encompassed in NAMING UNIT creation. Separate processes such as compounding and affixation are not differentiated. Therefore, a word such as *unrest* has been formed by the same process as *steamboat*.

Štekauer's (1998: 2) approach to word-formation, based in large part on Dokulil's theory, reflects a process of "referent-concept-meaning-form." A coiner selects the REFERENT to be named, defines the CONCEPT by way of prototypical semes (MEANING) that identify the referent, and gives FORM to the name by assigning morphemes to the semes. The onomasiological step, from which the theory is named, is contained in the process from concept to meaning. This central step involves conceptualizing the referent in terms of prototypical semantic features that are eventually assigned morphemes. (A more comprehensive discussion about morpheme-to-seme assignment can be found in section 2.6.) Language users must be able to infer the meaning behind word-formations in order to understand newly coined naming units, which means there is a level at which the meaning of any novel, context-free naming unit can be predicted.

My investigation draws on the methodology for predicting the meanings of novel, context-free naming units as outlined in Štekauer (2005a). This study aims to test the validity of Štekauer's meaning prediction methods by using nominals with recursive onomasiological structures and to test the meaning predictability of these recursive naming units. To study naming units of those found in news headlines, a large body of data comprising particularly lengthy nominals, Štekauer's theory must include a robust process of recursion. Štekauer's theory of onomasiology postulates that a complete complex

naming unit consists of a ternary structure in which all semantic elements are linguistically expressed. For the ternary structure to remain intact within his theory, the theory must be compliant with the integration of recursion, a process not addressed in Štekauer's outline of onomasiology. Otherwise, the ternary structure must be compromised in order to account for nominals with more than three constituents.

A second aim of this study is to test Štekauer's assertions that certain types of naming units are more easily predictable than others (Štekauer, 2005a: 260–1). The findings of my study indicate that Štekauer's conclusions may have been preemptive. What does seem to be consistent between the two studies is the overwhelming effect of extra-linguistic knowledge in the prediction of meaning. This, more than any other factor, contributes to ambiguity within a naming unit and increases or decreases the competition among probable meanings. Structure plays a limited role in meaning predictability indicating that analyses of word-formation processes from referent to word form may be more informative about name creation than merely studying word-formedness.

2 Literature Review

The early studies on word-formation in English were characterized by a generative approach, relating the syntactic elements of sentences to the smaller constituents of word formations. Lees (1960) began with a theory in which form dominated meaning, and he was widely rebuked for disregarding crucial semantic aspects of word-formation. Marchand (1969), one of Lees' most severe critics, emphasized a notion of semantic motivation to underlie his syntagma-based theory. Levi (1978) attempted to categorize the possible meaning relationships of complex nominals and proposed that the semantic material that is lost in Lees' account of transformations can in fact be recovered. Downing (1977) rejected previous attempts to limit the number of underlying semantic relationships in compounds and presented evidence that novel, context-free forms—forms that are truly

subjected to word-formation processes—contain much more fine-grained semantic detail than had previously been accounted for. She also provided a methodology of ranking proposed meanings that has been adopted for both Štekauer’s (2005a) investigation and my own. Dokulil (1962, 1978) introduced a meaning predictability aspect to word-formation and laid the groundwork for an onomasiology-based theory. It is from Dokulil (1978) and Horecký (1983) that Štekauer (1998, 2005a) presented his theory of onomasiology and conducted four experiments to test quantitative methods of predicting meaning in novel, context-free naming units.

2.1 Lees (1960)

Lees’ (1960) transformational account of English compounding attempted to apply the framework proposed in Chomsky’s (1957) *Syntactic Structures* to the formation of nominal compounds, focusing on generation rather than interpretation. His work paved the way for generative studies of compounding for about the following two decades (ten Hacken, 2009). The basic tenets of Lees’ work rely on the assumption that compound creation is analogous to nominal phrase construction. The transformational aspect of Lees’ work involves the use of underlying kernel sentences to derive compounds. According to Lees (1960), compound structures resemble key elements of the structure of the kernel sentences. These shared structure types include: Subject-Predicate (*girlfriend*), Verb-Object (*book review*), Subject-Prepositional Object (*garden party*), etc. The nominals discussed are, therefore, “noun-like versions of sentences” rather than sentences themselves (Lees, 1960: 54).

Lees (1960) makes a distinction between two types of nominals: subject clauses (*What he wants* was obvious (Lees’ example)) and compounds. Although he distinguishes these two types, both are derived by the same method, namely that of transforming underlying kernel sentences (ten Hacken, 2009). The following is the derivation of *population growth* via the transformational process (Lees, 1960: 138; ten Hacken, 2009: 56):

- (1) a. The population grows.
b. ...growth of the population...
c. ...population growth...

Working in a pre-lexicalist era, Lees did not consider semantic aspects of compound formation. Štekauer (2005a) highlights some of Marchand's (1974) main concerns about separating semantics from compound formation. First, Štekauer argues that compounding is an act of naming, the creation of a name for an entity that exists in the world. For this reason, semantics is integral to the word-formation process (Štekauer, 2005a: 3). He also cites Marchand (1974: 298) saying that "the aim of word formation is the production of new lexical units, not just the formation of new entities on grammatical patterns" (Štekauer, 2005a: 3). Second, the arbitrariness of the kernel sentences means that many underlying sentences can produce one compound, so that *windmill* can be derived from 'the wind powers the mill,' 'the wind activates the mill,' and 'the mill is activated by the wind' (Štekauer, 2005a citing Scalise, 1984). There is no motivation for any one kernel sentence. Finally, similar underlying sentences can turn out very different compound types. The examples he uses are *wading-bird* derived from 'the bird wades' and *population growth* from 'the population grows' (Štekauer, 2005a: 3). Both have identical underlying structures but different words are selected for placement as the head constituent. Finally, Lees does not address ambiguity in compounds. A *flour mill* is a mill that grinds flour while a *windmill* is a mill powered by the wind. By analogy, a *flour mill* could be a mill powered by flour as well (ten Hacken, 2009).

The power of Lees' compound rules has been largely criticized as well (ten Hacken, 2009 and Bauer, 1978). Lees (1960: 121) himself writes that "in a great many cases it will be possible to construct on the basis of the given transformations an indefinitely large number of compounds which do not occur in any extant corpus of English." The power of Lees compounding rules means "there are too many possible ways to relate the deep structure and the surface structure" (ten

Hacken, 2009: 62), an argument Marchand (1969) makes about the lack of motivating principles in Lees' theory.

Štekauer (2005a) argues that Lees' theory falls short of providing a necessary framework for explaining how meaning can be predicted from compounds. Ultimately, the fine-grained semantics of similar compound structures (i.e., *puppy dog* 'dog which is a puppy,' *watchdog* 'dog which watches something,' *sheep dog* 'dog which herds sheep') cannot be explained using the deep structure of underlying sentences that is central to Lees' transformational theory of compounding (Štekauer, 2005a: 2). Marchand (1969: 31), emphasizing the equal standing of form, meaning, and grammatical structure within word formation, does concede that compounds can be related to sentences, and in fact, describes the meaning of his examples by using underlying sentences.

2.2 Marchand (1969)

Marchand (1969: 11) begins his discussion regarding compound formation by introducing the notion of SYNTAGMA, the relationship between the binary compound constituents he labels DETERMINANT (modifier) and DETERMINATUM (head). The importance of a semantic aspect in the study of compounds is immediately addressed. Compounding arises from a need to differentiate subclasses of general categories: *steamboat* is a boat (general category), but one particularly run by steam (Marchand, 1969: 11). The meaning of the determinatum is restricted by the determinant.

A central concept in Marchand (1969) is motivation, and it is also one of his criticisms of Lees (Kastovsky, 2005: 101). Compounds, which Marchand calls COMPOSITES, are not arbitrary names like signs and morphemes are: "A composite rests on a relationship between morphemes through which it is motivated" (Marchand, 1969: 2). The relationship between form and meaning in complex linguistic structures is motivated because form and meaning "can be interpreted semantically on the basis of the knowledge of the meanings of their constituents and some general underlying pattern" (Kastovsky, 2005:

101). There are underlying relationships that motivate the formation of new complex linguistic units.

Marchand (1969: 53–9) outlines a pattern of analysis for compound structures that consist of five factors. The first is the MORPHOLOGICAL SHAPE which is determined by whether the contents of the compound are free or bound morphemes. The second is the MORPHOLOGICAL STRUCTURE which determines in what sequence the determinant and determinatum will be ordered (Kastovsky, 2005: 105). The third is CONTENT AT THE LEVEL OF GRAMMATICAL DEEP STRUCTURE which determines what syntactic functions the constituents of the compound perform at the level of the kernel sentence. The fourth factor is TYPE OF REFERENCE which includes topicalization of certain syntactic structures in the kernel sentence. It determines which word in the sentence becomes the determinant and which becomes the determinatum. Finally, the fifth factor is CONTENT AT THE MORPHOLOGICAL LEVEL. At this level, relevant semantic information is added to the compound that cannot be derived from the syntactic relationship alone. For example, a *baker* is not just someone who bakes but the meaning of OCCUPATION is added to the underlying semantics. This is a critical step in the lexicalization or idiomization process (Kastovsky, 2005: 106).

Marchand (1969) provided one of the most comprehensive accounts of compound structures ever attempted. His proposed word-formation processes are analyzed through the use of existing compounds rather than novel, context free compounds. The future research of Downing and of those from within the onomasiology perspective look at a new body of data: non-lexicalized compounds.

2.3 Levi (1978)

A major flaw in Lees (1960) is the power of his transformational rules which delete important semantic material that is unable to be accounted for in the meanings of the final nominals (Downing, 1977: 811). In response to this problem, Levi proposes what she calls RECOVERABLY DELETABLE PREDICATES (RDPs),

predicates that can be semantically recovered after the transformation takes place.

