Hans-Jörg Schmid

What is Creative to Whom and Why? Creativity in Word-Formation and Phraseology Against the Backdrop of Shared Conventions and Individual Routines

Abstract: Creativity can be defined in terms of a balance of originality and effectiveness. I interpret originality in terms of degrees of deviation from convention and effectiveness in terms of degrees of mutual understanding. The challenge for exploring creativity in the domains of word-formation and phraseology is that the schemas and templates that are available in these domains offer considerable scope for regular creativity (i.e. productivity), but also for daring deviations from existing conventions. I resolve this tension by relying on Sampson's (2006) distinction between F-creativity and E-creativity and my own Entrenchment-and-Conventionalization Model (Schmid 2020). I discuss systematic ways in which speakers exploit word-formation and phraseology for acts of linguistic creativity which are salient to different degrees and can be creative vis-à-vis different dimensions of regularity. Essentially, they can pertain to the symbolic dimension of regularity and affect form-meaning pairings, the syntagmatic dimension, affecting combinatorial regularities, and/or the socio-pragmatic dimension, affecting functional, situational and social regularities. The role of factors widening or limiting the scope for creativity, among them individual cognitive routines, is also discussed and included in the proposed definition of linguistic creativity.

1 Introduction

The notion of *creativity* is a slippery customer. According to frequently cited sources, it can be defined in terms of 'originality and effectiveness' (Runco and Jaeger 2012: 92) or 'novelty and appropriateness' (Sternberg and Lubart 1999: 3). Definitions of creativity in linguistics depend heavily on definitions of language and on the linguistic level in focus (see Bergs 2019; Hoffmann 2022; and Tin 2022 for surveys from different perspectives). Transferring the general definitions of creativity to language in a preliminary manner, utterances can be considered 'original' and 'novel' to the extent that they are unconventional; they are 'effective' to the

https://doi.org/10.1515/9783111321905-005

extent that they facilitate effortless mutual understanding and successful communication; and they are 'appropriate' to the extent that they are suitable for reaching relevant communicative goals and fit the current context.

When looking at creativity in word-formation and phraseology, the domains of linguistics at hand, it pays off to take into account Sampson's (2016) distinction between two types of creativity: *F-creativity* (short for *fixed creativity*) and *E-creativity* (i.e. *extending creativity*). In the present context, F-creativity refers to novel word-formation products or multi-word expressions that are based on existing word-formation patterns or phraseological templates, respectively. This notion largely corresponds to the more traditional notion of *productivity* (Bauer 2001; Plag 1999). E-creativity denotes original formations that go beyond regular applications of productive patterns in one way or other and tend to be salient enough to be recognized as being motivated by the intention to be creative. Anticipating examples to be discussed in more detail below, we can say that the attested but rare form *unrouteable* is F-creative, since it is based on the form of and meanings associated with the productive pattern *un-* [V *-able*]. In contrast, a fabricated form such as *unhappyable* may be considered E-creative, because the use of the adjective *happy* as a basis constitutes a more marked deviation from this pattern. It should be noted right from the start that the boundary between F-creativity and E-creativity is fuzzy. This does not mean, however, that it is pointless to make the distinction. Instead, it will remain useful for theoretical and argumentative purposes to distinguish between these two types of creativity, especially when it is important to highlight differences between them.

The notions of *effectiveness* and *appropriateness* mentioned in the definitions quoted above can be considered to stand for counter-forces that constrain novelty. Only novel utterances that are communicatively successful and considered appropriate count as creative, depending crucially, of course, on the context. Original ways of speaking which cross the limits set by effectiveness and appropriateness are no longer regarded as creative ways of using language, but as unsuccessful or downright wrong (see, e.g., Bergs 2019; Uhrig 2018). How a given utterance is assessed or experienced will depend on numerous other factors, in addition to context, e.g. the genre and register. The linguistic routines as well as the observed linguistic and stylistic competence of the speaker and the hearer will also affect whether an utterance is intended and perceived to be creative. An unconventional expression found in a verse written by a renowned poet is less likely to be viewed as 'unsuccessful' or 'aberrant' than a creative form used by a language learner who is obviously grappling with the grammatical and lexical resources needed to get across what they want to say.

Useful as the definitions just discussed certainly are, they come with a list of follow-up questions related to the notions which serve as criteria for defining creativity. These questions include the following:

1. If creativity is defined by unconventionality, novelty and originality, what is the benchmark or base line for deciding what is or is not conventional, novel or original?
2. What are the systematic ways in which creative utterances can deviate from this base line?
3. How do effectiveness and appropriateness constrain creativity, and what other forces working against creativity can be identified?
4. What are the factors that can loosen the constraints on creativity imposed by effectiveness, appropriateness and other potentially constraining forces?
5. What is the role of the cognitive routines of individual speakers in the production and perception of creative utterances?