According to Levi, complex nominals (which include compounds, etc.) are formed in two ways: the first is by DELETION of the predicate and the second is by NOMINALIZATION of the predicate. For example, *apple cake* ('cake MADE FROM apples') is an example where the predicate has been deleted from the underlying sentence, whereas in *film producer* ('someone who PRODUCES films), the predicate has been nominalized. Unlike Lees' theory, for the nominals formed by deletion, Levi asserts that certain predicates are recoverable in the final semantics of the nominal. Only the members of the set of predicates outlined by Levi may be deleted in the creation of complex nominals by deletion (Levi, 1978: 50). Levi wanted to consolidate the number of underlying semantic relationships for complex nominals in order to reduce ambiguity since complex nominals are obviously not wildly idiosyncratic. Theoretically, this outcome is desirable because it consolidates the number of meaning relationships. Her list of RDPs includes (Levi, 1978: 76–7):

- (2) 1. Cause
2. Have
3. Make
4. Use
5. Be
6. In
7. For
8. From
9. About

Downing (1977) argues that Levi is limiting the possible underlying semantic relationships too much by limiting the potential ambiguity in compounds to the number of RDPs. However, the RDPs also suggests that a given nominal is at least in nine ways semantically ambiguous, so it is up to "judicious exploitation of semantic clues, lexicalization, and pragmatic or extra-linguistic (encyclopaedic) knowledge" to help narrow down the meaning (Levi, 1978: 158). Downing further questions how it is possible to know which

underlying predicate goes with which compound. Some compounds can have several relevant RDPs. For example, *floral wreath* could just as easily be categorized under MAKE as it could BE or HAVE (Downing, 1977: 814). Furthermore, some RDPs can be relevant to complex nominals derived by nominalization rather than deletion. These include examples like *solar generator* (USE) and *viral infection* (CAUSE) (Downing, 1977: 814). Downing attempts to lessen the restrictions on meaning in complex nominals with her series of experiments on context-free compound.

2.4 Downing (1977)

Downing introduced the method of ranking proposed meanings of novel, context-free compounds, similar to the method used in Štekauer (2005a) and the current study. Downing frees herself from the former attempts at characterizing compounds by studying non-lexicalized forms (Downing, 1977). In her experiment, participants completed three tasks: context-free interpretation of novel compounds, a naming task where participants created a name for an object in a drawing, and a ranking task in which participants ranked the appropriateness of certain meanings.

She uses these tasks to illustrate two problems with previous studies of compounding. First, previous studies have not adequately considered the effects of lexicalization; second, they had not distinguished the function of compounds and full sentences (Downing, 1977: 819). Lexicalization often adds additional semantic information or subtracts existing semantic information. *Marshmallow* (Downing's example) is named for the plant it is derived from, but through lexicalization that meaning has been lost. The functional difference between compounds and full sentences, according to Downing, is that sentences 'assert' while compounds 'refer'. Compounds denote categories— they give a name rather than describe. "Sentential paraphrases are incomplete, for they cannot convey the fact that the speaker considers the referent worthy of a name rather than a description" (Downing, 1977: 824).

Previous studies also failed to capture the extent of the semantic richness of compounds. While there do exist some frequently occurring general underlying relationships such as location, time, composition, purpose, etc., Downing argues that the number of possible compounding relationships is not finite (Downing, 1977: 828). Former research by Levi (1975), Li (1971), and others sought to reduce the possible underlying semantic relationships of compounds to a finite list. Downing (1977) argues that these attempts eliminate the more fine-grained semantic detail compounds can encompass, as evidenced by the meanings her participants proposed. A compound like *frog-slime* was not merely defined as ‘slime from a frog’ but was more acutely articulated as “the slime that frogs exude to keep from dehydrating” (Downing, 1977: 826). Her results revealed that many participants included a real-world context in their definitions, a point that undergirds her assertion that compounds can carry in a rich and deep semantic structure. Rejecting many attempts made in former studies to address the underlying semantic relationships of compounds, Downing drew attention to a new body of data (non-lexicalized forms) that revealed many more and more complicated semantic structures than had previously been described.

2.5 Dokulil (1978)

Dokulil’s (1978) work marks the first attempt to associate meaning predictability with the constituent parts of a novel, context-free naming unit (Štekauer, 2005a: 38). Moreover, he writes from an onomasiological perspective and introduces the notions of the onomasiological CATEGORY, MARK, and BASE. The ONOMASIOLOGICAL CATEGORY is a way of defining how the subparts of a naming unit are ordered and expressed in language (Grzenga, 2009: 219). These subparts consist of the ONOMASIOLOGICAL BASE and the ONOMASIOLOGICAL MARK, analogous to Marchand’s DETERMINATUM and DETERMINANT. The base represents a more general classification of the notion to be named. The mark then serves as a specifying structure. It restricts the base according to what differentiates it from other

members of its class. Within the mark there are two elements: the DETERMINING and the DETERMINED constituents, which may not necessarily be expressed. Dokulil (1962) introduces four basic categories from which the onomasiological structure is made: SUBSTANCE, ACTION, QUALITY, and CIRCUMSTANCE. The relationship between the base and the determining element of the mark is what comprises the overall *concept*. Štekauer calls these the POLAR MEMBERS of the structure. “A concept of the category [for the base] of SUBSTANCE is determined by its relation to a concept of the category [for the mark] of (a) SUBSTANCE (*policeman*), (b) QUALITY (*blackberry*), (c) ACTION (*teacher*), (d) CONCOMITANT CIRCUMSTANCE (*evening paper*)” (Štekauer, 2005b: 210).

Dokulil does not elaborate on compounds because in his native Czech, compounding is a minor phenomenon (Grzenga, 2009: 221). However, unlike Štekauer, he does differentiate word formation processes such as compounding, affixation, conversion, etc. For Štekauer (2005a), the onomasiological terms are not limited to compounding but extend to all word formation processes. For example, in ‘blackberry’ *black* is the specifying mark of the base *berry*, where *berry* is “common to the whole conceptual group of various *berries*” (Štekauer, 2005b: 210). In the derivation-type process that creates *singer*, the mark is *sing* while the base is agentive *-er* (Grzenga, 2009: 220). Dokulil (1962) informed much of the future work conducted within the onomasiological framework and provided the basis for Štekauer’s work in word-formation studies.

2.6 Štekauer (2005a)

Štekauer (1998) outlines a theory of word-formation within the framework of onomasiology, basing much of his theory on Dokulil (1962). He begins with a definition of word-formation as “the process of coining new naming units in accordance with the naming needs of a speech community” (Štekauer, 1998: 8). Words are not formed in isolation from extra-linguistic reality. The basis for word-formation then is not the word but the concept behind the word. Cognitive and

conceptual aspects, therefore, play a much larger role in onomasiology than in traditional generative theories of word-formation. Štekauer's onomasiological theory focuses on a triadic relationship between "extra-linguistic reality (object to be named), speech-community (coiner) and word formation..." (Štekauer, 2005a: 44).

The onomasiological model, first established by Horecký (1983), comprises a formation process that begins with an extra-linguistic reality and results in a complex naming unit. The model consists of six levels (Štekauer, 1998: 8):

- (3) 1. Extra-linguistic reality
2. Conceptual level
3. Semantic level
4. Onomasiological level
5. Onomatological level
6. Phonological level

EXTRA-LINGUISTIC REALITY is the notion that needs to be named. The CONCEPTUAL LEVEL, then, involves generalizing the extra-linguistic reality by classifying it into one of Dokulil's (1962) four CONCEPTUAL CATEGORIES: SUBSTANCE, ACTION, QUALITY, and CONCOMITANT CIRCUMSTANCE (Štekauer, 2005a). The SEMANTIC LEVEL assigns semes, or prototypical features, to the concept. The ONOMASIOLOGICAL LEVEL determines the structure of the word with three components: the determining and determined constituent of the mark and the base. "At the onomasiological level proper, one of the semes is selected to function as an ONOMASIOLOGY BASE denoting a class, gender, species, etc., to which the object belongs, and one of them is selected to function as an ONOMASIOLOGY MARK which specifies the base" (Štekauer, 2005a: 9).

The ONOMATOLOGICAL LEVEL is the level at which semes are assigned morphemes. Štekauer calls this the MORPHEME-TO-SEME ASSIGNMENT PRINCIPLE (MSAP). It is from this assignment that Štekauer finds justification in dispensing with the formal divisions in word formation such as compounding, affixation, back-formation, etc. (Štekauer, 2005a: 51):

In addition to reducing the number of classificatory criteria to a single criterion (i.e., MSAP), the proposed approach makes it possible to reveal the naming strategies, and to show, for example, that formally different naming units may result from one and the same naming strategy.

For example, *-ist*, *-ian*, *-er*, *man*, *-ant*, etc. can all represent AGENT. Thus, the process of assigning morphemes to semes is the same in any act of word-formation, but the number of morphemes that can be assigned to a given seme varies. Finally, the PHONOLOGICAL LEVEL is where stress pattern and pronunciation are assigned.

My study is based on Štekauer (2005a) in which he conducts a set of experiments testing to what extent the meaning of context-free naming units can be predicted. With the results of his studies, Štekauer shows that the number of possible meanings of a naming unit is not infinite as Downing (1977) asserted because all of the naming units have a clear dominant reading which will rule out most other readings for people interpreting novel naming units. Extra-linguistic factors also play a large part in narrowing down the potential meanings. The following section will identify crucial tenets of onomasiology from Štekauer's perspective, define important concepts within the theory, and lay a framework for my investigation.

3 The Theory of Onomasiology

3.1 Compounding and Onomasiology

Defining the notion of COMPOUND has been a difficult task for linguists, and many of them disagree over the status of compounding as a word-formation process. Throughout the history of word-formation studies, the definition of a compound has been argued for on morphological, phonological, semantic, syntactic, and cognitive grounds, and there is little agreement. Some linguists concede that no satisfactory definition of a compound really exists (Bauer, 1978; Lieber and Štekauer, 2009; Plag, 2003; etc.). This has led some linguists to accept the limitations of the definition and proceed with research anyway, while others have disposed of notions of

compounding as a distinct process and subcategorize it under other word-formation process (Marchand, 1969; Levi, 1978; Štekauer, 1998).

Marchand (1969) argues for two process types, expansion and derivation, where compounding is not a separate word-formation process. The distinguishing features of expansion and derivation are whether or not the determinatum (head) is an independent or bound morpheme (respectively). Thus, compounding and prefixation fall into the expansion category while suffixed words fall into the derivation category (Lieber and Štekauer, 2009). Bauer (1978: 54), on the other hand, treads lightly in his attempt to define a compound, saying that the tendencies, rather than rules, of compounding are that compounds are “made up of two words (lexemes) acting as a single word (lexeme),” and that compounds seem to function as syntactically and semantically isolated units. Levi (1978), however, argues that no adequate criteria exist to isolate compounding as a distinct word-formation process from the construction of other complex nominals.

Štekauer (1998, 2005a) and other researchers within the onomasiological perspective argue for a more unified theory of word-formation that relies on semantics and cognition, grouping many word-formation processes that are differentiated in generative grammar into one process. In his approach, Štekauer seeks to unify all word-formation processes under onomasiology. Therefore, he does not talk of compounds but rather NAMING UNITS, a term that will be adopted for the rest of this study. The term NAMING UNIT presupposes a central aspect of Štekauer’s theory of word-formation: that word-formation is word-CREATION, and formation processes create new words which invariably name objects of extra-linguistic reality. These words can arise out of the spontaneous need of a speaker at a given moment never to be used again, or they can become a lexicalized part of the language.

The WORD-FORMATION COMPONENT of onomasiological theory comprises Horecký’s (1983) model (outlined in section 2.6) which draws elements from the LEXICAL COMPONENT of a speaker’s grammar to create new naming units. These units can then be added to the

Lexical Component once they are established in the language. The Word-Formation Component is then completely productive, as it uses pre-lexicalized linguistic units in the formation of new naming units through its interactive relationship with the Lexical Component. (See section 5.2 for more discussion on the two components.)

As mentioned above, the onomasiological level of the word-formation component is the level at which it is decided what structural components will be assigned morphemes. The onomasiological category, or structure, is a ternary structure consisting of the onomasiological mark—divided into determining and determined constituents—and the onomasiological base. According to Štekauer (2001: 21-2):

It should be emphasized...that the [base] always refers to a class of objects, a genus, etc. Consequently, rather than identifying [the] head either positionally or morphologically (a particular morpheme of a naming unit) the proposed approach shifts the criterion of headedness to the extralinguistic [*sic*] level, in particular, to the conceptual level of coining new naming units. By implication, head can be a suffix, a prefix, or a word-formation base.

The following is a representation of the complex word structure:

(4) Determining constituent – determined constituent – base

While the base does typically appear on the right, it is not limited to this position. A word such as *untrue* consists of a BASE-MARK structure.

At the onomatological level, the structure is assigned morphemes by way of the MORPHEME-TO-SEME ASSIGNMENT PRINCIPLE (or MSAP). If we use *language teacher* as an example, the morphemes would fit into their respective slots in the following way:

(5)	determining	–	determined	–	base
Onomas. level:	Object	–	Action	–	Agent
MSAP:	<i>language</i>	–	<i>teach</i>	–	<i>er</i>

Language is the determining constituent (determines what is being taught), *teach* becomes the determined constituent (the actional element), and *-er* denotes the base (the Agent). The determining constituent can sometimes contain the specifying and specified elements. For example, *transformational grammarian* is parsed as follows:

- (6)
- | | | | | | |
|----------------|---------------------------------|---|------------|---|------------|
| | determining | – | determined | – | base |
| Onomas. level: | Object | – | (Action) | – | Agent |
| MSAP: | <i>transformational grammar</i> | – | ∅ | – | <i>ian</i> |

Transformational grammar is what the agent (*-ian*) studies; however, the actional element of ‘study’ is not linguistically expressed (denoted by ∅). *Transformational* represents the specifying element and *grammar* is the specified element. In other words, *transformational* restricts the meaning of *grammar* but as a whole they function as the determining constituent, the object of study.

Štekauer’s onomasiological types (hereafter OT) are categorized based on the constituents that are linguistically expressed (are assigned morphemes) for a given word formation. Certain word formations express all three elements, some only two elements. The five types are listed below with relevant examples (DG = determining constituent; DD = determined constituent) (Štekauer, 2005a: 52):

Type 1: All three constituents (the determining constituent, the determined constituent, and the base) are assigned morphemes. Examples include *language teacher* ($\text{language}_{\text{DG}} - \text{teach}_{\text{DD}} - \text{er}_{\text{BASE}}$), *weather warning* ($\text{weather}_{\text{DG}} - \text{warn}_{\text{DD}} - \text{ing}_{\text{BASE}}$), and *spring lock-nut* ($\text{spring}_{\text{DG}} - \text{lock}_{\text{DD}} - \text{nut}_{\text{BASE}}$).

Type 2: Only the determined constituent of the mark and the base are assigned morphemes. This type can be extended to type 1 by adding a determining constituent. Examples include *teacher* ($\emptyset - \text{teach}_{\text{DD}} - \text{er}_{\text{BASE}}$) and *stop-watch* ($\emptyset - \text{stop}_{\text{DD}} - \text{watch}_{\text{BASE}}$).

Type 3: Only the determining constituent of the mark and the base are expressed. The actional element (determined constituent) is missing. Examples include *honey bee* ($\text{honey}_{\text{DG}} - \emptyset - \text{bee}_{\text{BASE}}$) and *transformational grammarian* ($\text{transformational grammar}_{\text{DG}} - \emptyset - \text{ian}_{\text{BASE}}$).

Type 4: The mark cannot be broken down into determining and determined constituents, although the mark and the base are still expressed. Examples include *blue-eyed* ($\text{blue eye}_{\text{MARK}} - \text{ed}_{\text{BASE}}$) and *untruth* ($\text{untrue}_{\text{MARK}} - \text{th}_{\text{BASE}}$).

Type 5: This type accounts for the phenomenon known as conversion. It is not relevant for the present study, however, and will not be discussed.

3.2 Predictability

Štekauer (2005a) investigates the possibility that the meaning of novel, context-free naming units can be predicted. This is based on the assumption put forth by van Lint (1982: 136 echoed by Štekauer, 2005a: 43) that “interpretation presupposes a production system which only produces interpretable strings.” It is possible to predict the meaning of a naming unit because the naming unit was created with the intention that it can be interpreted. Meaning predictability presumes that speakers and listeners have world knowledge about how objects relate to one another and that the construction of naming units recognizes these relationships. Therefore, naming units are predictable insofar as speakers and listeners recognized these relationships in the world and understand how these relationships can be represented on a linguistic level.

When listeners interpret a new naming unit, the range of possible meanings is limited to the “range of general semantic relationships between the constituents” of the naming unit (Štekauer, 2005a: 71). The listener selects the most probable relationship as the most probable application of the MORPHEME-TO-SEME ASSIGNMENT

PRINCIPLE during the word-formation process (Štekauer, 2005a: 71-2). Semes represent the prototypical features of an object and, presumably, most speakers and listeners will recognize these features. Štekauer (2005a: 71) explains with regard to listeners' interpretations of novel forms:

Based on mutual relations between these semantic constituents, it is possible to predict the possible onomasiological structure, or in other words, to employ one's competence with regard to the [Word-Formation] Rules and patterns to predict the particular lexical meaning of a novel complex naming unit on the basis of the identification of the *general word formation meaning* underlying the coinage.

Štekauer has laid out a quantitative method for calculating the predictability of meanings for novel, context-free naming units based on participants' rankings of meanings they propose. Since what is an acceptable and unacceptable meaning varies from speaker to speaker, the predictability is a cline (Štekauer, 2005a: 94). Therefore, Štekauer uses a scale from 1 to 10 for judging the acceptability of a meaning (I employ this method as well).

3.3 Defining PREDICTABILITY RATE, PREDICTABILITY RATE GAP, and OBJECTIFIED PREDICTABILITY RATE

The rate of predictability—the probability that a certain meaning will be selected as the most probable representation of the underlying semantic relationship— of a given naming unit can be quantified through the formula given by Štekauer (2005a). The PREDICTABILITY RATE (PR) is calculated in the following way (Štekauer, 2005a: 94-5):

$$(7) \quad PR = \frac{r}{R_{MAX}} \times \frac{p}{P_{MAX}}$$

where

r = the number of informants identifying a particular meaning as acceptable

R_{MAX} = the total number of informants
 p = the sum total of points assigned to a given meaning by all informants (on a scale of 1 to 10, where 10 stands for the highest acceptability of the meaning)
 P_{MAX} = the maximum possible number of points assignable by all informants

For example, if the meaning ‘someone who studies transformational grammar’ for *transformational grammarian* was proposed by 16/20 participants (a frequency of 0.8), and the total scores from these participants equals 168 out of 200 possible points, then the PR is 0.672.

$$0.672 = (16/20) \times (168/200)$$

All of the proposed meanings for *transformational grammarian* can be ranked in this way. The PR can also be weighed against the scores of other naming units to determine the OBJECTIFIED PREDICTABILITY RATE (OPR).

The OPR indicates how predictable a meaning for a given naming unit is in comparison to another naming unit. For example, if the highest rated meanings for two naming units X and Y have identical PRs, it is not enough to say these meanings are equally predictable for their respective naming units. The PREDICTABILITY RATE GAP between the meaning with the highest (TOP) score and the meaning with the second highest score, etc. determines which of the naming units X and Y is more predictable. Štekauer’s example is as follows (2005a: 96):

(8)	Naming unit X		Naming unit Y	
		PR		PR
	X ₁	0.486	Y ₁	0.486
	X ₂	0.194	Y ₂	0.362
	X ₃	0.088	Y ₃	0.088

Where X₁, X₂, X₃ and Y₁, Y₂, Y₃ are the three highest ranked meanings for each naming unit. Plugging these predictability rates into the OPR formula below results in the OPR_X of 0.633 and the OPR_Y of 0.519:

$$(9) \quad \text{OPR} = \frac{\text{PR}^{\text{TOP}}}{\text{PR}^{\text{TOP}} + \text{PR}^{\text{TOP-1}} + \text{PR}^{\text{TOP-2}}}$$

The PREDICTABILITY RATE GAP between the highest predicted readings and the subsequent readings is higher for naming unit X (0.633) than Y (0.519), meaning the most predictable reading for X is MORE predictable than the most predictable reading for Y. It follows, then, that “the higher the number of competing predictable readings...the lower the OPR” (Štekauer, 2005: 97). These two formulas will be used to determine the PRs and OPRs of the naming units in my study.

4 Research Questions and Hypotheses

4.1 Research Questions

The foundation for this study is based on three research questions that the data from the study will answer.

1. *What is the role of recursion in a theory of onomasiology?* As will become further clarified in section 5.2, for naming units such as *childcare benefit cuts warning*, recursion must be a feature of word-formation that fits within the onomasiological theory. Certain structures, such as the one named above, do not fit neatly into a tripartite word structure without recursion. How can *childcare benefit cuts warning* fit into a ternary word structure when it contains four distinct words (*childcare* even being a compound, in generative terms) unless more than one word can act together as a single constituent?
2. *How well can Štekauer’s theory and methods of predictability account for actual recursive naming units?* Again, as explained in section 5.2, recursion fits quite well into onomasiological theory. Furthermore, it seems that all four onomasiological types can be embedded within both OT1 and OT3 structures. As OT1 and OT3 were the most common overall structure

types found, it is these two types that are the focus of my study. This does not exclude the possibility of an overall OT2 or OT4 structure, but merely limits the scope of my study. As Štekauer (2005a) has shown in his meaning predictability studies, his methods for quantifying predictability work quite well. An analysis of his methods within the scope of recursive naming units will be made using the data from my study.