In this contribution, I would like to demonstrate how the Entrenchment-and-Conventionalization Model (or EC-Model, see Schmid 2015, 2020) can contribute to answering these questions. This model provides a unified account of how linguistic conventions emerge and are sustained in a community of speakers and how linguistic knowledge emerges and is sustained in the minds of individual speakers. As such, it provides a good starting point for answering question 1), and it will also turn out that it is helpful for answering the other four questions.

The two linguistic domains that are in the focus of this volume and the present contribution can be demarcated as follows. The domain of word-formation encompasses recurrent patterns that can be used to create new lexemes from existing morphological resources: derivation, compounding and conversion, as well as blending, clipping, back-formation and the formation of acronyms. Regarding creativity, it is important to keep in mind that the core morphemic processes are either partly or fully variable patterns, in the sense of construction grammar (see, e.g., Goldberg 2019). This means that in contrast to lexically or morphologically specific items and chunks, they include at least one variable slot (in the case of partly variable patterns) or consist only of variable slots (in the case of fully variable schemas):

– derivations are partly variable patterns, because the prefixes and suffixes used are specific morphological elements (the fixed part) that are combined with different bases (the variable slot; e.g. *un-* + Adj or V + *-able*);
– conversion is a fully variable pattern, because it is not marked by any type of constant morphological substance and therefore does not include any fixed elements (e.g. *hammer*$_{N}$ and *hammer*$_{V}$);

– compounding varies from fully variable patterns, such as N + N (*flower pot, paper tissue*) or Adj + N (*hotbed, blackbird*), to partly specific ones, e.g. V + *-ing* + N (*swimming pool, walking stick*) or N + V + *-er* (*laser printer, dish washer*).

The domain of phraseology encompasses a wide range of linguistic phenomena sharing the property that several words are combined to build more or less fixed chunks bearing more or less idiomatic meanings. On one pole of these continua, proverbs, commonplaces, routine formulae, idioms and lexical bundles are relatively fixed and lexically 'unique', so to speak. For example, the proverbs *a stitch in time saves nine* and *the early bird catches the worm* are only comparable regarding abstract properties such as their functions and usage contexts, but they do not share any lexically specific material. On the other end, phraseological templates, multi-word verbs, prepositions and conjunctions as well as collocations tend to occur in numerous variant forms and therefore lend themselves to being regarded as productive patterns, due to the variable slot or slots they include. For example, the commonality shared by utterances of the type *as white as a sheet, as free as a bird* or *as light as a feather* can be captured as a partly variable phraseological template of the form *as* Adj *as a* N, which can serve as a productive source for new formations.

It is precisely the tension between complex lexemes and fixed expressions which are established as lexically specific items, partly variable patterns and fully variable patterns that makes the domains of word-formation and phraseology so interesting for the study of linguistic creativity. This tension provides speakers with the opportunity to be creative to different degrees, in a variety of ways and in the pursuit of different goals. One of the main challenges to be tackled here is to do justice to this flexibility and yet manage to come up with a unified concept of creativity.

To enrich the discussion with examples from word-formation and phraseology, I focus on two patterns (but will add other examples where necessary or helpful): the word-formation pattern *un-* + X + *-able*$_{\text{Adj}}$ (e.g. *uncomfortable, unpredictable*) and the phraseological template *what the N* (e.g. *what the devil, what the hell*).

2 Two Exemplary Patterns: *un-* + X + *-able*$_{\text{Adj}}$ and *what the N*

Adjectives exhibiting the form *un-* + X + *-able*$_{\text{Adj}}$ can have as many as five different constituent structures:

1. The frequent binary right-branching verb-based structure *un*-[V-*able*]$_{\text{Adj}}$ found in *unavoidable* or *unpredictable*, i.e. *un*-prefixations of adjectives that are derived from verbs by means of the suffix *-able*.
2. The rarer binary left-branching verb-based structure [*un*-V]-*able*$_{\text{Adj}}$ as in *unscrewable*, i.e. verbs that are prefixed by reversative *un*- and then transformed into adjectives by means of the suffix *-able*.
3. The binary right-branching noun-based structure *un*-[N-*able*]$_{\text{Adj}}$ as in *unknowledgeable* or *unpleasurable*, i.e. negated adjectives that are derived from nouns with the suffix *-able*.
4. The ternary, 'synthetic' structure [*un*- X -able]$_{\text{Adj}}$ for forms which are difficult or impossible to align with one of the two possible binary analyses, as in *unskillable* or *unseatable*.
5. A structure derived from a bound root or base of unclear status, as in *untenable*, *unconscionable* or *unviable*.