3. *To what extent does structure affect meaning predictability?*

Štekauer (2005a: 253) concludes that onomasiological types 1, 2, and 4 are PREDICTABILITY BOOSTING structures while onomasiological types 3 and 5 are PREDICTABILITY REDUCING structures. Since all the naming units in my study have an overall structure of either OT1 or OT3, contrasting structures in their predictability, the data should indicate whether Štekauer's claims can account for larger recursive structures as well.

4.2 Hypotheses

My five hypotheses are based on the research questions above regarding the interaction of structure and predictability:

1. Extra-linguistic knowledge about the components (i.e., an action requiring a HUMAN agent) will create a higher OPR even if there are many possible semantic combinations. Therefore, extra-linguistic knowledge will affect the OPR when the structure does not.
2. When a constituent can be interpreted as a modifier, the OPR will decrease due to greater competition from other meanings. For example, *primary pupil* could be 'a pupil who attends primary school' or 'a very important pupil' depending on the interpretation of *primary*.
3. Data from this study will corroborate Štekauer's claims about onomasiological type and predictability, namely that the overall OT1 structures will have higher OPR scores than those naming

units with an OT3 structure due to the presence of the determined constituent in OT1. The structure of the recursive element will affect the predictability if the recursive structure contains an expressed actional element.

4. The actional nature of a determining constituent or base will 'fill in' for the missing actional seme in OT3 structures creating a higher OPR than for other OT3 structures (Štekauer, 2005a: 214).
5. The PRs for naming units with recursive structures will be higher than those found for analogous structures in Štekauer's data because more semes are linguistically expressed which allows for additional specification.

5 The Study

5.1 Methodology

5.1.1 Participant Selection

The participants for this investigation voluntarily responded to an e-mail I sent to friends and family in America. The email consisted of the instructions in (10) and the survey. A copy of the e-mail can be found in Appendix A. The first twenty completed surveys were chosen for the study. The survey began with an informed consent section followed by instructions with an example and then the eight naming units listed in (11). Participants were instructed with the following paragraph and example:

- (10) The following is a list of compound words found in online news headlines. Propose as many meanings as you can think of for each of the compound words. Not every meaning you propose may be straightforward or make much sense, especially if you are fairly sure of the intended meaning. List these anyway as potential meanings. Assign a score (from 1 to 10) to each of the meanings you propose, with a 10 indicating the highest probability that the meaning proposed would occur as the intended meaning (the most predictable), and assign a 1 to the meaning with the minimum chance of occurring as the intended meaning.

For example:

Water-mill

–a mill driven by water	10
–a mill near water	4
–a mill made of water	2
–a mill the color of water	1
–a mill that grinds water	1
(like a pepper mill grinds pepper)	

The instructions are modeled on Štekauer's (2005a: 100) instructions but some rewording for the sake of clarity has been made. The example is taken from Štekauer (2005a: 101) with the exception of '*a mill that grinds water*' which was added to encourage participants to think of more than just the obvious senses of words.

5.1.2 Naming Unit Selection

The critical distinction between novel, context-free naming units and lexicalized ones is that language users must use their own linguistic knowledge and extra-linguistic knowledge in order to interpret novel, context-free naming units. Lexicalized units have received a meaning agreed upon by society that may or may not be related to their structure. Furthermore, syntactic structures such as phrases are interpreted using a very different process than that used for interpreting new naming units (Jackendoff, 2009: 110):

Standard syntax pretty well determines the meaning of a novel phrase from the meanings of the words and the principles for mapping syntactic structure into combinatorial relations among constituent meanings. In the case of a novel compound, though, the general principles yield only a vast range of possible meanings. The language user must home in on the intended meaning of a novel compound by making use of (a) the semantic details of the constituent words and (b) the discourse and extralinguistic [*sic*] context.

Although combining the respective meanings of constituents within a novel naming unit is a critical part of the interpretation process, extra-linguistic knowledge is always a factor.

From a historical perspective, there is no reason to presume that words created centuries ago should have come about by way of the same derivational processes as new words today, “that is, a synchronic grammar may no longer be able to produce compounds which were first generated several hundred years ago and have, as a result of the rules operating at that date, become fixtures in the language” (Bauer, 1978: 74-5). Thus, to get an accurate picture of the words being produced today, synchronic word-formation rules—those that also produce today’s news headlines—should be the focus of study. Therefore, lexicalized naming units alone—or those words that came about through diachronic processes—will not suffice in a study of the meaning prediction of newly formed words. However, lexicalized units are used in the creation of new naming units. Depending on the reading, in *wedding day rehearsal* and *primary pupil booklist*, *wedding day* and *primary pupil* may very well be lexicalized for some people. If that is the case, the most probable reading will feature a meaning in which *wedding day* and *primary pupil* are seen as inseparable units.

News headline editors are coiners of new naming units, and they rely on the assumption that their naming unit will be interpretable to a reader or listener. News headlines provide a large body of data in which naming units are continuously coined for the purpose of attracting readers’ attention to the content of the article. Their purpose is not so far removed from any other coiner’s purpose: to create a naming unit with one meaning in mind which is to be interpreted by anyone who reads or hears it.

Of course, with regard to news headlines, one must recognize the particular usage of these naming units as units which fit a maximum amount of information in a minimum amount of space. Bauer (1978: 15-6) cautions against using headlines because they are “extreme examples” of compound creation: “In many cases one suspects that rules for compounding are relaxed in newspaper headlines just so that a ‘new, striking expression’ which takes up a minimum of space can be created.” He calls headlines “misleading” (Bauer, 1978: 16). The nature and purpose of headlines is something to

be aware of, but from an onomasiological perspective, the creation of these naming units abides by word-formation rules. In phrase creation, news headlines are relaxed, often leaving out certain verbs that can be implied. However, these extreme cases are not limited to headlines (Bauer, 1978). A friend informed me of two new services offered at her retail store officially titled *complimentary self-gift-packaging station* and *upcoming in-store registry completion event*. One only needs to look at these or the names of various government committees to see that ‘extreme cases’ are far less extreme and far more common than Bauer asserts.

It is important to address the question of why to use context-free naming units when the creation of them occurs within some context. Štekauer (2005a: 56) notes that “context-free interpretation admits generalisation, this being a crucial condition for the meaning-prediction process.” Without a contextual environment, Štekauer argues, the linguistic and cognitive effects on predictability can more easily be measured. Naming units like those found in news headlines, while sometimes tied to ongoing events, can often be free of context. Unlike Štekauer’s studies in which he creates context-free naming units, my study considers a real-life situation in which language users encounter context-free naming units. Therefore, I have selected news headlines from various news websites for the naming units used in my survey of predictability.

5.2 Recursive Structures

News headlines are rarely one to two constituents in length. Often, when not structured as phrases, headlines consist of several related nouns that create a multi-constituent naming unit. In onomasiology, the number of words in a naming unit is not relevant. It is the constituents represented by morphemes, either bound or free, that are relevant to the structure. In this theory, naming unit structure is limited to three basic constituents (as noted above, the determining and determined constituents of the mark and the base). In order for naming units of headline-length to be satisfactorily analyzed, recursion must be

a robust process in the onomasiological account of word-formation. Recursion is the basis from which the naming units for my study were selected. The naming units listed in (11) were taken from various online news websites with the exception of *growth inhibitor*, which was of my invention (See Appendix B for the sources):

- (11)
1. Hand washing warning
 2. Wedding day rehearsal
 3. Childcare benefit cuts warning
 4. Growth inhibitor
 5. Snowstorm face protectors
 6. Employee sick absence
 7. Primary pupil booklist
 8. Arab world unrest

In onomasiology, the base typically appears on the right with some exceptions including the negating element *un-*. By deconstructing the above words, the necessity of recursion in a theory of onomasiology will become apparent. Although Štekauer makes no explicit mention of recursion, Grzenga (2002: 10) implies that recursion is incorporated into Štekauer's theory:

Let us now have a look at the term *brimstone butterfly*. Here we can't assume a typical three-fold distinction *brimstone-butter-fly*, with *brimstone* being the determining constituent and *butter* the determined constituent. It is rather the case that *brimstone* specifies *butterfly* as a whole. In this case, it only makes sense to assume that *butterfly* is the onomasiological base and *brimstone* the onomasiological mark. This already seems to be covered by Štekauer's model, but it seems important to me to show the difference between "bi-partite" compounds and "pluri-partite" compounds.

Based on Grzenga's interpretation of *brimstone butterfly* it would seem that a complex structure can serve as any of the constituents in another complex structure. For the nominals found in headlines and elsewhere, it seems necessary to postulate the recursion of Štekauer's onomasiological types. As Grzenga does not seem to think this is out of line with Štekauer's thinking, the naming units for my study will be analyzed according to their recursive structures.

From a theoretical standpoint, Štekauer's model of the Word-Formation and Lexical Components (Fig. 1) allows for cyclicity. It is not enough to say that the cognitive process selects lexicalized elements from the Lexical Component to be assigned to semes at the same level because units such as *childcare benefit cuts* or *face protectors* are surely not lexicalized. Words must be formed and subsequently temporarily stored in the lexicon so they can be added to more complex formations. Štekauer's structure of the Word-Formation and Lexical Components allows for an interconnected relationship. The Lexical Component feeds the Word-Formation component and the Word-Formation Component feeds the Lexical Component. The Lexical Component is tied to the Syntactic Component for insertion of the new naming units into syntactic structures. The Lexical Component/Word-Formation Component interface must allow newly formed naming units to be drawn from the Lexical Component where they are temporarily stored for the creation of a more complex, recursive unit. Therefore, the word-formation process in onomasiology must be cyclical to allow for complex structures to be placed within a larger overall structure.

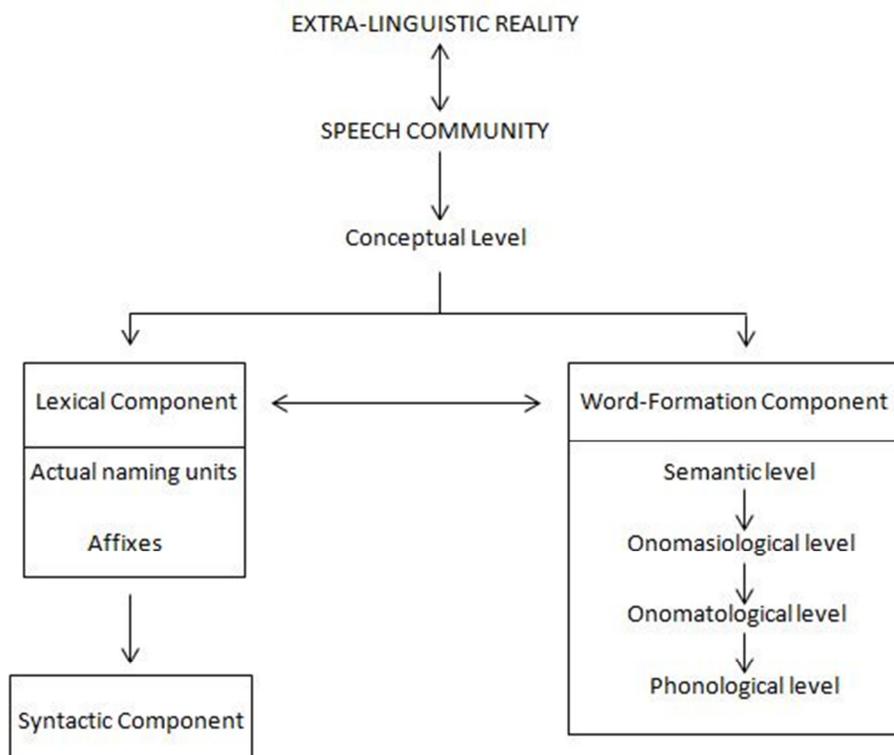


Figure 1 Word Formation and Lexical Components (Štekauer: 2005a:45)

Bisetto (2010: 14) defines recursion as follows: “A linguistic entity is recursive when it has a complex structure that can be decomposed into two or more entities of the same type.” For the naming units in this study, a naming unit of any onomasiological type can serve as a constituent within an overall structure of the same or a different onomasiological type. Each naming unit was chosen with a specific structure in mind which was based on my reading of the naming unit. For the most part, my intuitions matched the data, as will be apparent in the following sections. Naming units (1)-(4) are characterized by a final OT1 structure, with all elements (determining, determined and base) assigned morphemes. Naming units (5)-(8) are characterized by an overall OT3 structure, where the determined constituent (actional element) is not assigned a morpheme. It was found that many news headlines that consisted of recursive structures had an overall structure of either OT1 or OT3. No naming units with an overall OT2 or OT4 structure were found in the search through news headlines; therefore, four OT1 structures and four OT3 structures were selected with various embedded structures. Naming units (1) – (4) were selected for their embedded structures of OTs 1, 2, 3, and 4, respectively, and (5) – (8) were also selected for their embedded structures of OTs 1, 2, 3, and 4, respectively. It is important to note that some of the above naming units contain more than one recursive structure. For example, *snowstorm face protector* consists of *face protector* which is an OT1 structure and *snowstorm* which is an OT3 structure. *Wedding day rehearsal* in fact contains three OT2 structures: *wedding*, *wedding day*, and *rehearsal*.

5.3 Data and Analysis

The following sections consist of explanations of the structure of each naming unit followed by the results from the study for the given naming unit and an analysis of the data. Many naming units received a meaning in which one of the constituents was interpreted as a verb. In the case of *hand washing warning*, many participants listed ‘a hand washing a warning.’ Proposed meanings such as this were

ignored because they give a phrasal, and therefore syntactic, reading to the naming units. Given that most of these naming units were selected from news headlines, there is no reason that a phrasal reading could not be possible; however, for the purpose of testing the meaning predictability of NAMING UNITS rather than phrases, verbal interpretations were ignored.

5.3.1 Hand washing warning (OT1 within OT1)

Structure

The overall structure of this naming unit is type 1 where *warning* comprises the determined constituent of the mark and the base. With the absence of a determining constituent, *warning* is an OT2 structure, where *-ing* is the base and *warn* is the determined constituent of the mark:

(12)		determining	–	determined	–	base
	Onomas. level:			Action		Process
	MSAP:	∅	–	warn	–	ing

An OT2 can be turned into an OT1 by adding a determining constituent. In the case of *hand washing warning*, *hand washing* fills the determining constituent slot to create the overall OT1 structure:

(13)		determining	–	determined	–	base
	Onomas. level:	Object		Action		Process
	MSAP:	hand washing	–	warn	–	ing

Even though it is placed in the position of determining constituent, *hand washing* itself is an OT1 structure:

(14)		determining	–	determined	–	base
	Onomas. level:	Object		Action		Process
	MSAP:	hand	–	wash	–	ing

Hand washing as a unit then acts as the determining constituent of *warning* making *hand washing warning* a complete OT1 structure which encompasses a recursive OT1 structure within it.

Ranking

The top three meanings as proposed by the participants have been listed below along with the frequency of occurrence (FO), the Predictability Rate of the given meaning, and the Objectified Predictability Rate of the top meaning (Štekauer, 2005a). The FO indicates how many participants proposed the given meaning. The ranking for these meanings are based on PR score with the meaning with the highest PR ranked as 1. Only the top meaning is given an OPR since it is the top—and most probable—meaning that is used to compare the predictability of one naming unit to others.

1. *'a warning to wash your hands'*

FO: 16/20

PR: 0.562

OPR: 0.576

2. *'a warning not to wash your hands'*

FO: 15/20

PR: 0.341

3. *'a warning that you should wash items by hand'*

FO: 7/20

PR: 0.072

Comments

The relatively low ranking of the OPR for meaning (1) is due to high competition from meaning (2). Most participants differentiated between a warning TO wash hands and a warning NOT TO wash hands, as is evident by the frequency of occurrence for both meanings. A meaning of 'a warning about washing hands' seemed to be too general a definition, and two-thirds of participants listed both meaning (1) and meaning (2) to specify the meaning further. It was fairly clear to participants, however, that someone might be warned TO wash his or her hands more often than NOT TO wash his or her hands. Since there was no linguistic marker for the distinction between the two meanings, the PR Gap between the two is the lowest of all naming units.

In terms of onomasiological type, the top three meanings receive the OT1 within an OT1 recursive structure interpretation.

However, in meanings (1) and (2), *hand washing* was interpreted with a reflexive sense of washing one's own hands, while, in meaning (3), it was interpreted with the hands being the instrument used to wash something else rather than themselves. Obviously, *warning* did not pose an interpretation problem but rather the internal semantics of *hand washing* did. In meanings (1) and (2), *hand* received an OBJECT interpretation:

(15) Onomas. Level: Object Action Process
 MSAP: *hand* – *wash* – *ing*

In meaning (3), however, *hand* was interpreted as the INSTRUMENT used for *washing*:

(16) Onomas. level: Instrument Action Process
 MSAP: *hand* – *wash* – *ing*

Clearly, extra-linguistic knowledge about hand washing plays a large role in the interpretation of this naming unit. It is not the overall structure that has an effect on the interpretation but rather how *hands* are related to *washing* in the real world.

5.3.2 Wedding day rehearsal (OT2 within OT1)

Structure

Wedding day rehearsal consists of several embedded OT2 structures. The first is *wedding*:

(17) determining – determined – base
 Onomas. level: Action Process
 MSAP: \emptyset – *wed* – *ing*

Wedding day is another OT2 structure, with the embedded OT2 structure *wedding* serving as the determined constituent:

(18) determining – determined – base
 Onomas. level: Process Time
 MSAP: \emptyset – *wedding* – *day*

Wedding day rehearsal is then a complete OT1 structure when *wedding day* becomes the determining constituent:

(19)		determining	–	determined	–	base
	Onomas. level:	Theme		Action		Process
	MSAP:	<i>wedding day</i>	–	<i>rehearse</i>	–	<i>al</i>

Ranking

1. ‘*a rehearsal for the events of a wedding*’

FO: 18/20

PR: 0.801

OPR: 0.700

2. ‘*a daytime rehearsal for a wedding*’

FO: 16/20

PR: 0.192

3. ‘*a rehearsal that is on a wedding day*’

FO: 11/20

PR: 0.151

Comments

Meaning (1) is the expected meaning, outlined in (19), where *wedding day* is the THEME. For meaning (2), *wedding day rehearsal* received an OT3 reading in which *wedding* is the THEME of *day rehearsal*:

(20)		determining	–	determined	–	base
	Onomas. level:	Theme				Event
	MSAP:	<i>wedding</i>	–	∅	–	<i>day rehearsal</i>

Day rehearsal is then an embedded OT1 structure in which *day* is mapped onto TIME:

(21)		determining	–	determined	–	base
	Onomas. level:	Time		Action		Event
	MSAP:	<i>day</i>	–	<i>rehearse</i>	–	<i>al</i>

This meaning may be more closely related to *daytime wedding rehearsal* rather than *wedding day rehearsal*. Meaning (3), while still an OT2 inside an OT1, reveals an interpretation of *wedding day* as TIME rather than THEME.

In this case, meanings (1) and (3) are interpreted within the intended overall structure (OT2 in OT1), but meaning (2) is interpreted within an overall OT3 structure. The competition from the high FOs of meanings (2) and (3) pull the OPR of meaning (1) down. This indicates that there might be some ambiguity as to what the overall structure of this naming unit actually is. Extra-linguistic knowledge must play a part in determining the most probable meaning because the PR of the meaning (1) is still within a high level of predictability.

The OPR for this naming unit was affected by competition from meanings (2) and (3). The FO for meaning (2) was quite close to the FO for meaning (1) which contributed to a moderately high PR score for meaning (2). Meaning (3) received the second highest PR and FO for third meanings out of all the naming units, which also lowered the OPR.

5.3.3 Childcare benefit cuts warning (OT3 within OT1)

Structure

Childcare benefit cuts is an OT3 structure:

(22)

	determining	–	determined	–	base
Onomas. level:	Theme				Agent
MSAP:	<i>childcare benefit</i>	–	∅	–	<i>cuts</i>

Childcare benefit cuts is then embedded in the OT1 structure as the determining constituent, and *warning* acts in the same way as it does for *hand washing warning*: *warn* represents the ACTION while *-ing* is the base:

(23)

	determining	–	determined	–	base
Onomas. level:	Theme		Action		Agent
MSAP:	<i>childcare benefit cuts</i>	–	<i>warn</i>	–	<i>ing</i>

Ranking

Since all meanings beyond (2) received an FO of 1/20, no other meanings than the top two are given.

1. ‘a warning issued about cuts to childcare benefits’

FO: 20/20

PR: 0.995

OPR: 0.997

2. ‘a warning against injury from childcare benefits’

FO: 2/20

PR: 0.002

Comments

There is very little competition for meaning (1) in this instance. Every participant proposed the meaning ‘a warning issued about cuts to childcare benefits.’ All but one participant gave this meaning a 10 ranking (that participant ranked it 9). This naming unit received the highest OPR of all the naming units. It is my suspicion, given that the highest FO after the first reading was 2/20 and the FO for all subsequent meanings was 1/20, that participants were finding it difficult to interpret this naming unit in any other plausible way. The instruction for participants to come up with as many meanings as they could think of may have been problematic in this case. All the other meanings were only proposed by one person at most and were only given a score of 1. No participant favored another meaning as the top meaning over ‘a warning issued about cuts to childcare benefits.’

Given that meaning (2) scored very low in both FO and PR, there seems to be almost no ambiguity over the interpretation of this naming unit. This is perhaps due to the actional nature of *cuts* in the OT3 structure *childcare benefit cuts*. The actional element (determined constituent) specifies the relationship between the other members of the onomasiological structure (Štekauer, 2005a: 52). For example, *drive* in *truck driver* specifies the relationship of the AGENT to the *truck* (he *drives* it). In some cases, such as *childcare benefit cuts*, the actional nature of either the determining constituent or base can fill in for the missing determined constituent. *Cuts* implies an action of cutting the benefits. This means that an overtly expressed determined constituent is not necessary in order to determine the relationship between the determining constituent and base.

Both meanings (1) and (2) receive the OT3 within OT1 reading, so the ambiguity seems to lie in the meaning of *cuts*. In meaning (1), *cuts* is interpreted as ‘reduction,’ while in meaning (2), it is interpreted as ‘minor injury.’ Rather than *cuts* being the AGENT, as in meaning (1), *childcare benefit* is the AGENT and *cuts* is the THEME:

(24)		determining	–	determined	–	base
	Onomas. level:	Agent				Theme
	MSAP:	<i>childcare benefit</i>	–	∅	–	<i>cuts</i>

The low PR and FO scores of meaning (2) indicate that extra-linguistic knowledge blocks *childcare benefits* from legitimately being interpreted as AGENT.

5.3.4 Growth inhibitor (OT4 within OT1)

Structure

This naming unit contains the OT4 structure *growth*:

(25)		mark	–	base
	Onomas. level:	Action		State
	MSAP:	<i>grow</i>	–	<i>th</i>

Growth becomes the determining constituent of the overall OT1 structure:

(26)		determining	–	determined	–	base
	Onomas. level:	Process		Action		Agent
	MSAP:	<i>growth</i>	–	<i>inhibit</i>	–	<i>or</i>

Ranking

1. ‘*something that inhibits growth*’
FO: 20/20
PR: 0.980
OPR: 0.773
2. ‘*something that inhibits a (bodily) growth from forming/enlarging*’
FO: 14/20
PR: 0.280
3. ‘*a growth that inhibits something*’
FO: 4/20
PR: 0.007

Comments

Growth inhibitor received the fifth lowest OPR score. The low OPR is due to high PR and FO scores for meaning (2). Meaning (3) had very little effect on the OPR with its low FO and PR scores. While four participants proposed meaning (3), most gave it a ranking of 1. The ambiguity between meanings (1) and (2) does not lie with the structure but rather at the onomasiological level. In meaning (1), *growth* is assigned to PROCESS (shown in (26)) while in meaning (2) it is assigned to PATIENT:

(27)		determining	–	determined	–	base
	Onomas. level:	Patient		Action		Agent
	MSAP:	<i>growth</i>	–	<i>inhibit</i>	–	<i>or</i>

For meaning (3), *growth* is reflexively expressed as the AGENT:

(28)		determining	–	determined	–	base
	Onomas. level:	Agent		Action		Agent
	MSAP:	<i>growth</i>	–	<i>inhibit</i>	–	<i>or</i>

It would seem unlikely that anyone would coin *growth inhibitor* for the meaning ‘a growth that inhibits something,’ as the AGENT expressed twice. It is analogous to coining *bus driver* for the meaning ‘a bus that drives,’ rather than coining *driving bus*.

5.3.5 Snowstorm face protectors (OT1 within OT3)

Structure

Snowstorm face protectors characterizes an OT1 structure inside an OT3 structure. *Face protector* is the OT1 structure ($face_{DG} - protect_{DD} - or_{BASE}$) which acts as the base for the OT3 structure:

(29)		determining	–	determined	–	base
	Onomas. level:	Target				Agent
	MSAP:	<i>snowstorm</i>	–	∅	–	<i>face protector</i>

A more comprehensive discussion of how structure relates to predictability can be found in section 5.4.2.

5.3.6 Employee sick absence (OT2 within OT3)

Structure

Employee is an OT2 structure ($\emptyset - \text{employ}_{\text{DD}} - \text{ee}_{\text{BASE}}$) where *-ee* maps onto PATIENT. *Sick absence* itself maps onto a CAUSE – CIRCUMSTANCE relationship in an OT3 structure ($\text{sick}_{\text{DG}} - \emptyset - \text{absence}_{\text{BASE}}$). *Sick absence* becomes the base while *employee* becomes the determining constituent:

(31)	determining	–	determined	–	base
Onomas. level:	Patient				Stative
MSAP:	<i>employee</i>	–	\emptyset	–	<i>sick absence</i>

Ranking

1. ‘*an employee’s absence due to being sick*’
FO: 19/20
PR: 0.869
OPR: 0.975
2. ‘*a disgusting reason for an employee’s absence*’
FO: 7/20
PR: 0.014
3. ‘*an absence of sick employees*’
FO: 3/20
PR: 0.008

Comments

Meaning (1) received an interpretation in line with the intended structure and semantics, although it is difficult to distinguish whether *sick* was interpreted as a modifier or a clipping of *sickness* for meaning (1) (‘sick employee absence’ versus ‘employee sickness absence’). Meaning (2) received a figurative interpretation of *sick* meaning ‘gross/disgusting.’ The ambiguity between meanings (2) and (3) is in regard to which constituent *sick* modifies, *absence* or *employee*, but these meanings had little effect on the OPR.

It is possible that participants were confused over the placement of *sick* within the naming unit and chose to ignore the order of the constituents when interpreting the naming unit. It may be that participants were interpreting this naming unit as if it were *sick employee absence*. Jackendoff (2009: 111) says that the productivity of compounding is fragile. Sometimes listeners parse out complex word structures incorrectly or a re-wording often will make more sense, which is probably the case here. Whether read as *sick employee absence* or *employee sick absence*, the structure remains the same: OT2 embedded in OT3. *Sick* is not consequential to the overall word structure of the top three meanings. When *sick* is removed, the remaining word structure is still that of OT2 within OT3:

(32)

	determining	–	determined	–	base
Onomas. level:	Patient				Stative
MSAP:	<i>employee</i>	–	∅	–	<i>absence</i>

With the second highest OPR and an almost unanimous FO, however, it was clear to participants what the most probable meaning was. The very low PR and FO scores meanings (2) and (3) support this claim.

5.3.7 Primary pupil booklist (OT3 within OT3)

Structure

For this naming unit to get an OT3 reading, *primary pupil* must be regarded as a pupil from a primary school. *Primary*, in this case, serves as a clipping of *primary school*, similar to *grad student* being a clipping of *graduate school student*. *Primary pupil* is the determining constituent and *booklist*, an OT3 structure ($\text{book}_{\text{DG}} - \emptyset - \text{list}_{\text{BASE}}$), is the base:

(33)

	determining	–	determined	–	base
Onomas. level:	Target				Theme
MSAP:	<i>primary pupil</i>	–	∅	–	<i>booklist</i>

Ranking

1. '*a list of books for primary school students*'
FO: 18/20
PR: 0.756
OPR: 0.531
2. '*the main list of books for students*'
FO: 16/20
PR: 0.504
3. '*a list of books for the most important student*'
FO: 13/20
PR: 0.163

Comments

The obvious ambiguity for this naming unit is over the status of *primary* either as an adjective or a clipping. Furthermore, there is ambiguity over whether *primary* specifies *pupil* or *booklist*. As all participants were American, the high FO and PR scores for meaning (2) might be due to the selection of this naming unit from a British news outlet as *primary school* in British English is *elementary school* in American English. However, extra-linguistic knowledge plays a clear role in the interpretation for meaning (1) because most participants were familiar with *primary school* as a name for the school that younger children attend. This does not downplay the fact that several participants proposed what they thought to be valid alternate meanings that provided tough competition for the most probable meaning. This resulted in the lowest OPR score for all naming units.

The second and third meanings reveal *primary* as an adjective modifying *list* and *pupil*, respectively. Regardless of the ambiguity, the structure remains an OT3 for the top three meanings, which means OT structure did not contribute to ambiguity. Only one participant offered an OT1 reading for the naming unit:

(34) '*A list of the most important students who have been arrested (booked)*'

	determining	–	determined	–	base
Onomas. level:	Object		Action		Patient
MSAP:	(primary) pupil	–	book	–	list

In most cases, however, *booklist* was interpreted as an OT3 structure, the intended recursive structure.

5.3.8 Arab world unrest (OT4 within OT3)

Structure

Unrest is an example of an OT4 structure where the mark cannot be separated into constituents:

(35)		base	–	mark
	Onomas. level:	Negation		State
	MSAP:	<i>un</i>	–	<i>rest</i>

Un- serves as the base in this case because, conceptually, the negating element is a general class of things that are NOT, and *rest* specifies what is NOT. *Unrest* serves as the base in the overall OT3 structure:

(36)		determining	–	determined	–	base
	Onomas. level:	Theme				State
	MSAP:	<i>Arab world</i>	–	∅	–	<i>unrest</i>

Ranking

1. '*conflict in Arabian nations*'
FO: 19/20
PR: 0.898
OPR: 0.796
2. '*conflict in the world created by Arabs*'
FO: 13/20
PR: 0.172
3. '*lack of sleep among Arabs*'
FO: 8/20
PR: 0.058

Comments

Similar to *primary pupil booklist*, the ambiguity between meanings (1) and (2) lies in which constituent *Arab* specifies. In meaning (1), it is unrest for the Arab world, and in meaning (2), it is world unrest caused by Arabs. Despite a moderately high FO for the

latter meaning, the OPR for meaning (1) is relatively high. I chose this naming unit on the assumption that *Arab world* was lexicalized for most people and that any interpretation of *Arab* as a modifier would not have a great impact on the top ranked meaning. Indeed, while the meaning (2) has an FO of more than half the participants, it is not ranked highly. With an almost unanimous proposal of meaning (1), it seems to be the case that *Arab world* is interpreted as a unit for most people.

The OT structure does not affect the interpretability as each of the top three meanings exemplifies an OT4 within an OT3 structure. The third meaning seems to include an interpretation of *unrest* where *rest* implies ‘sleep’ rather than ‘peace.’ It may be that participants were, again, feeling challenged by the naming unit and creatively came up with a possible meaning in which *unrest* is not defined as ‘turmoil’ or ‘disturbance’ but as ‘lack of sleep.’ Although 8/20 participants proposed an interpretation of *unrest* as ‘lack of sleep,’ a more plausible naming unit for meaning (3) would be *Arab world sleeplessness* since *unrest* has no dictionary definition¹ of ‘lack of sleep.’ It is possible, however, for a coiner to invent a word like *unrest* or *unsleep* for the meaning found in (3). Most interestingly, three participants ranked this meaning as a 5 and one even gave it a 10! A check for this sort of proposed meaning would be to give participants the meaning ‘lack of sleep among Arabs’ and ask them to propose naming units for such a meaning to see if anyone coins *unrest* or *unsleep*.

5.4 Discussion

5.4.1 Predictability of the Naming Units

Štekauer divides the predictability rates of his naming units into four levels in order to evaluate his findings (2005a: 153):

(37)	0—0.25	unpredictability level
	0.26—0.50	low predictability level
	0.51—0.75	medium predictability level
	0.76—1.00	high predictability level

¹ Verified with OED online and Merriam-Webster online.

All of the naming units in this study scored above the high predictability level except *hand washing warning*, which scored in the medium predictability level. No naming units scored below medium. All OT3 structures scored in the high range with *primary pupil booklist* scoring just at the boundary between medium and high. Incidentally, although *hand washing warning* received the lowest PR, *primary pupil booklist* received the lowest OPR. This is one indication that PR is not a direct indicator of OPR.

Table 1 also shows that the PR score and the PR Gap for the top reading are not necessarily indicative of the OPR ranking. As shown in the Table 1, *childcare benefit cuts warning* received the same PR score as *snowstorm face protectors*; however, the PR Gap ratios are significantly different, resulting in varied OPRs for the two naming units. The two naming units with the lowest PR scores are *hand washing warning* and *primary pupil booklist*. While the PR value for *hand washing warning* is lower than that of *primary pupil booklist* by 0.194, the OPR for *primary pupil booklist* is lower than that of *hand washing warning* by 0.045. Furthermore, even though *employee sick absence* ranks fifth in PR value, it ranks second in OPR value due to almost no competition from its meanings (2) and (3). *Snowstorm face protectors* faces more competition from its meanings (2) and (3) so its rank goes from 1 for PR value to 3 for OPR value.

Most predictable reading for each naming unit	PR	PR-based rank	PR Gap: Absolute Value	PR Gap: R1/R2 ratio	OPR	OPR-based rank
Hand washing warning (OT1 in OT1)	0.562	8	0.221	1.648	0.576	7
Wedding day rehearsal (OT2 in OT1)	0.801	6	0.609	4.172	0.700	6
Childcare benefit cuts warning (OT3 in OT1)	0.995	1	0.993	497.5	0.997	1
Growth inhibitor (OT4 in OT1)	0.980	3	0.700	3.500	0.773	5
Snowstorm face protectors (OT1 in OT3)	0.995	1	0.938	17.46	0.941	3
Employee sick absence (OT2 in OT3)	0.869	5	0.855	62.07	0.975	2
Primary pupil booklist (OT3 in OT3)	0.756	7	0.252	1.5	0.531	8
Arab world unrest (OT4 in OT3)	0.898	4	0.726	5.221	0.796	4

Table 1 (Štekauer, 2005a: 154)

The competition of meaning (2), both in FO and PR, has a distinct effect on the OPR. Remember that it is the OPR, rather than the PR or FO, that actually indicates how predictable the meaning of the naming unit is against the meaning of other naming units. The PR value only says how strong a given meaning is in relation to the other proposed meanings for that naming unit.

5.4.2 To What Extent Does Structure Affect Meaning Predictability?

The answer to this research question can be explained with regard to the overall onomasiological type, the onomasiological type of the recursive structure, and the structures of the naming units with embedding in comparison to Štekauer's unembedded naming units. For the current study, structure plays some part in determining predictability, but it is extra-linguistic knowledge that dominates meaning predictability.

Only in *wedding day rehearsal* and *snowstorm face protectors* did any difference in structural interpretation affect the meaning predictability. For both naming units, meanings (1) and (2) differed in structural interpretation, while meaning (3) was interpreted with the intended structure. The structure of *wedding day rehearsal* was interpreted as an overall OT1 for meaning (1) and an overall OT3 for meaning (2). Meaning (1) was interpreted with a recursive OT2 within an OT1, and meaning (2) was interpreted as a recursive OT1 within an OT3. When *day* and *rehearsal* were combined into the unit *day rehearsal* and interpreted as the base (meaning (2)), the overall OT1 structure lost its determined constituent (rehearse) and, consequently, became an overall OT3 structure with *wedding* as the determining constituent (see (38)).

For *snowstorm face protectors*, meaning (1) was interpreted with an overall OT3 structure, and meaning (2) was interpreted as an overall OT1 structure. Given the presence of an ACTION within the base (*protect*), it is hardly surprising that participants interpreted *protectors* as the determined constituent plus base (\emptyset – $\text{protect}_{\text{DD}}$ – OTS_{BASE}) and

combined *snowstorm face* into the determining constituent for meaning (2) (see (39)).

It seems that there can be ambiguity between an overall OT1 and OT3 reading in these cases because the actional element can be interpreted as the base when, as a recursive structure, it contains a determining constituent (*face* and *day* in (38)).

(38) Overall OT3 reading:

	determining	–	determined	–	base
Onomas. level:	Target				Agent
MSAP:	<i>snowstorm</i>	–	∅	–	<i>face protector</i>
	determining	–	determined	–	base
Onomas. level:	Theme				Event
MSAP:	<i>wedding</i>	–	∅	–	<i>day rehearsal</i>

Face protector and *day rehearsal* are complete OT1 structures inserted as the base. When *face* is grouped with *snowstorm* and *day* is grouped with *wedding* (see (39)), the OT1 structure of the base is disassembled and the actional element (*protector* and *rehearsal*) is divided into the determined constituent plus base.

(39) Overall OT1 reading:

	determining	–	determined	–	base
Onomas. level:	Object		Action		Agent
MSAP:	<i>snowstorm face</i>	–	<i>protect</i>	–	<i>or</i>
	determining	–	determined	–	base
Onomas. level:	Theme		Action		Process
MSAP:	<i>wedding day</i>	–	<i>rehearse</i>	–	<i>al</i>

My hypothesis that *the actional nature of a determining constituent or base will “fill in” for the missing actional element in OT3 structures creating a higher OPR than for other OT3 structures* seems to be verified by *snowstorm face protectors* insofar as its meaning (1) received the highest PR score (tied with *childcare benefit cuts warning*) overall. However, due to the actional nature of the base, meaning (2) of *snowstorm face protectors* brought the OPR score

down to a number 3 ranking overall for OPR because of structural ambiguity. The presence of an actional element in one of the constituents, while possibly contributing to a higher PR score, actually decreased the OPR score in this case.

Overall, the onomasiological type did not seem to affect the predictability rate in a significant way. Similarly, recursive structures of OT2 or OT4 were always interpreted as OT2 and OT4 structures, which means their interpretations did not affect the predictability rates. As shown in (38) and (39), OT3s were sometimes interpreted as OT1s and vice versa, which may have affected their OPR scores. As Table 1 shows, the OPR rankings for combinations of OT1 and OT3 structures were as follows:

- (40) OT1 within OT1 (*hand washing warning*) – rank 7
- OT3 within OT1 (*childcare benefit cuts warning*) – rank 1
- OT1 within OT3 (*snowstorm face protectors*) – rank 3
- OT3 within OT3 (*primary pupil booklist*) – rank 8

The rankings of only one of these combinations can be explained by structure: OT1 within OT3 (*snowstorm face protectors*). Incidentally, however, the OPR for *snowstorm face protectors* is still highly ranked as a result of extra-linguistic information.

Primary pupil booklist, ranked 8, received a low OPR because of ambiguity over the status of *primary*, which is an extra-linguistic issue rather than a structural issue. For *hand washing warning*, ranked 7, the low OPR was due to a split in the meaning of a semantic element that was missing (to do or not to do something) rather than in the interpretation of one of the meanings as OT3. Again, this was due to extra-linguistic knowledge rather than structure. *Childcare benefit cuts warning* received the highest OPR and was only interpreted as OT3 within OT1. These cases reaffirm the hypothesis that *extra-linguistic knowledge will affect the OPR when the structure does not*.

Štekauer (2005a: 255) concludes that onomasiological type 3, with the absence of the actional element, is a “serious obstacle” to meaning predictability. It was my hypothesis that data from the OT3 naming units would affirm this conclusion; however, my hypothesis is

not supported by the data in Table 1. The data reveals an average PR score of 0.835 for OT1 structures and an average PR score of 0.880 for OT3 structures, a difference of 0.045. Moreover, the average OPR for OT1 structures is 0.762 while the average OPR for OT3 structures is 0.811, a difference of 0.049. While the differences are small, they do indicate that when compared to OT1 structures, OT3 structures tend to be slightly MORE predictable. To verify this tendency, a longer list of words with OT1 and OT3 structures should be tested.

While Štekauer's argument that OT3 structures are PREDICTABILITY REDUCING structures based on their lack of an actional element, it may be that actional elements within the recursive structures provide enough information to counterbalance the lack of an overt determined constituent. *Snowstorm face protectors* received the top PR score indicating that the intended meaning was very clear to all participants. That *face protectors* contains the action *protect* may have helped the predictability rate. My hypothesis that *the structure of the recursive element will affect the predictability if the recursive structure contains an expressed actional element* is validated by this naming unit. The PR of *snowstorm face protectors* may have been enhanced by the actional element of the recursive structure while at the same time the OPR may have been reduced due to the actional element. Predictability is, therefore, affected by the actional element regardless of whether the effect is positive or negative.

Primary pupil booklist is the only clear example that supports the following hypothesis: *When a constituent can be interpreted as a modifier, the OPR will decrease due to greater competition from other meanings.* This naming unit in particular received a lower OPR due to high competition from meanings (2) and (3) regarding the status of *primary* as a modifier. However, in *Arab world unrest*, *Arab* was also treated as a modifier in meaning (2) but the PR score for this meaning was fairly low. Although the FO for meaning (2) was above half of the participants (13/20), the PR score for *Arab world unrest* seems to be affected more by the lexicalization of *Arab world*, which is a matter of extra-linguistic knowledge. Therefore, *Arab* acting as a modifier of

world unrest hardly provided competition against meaning (1). Similarly, meanings (2) and (3) for *employee sick absence* were based on the status of *sick* as a modifier of *employee absence*. Further testing on ambiguous modifiers would provide better evidence in support of the above hypothesis.

Finally, the hypothesis that *the PRs for naming units with recursive structures will be higher than those found for analogous structures in Štekauer's data* has been validated by the data in Table 1. In one of his experiments regarding OT3 structures, Štekauer (2005a: 255) makes the claim that OT3s are prone to low PR scores and can be categorized within the low predictability level. The average PR score for his naming units is 0.322. The average PR score for the OT3 naming units in my study is 0.880, a rating well within the high predictability level and a very different outcome than Štekauer's. This could be attributed to the additional linguistically-expressed elements of the recursive structures in my naming units.

The high PR scores for my study indicate that participants agreed on what the meaning of a given naming unit might be. The low PR scores for Štekauer's naming units indicated that many participants had a difficult time locating a meaning they considered to be an acceptable reading. Štekauer (2005a: 110) himself says that this may be due to incompatibility of the semes for a naming unit (i.e., how a spade might be related to a dog in *dog spade*). The high PRs for my study seem to indicate that participants found the elements in the naming units compatible.

Although, Štekauer's and my scores differ in PR value for OT3 structures, our OPR scores are quite similar. In Štekauer's experiment (2005a: 212), the average OPR for the OT1 naming units was 0.691, and the average OPR for the OT3 naming units was 0.775. These numbers are 0.071 and 0.036 lower than mine (respectively), which is not a considerable difference. The significance of these findings is that the OT3 structures in my experiment tend to have much more obvious intended meanings than Štekauer's naming units but the meanings are only slightly more predictable for my naming units than for Štekauer's.

6 Conclusion

Štekauer (2005a: 6) asserts—and it has also been shown in this study—that a crucial element of understanding and predicting the meaning of compounds is the extra-linguistic knowledge of each speaker. For example, *apartment dog* is less likely to be interpreted as ‘a dog that MAKES an apartment’ or ‘a dog that HAS an apartment,’ because dogs do not build or own apartments. More likely, the most probable meaning of *apartment dog* would be ‘a dog suitable for apartment-style living (i.e., smaller, quieter, etc.).’ This meaning is based on a speaker’s knowledge of dogs and the possible interactions they have to apartments. Extra-linguistic knowledge restricts the possible meanings for a given naming unit. Štekauer’s own experiments, then, give rise to three assumptions regarding language users’ abilities to predict meanings. A language-user (Štekauer, 2005a: 251):

1. must know the lexical meanings of the motivating constituents [of a naming unit];
2. must be able to conceptually analyze the objects of extra-linguistic reality which are covered by these lexical meanings; and
3. identify their possible relations based on his/her knowledge and/or experience.

Bauer (1978: 15 quoting Carr, 1939) notes that “compounds are more vague and less precise than syntactic phrases, but what the compounds lose in precision they gain in flexibility and suggestiveness.” It is flexibility and suggestiveness that allows speakers to coin new naming units with a specific meaning in mind. The coiner homes in on a SPECIFIC relationship between SPECIFIC features of the constituents of the naming unit. When interpreting a naming unit, however, the specificity of a meaning is up to the interpreter. Neither Levi (1978) nor Downing (1977) is correct in their assertions about compound meaning. The former limits too drastically the number of probable meanings while the latter leaves room for infinite meanings. As the data presented in this investigation shows,

while the number of POSSIBLE meanings may be infinite, the number of PROBABLE meanings is limited. In fact, for such recursive naming units as the ones in this study, only one probable meaning exists.

Naturally, a consequence of the flexibility and suggestiveness of compounds is ambiguity, and “the richness of conceptual structures is the main source for compound ambiguity” (Meyer, 1993: 103). The data presented above illuminate three factors that contribute to PR- and OPR-reducing ambiguity:

1. Extra-linguistic factors such as multiple semantic roles for one constituent and which relationships between constituents are deemed plausible
2. Structural ambiguity regarding combinations of recursive OT1 and OT3 structures within overall OT1 and OT3 structures
3. Interpretation of certain constituents as both modifiers and nominals

When discussing PR- and OPR-reducing factors, one must bear in mind that all but one naming unit from the study scored in the high level of predictability for PR. What has been shown in this study is that naming units with recursive structures seem to score much higher than those without recursive structures. Because of additional components within the structure, the number of plausible relationships between all the constituents is reduced, creating a higher probability that a certain meaning is dominant. The interaction of the recursive structure and overall structure played only a slight role in affecting predictability. Štekauer (2005a) made assumptions about the effects of onomasiological type on predictability and the current study challenged those assumptions. In order to make conclusive remarks regarding the effects of structure type for recursive naming units, a longer study with a more extensive list of naming units must be conducted. What is evident in both studies, however, is that extra-linguistic knowledge is the most important factor in interpreting new naming units.

References

- Bauer, L. (1978) *The Grammar of Nominal Compounding*, Odense: Odense Universitets Publikationskonto.
- Downing, P. (1977) 'On the creation and use of English compounds', *Language*, vol. 53, December, pp. 810-842.
- Grzenga, J. (2002) 'Some thoughts on a cognitive onomasiological approach to word-formation with special reference to English', *Onomasiology Online*, vol. 3, pp. 1-29.
- Grzenga, J. (2009) 'Compounding from an onomasiological perspective' in Lieber, R. and Štekauer, P. (ed.) *The Oxford Handbook of Compounding*, Oxford: Oxford University Press.
- Horecký, J. (2007) 'An onomasiological interpretation of word-formation', *SKASE*, vol. 4, pp. 21-7.
- Jackendoff, R. (2009) 'Parallel architecture and conceptual semantics' in Lieber, R. and Štekauer, P. (ed.) *The Oxford Handbook of Compounding*, Oxford: Oxford University Press.
- Kastovsky, D. (2005) 'Hans Marchand and the Marchandean' in *Handbook of Word-Formation*. Štekauer, P. and Lieber, R. (ed.) Dordrecht: Springer.
- Lees, (1960) 'The grammar of English nominalizations' in *International Journal of American Linguistics*, vol. 26, July, pp. 1-205.
- Levi, J. (1978) *The Syntax and Semantics of Complex Nominals*. San Francisco: Academic Press.
- Lieber, R. & Štekauer, P. (2009) 'Introduction: status and definition of compounding' in Lieber, R. and Štekauer, P. (ed.) *The Oxford Handbook of Compounding*, Oxford: Oxford University Press.
- Marchand, H. (1969) *The Categories and Types of Present-Day English Word Formation*. München: Beck.
- Plag, I. (2003) *Word-Formation in English*. Cambridge: Cambridge University Press.
- Štekauer, P. (1998) *An Onomasiological Theory of English Word-Formation*. Amsterdam: John Benjamins.
- Štekauer, P. (2001) 'Fundamental principles of an onomasiological theory of English word-formation', *Onomasiology Online*, vol. 2, pp. 1-42.

Stekauer, P. (2005a) *Meaning Predictability in Word Formation*.
Amsterdam: John Benjamins.

Stekauer, P. (2005b) 'Onomasiological approach to word-formation',
in Stekauer, P. & Lieber, R. (ed.) *Handbook of Word-
Formation*, Dordrecht: Springer.

Ten Hacken, P. (2009) 'Early generative approaches' in Lieber, R. and
Štekauer, P. (ed.) *The Oxford Handbook of Compounding*,
Oxford: Oxford University Press.

Appendix A

Informed Consent

By typing my name and the date below, I acknowledge my willingness to participate in the following study, and I agree that my answers may be used for this specific research project *Meaning Predictability and Recursion in Onomasiology* conducted by [REDACTED]. I consent to allow my answers to be used in teaching or research-related presentations or publications by the above-named researcher as well as other researchers. I understand that my identity will remain anonymous and unconnected with my answers. I understand that if I have any questions regarding the following task that I may contact [REDACTED] at [REDACTED].

Full name:

Date:

Instructions – Please read carefully!

The following is a list of compound words found in online news headlines. Propose as many meanings as you can think of for each of the compound words. Not every meaning you propose may be straightforward or make much sense, especially if you are fairly sure of the intended meaning. List these anyway as potential meanings. Assign a score (from 1 to 10) to each of the meanings you propose, with a 10 indicating the highest probability that the meaning proposed would occur as the intended meaning (the most predictable), and assign a 1 to the meaning with the minimum chance of occurring as the intended meaning.

For example:

Water-mill

–a mill driven by water	10
–a mill near water	4
–a mill made of water	2
–a mill the color of water	1
–a mill that grinds water	1

(like a pepper mill grinds pepper)

1. hand washing warning
2. wedding day rehearsal
3. childcare benefit cuts warning
4. growth inhibitor
5. snowstorm face protectors
6. employee sick absence
7. primary pupil booklist
8. Arab world unrest

Appendix B

Hand washing warning: BBC News. Found at www.bbc.co.uk/news/ on June 6, 2011

Wedding day rehearsal: BBC News. Found at www.bbc.co.uk/news/ on April 27, 2011

Childcare benefit cuts warning: BBC News. Found at www.bbc.co.uk/news/ on June 6, 2011

Snowstorm face protectors: The Atlantic. Found at www.theatlantic.com/technology/archive/2010 on August 24, 2010

Employee sick absence: BBC News. Found at www.bbc.co.uk/news/ on June 6, 2011

Primary pupil booklist: BBC News. Found at www.bbc.co.uk/news/ on May 7, 2011

Arab world unrest: The New York Times. Found at www.nytimes.com on May 5, 2011