Regarding the semantic structure of items sharing the form *un*- + X + *-able*, we can distinguish three dominant types (see Bauer, Lieber and Plag et al. 2013: 307–309):

– subject reference, as in *uncomfortable* or *unknowledgeable*,
– object reference, as in *unpredictable* or *unmarriageable*,
– and, more rarely, adverbial reference, as in *unpigeonholeable*.

As far as meanings associated with this pattern are concerned, the literature on prefixation by *un*- and suffixation by *-able* (e.g. Bauer, Lieber and Plag 2013: 291–303, 354–389; Hansen et al. 1985: 85, 112–113; Plag 2004) provides at least the following options (to which the adjectives can often not be allocated unambiguously):

– 'that cannot be V-ed', 'that is unable/unfit to be V-ed', e.g. *unavoidable, undeniable*
– 'that is not worthy of being V-ed', 'that should not be V-ed', e.g. *unacceptable, unlikeable*
– 'that does not V', 'that is unlikely to V', e.g. *unsuitable, unalterable*
– 'that is unable/unfit/unlikely to be subjected to N', e.g. *unmarriageable, unobjectionable*
– 'not having/being characterized by N', e.g. *unknowledgeable, unreasonable*
– 'not being in accordance with N', e.g. *unfashionable, unreasonable*
– 'not being capable of giving/bringing N', e.g. *uncomfortable, unpleasurable*.

To provide data illustrating this pattern from the British National Corpus (BNC), I ran the simple search `un+able_AJ0` to extract examples of adjectives beginning with *un*- and ending in *-able* (excluding *unable* by means of the wildcard "+",

Table 1: Extract of the frequency list of items harvested by the search `un+able_AJ0` from the BNC.

Rank	Item	Freq.	Prop.	Rank	Item	Freq.	Prop.	Rank	Item	Freq.	Prop.
1	*uncomfortable*	1327	8.33%	123	*unanalysable*	7	0.04%	382	*unscrewable*	1	0.01%
2	*unacceptable*	1219	7.66%	124	*unassimilable*	7	0.04%	383	*unseatable*	1	0.01%
3	*unreasonable*	978	6.14%	125	*unattributable*	7	0.04%	384	*unseverable*	1	0.01%
4	*understandable*	829	5.21%	126	*unforgiveable*	7	0.04%	385	*unsexable*	1	0.01%
5	*unstable*	697	4.38%	127	*unmatchable*	7	0.04%	386	*unshaveable*	1	0.01%
6	*unsuitable*	690	4.33%	128	*unrealisable*	7	0.04%	387	*unshirkable*	1	0.01%
7	*unpredictable*	673	4.23%	129	*unresectable*	7	0.04%	388	*unsignable*	1	0.01%
8	*undesirable*	617	3.88%	130	*unteachable*	7	0.04%	389	*unsingable*	1	0.01%
9	*unbelievable*	532	3.34%	131	*unverifiable*	7	0.04%	390	*unskillable*	1	0.01%
10	*unreliable*	476	2.99%	132	*unclimbable*	6	0.04%	391	*unslakeable*	1	0.01%
11	*unmistakable*	402	2.52%	133	*undrinkable*	6	0.04%	392	*unsliceable*	1	0.01%
12	*unavoidable*	394	2.47%	134	*unmarriageable*	6	0.04%	393	*unsoftenable*	1	0.01%
13	*unavailable*	377	2.37%	135	*unnavigable*	6	0.04%	394	*unspellable*	1	0.01%
14	*unthinkable*	351	2.2%	136	*unpleasurable*	6	0.04%	395	*unstemmable*	1	0.01%
15	*unbearable*	337	2.12%	137	*unputdownable*	6	0.04%	396	*unstereotypeable*	1	0.01%

which requires the variable slot to be filled by at least one letter). Table 1 reports three segments of the frequency list of the 424 types found: the 15 top-ranking items, followed by two sets of 15 items selected opportunistically from the middle range and from the hapax legomena of this pattern in this corpus. Frequencies range from 1,327 attestations for the form *uncomfortable* to a tail of as many as 189 hapaxes.

According to Stangl (2021), the phraseological template *what the N* started out in the late 15th century in the specific form *what the devil*, probably calqued from French *que diable*. Over time, the fixed form became variable in the nominal slot (see Figure 1), probably mainly due to euphemistic strategies for avoiding the taboo word *devil* (and later for dysphemistic strategies of promoting the use of taboo words).

As the figure shows, more and more new variants of the pattern were introduced, but all innovations have fairly strong formal and/or semantic associations to existing variants. For example, *deuce* and *dickens* are euphemistic substitutes for *devil*; their choice is motivated by formal similarity. *Hell* shows a strong semantic connection to *devil* and in turn gives rise to the formally associated substitute *heck*. The latter is a good candidate for serving as the motivation for *fuck* and formally and/or semantically related words derived from it. While Figure 1 lists frequent variants of the pattern, we will see that there are also rare and highly creative modifications. According to the Corpus of Contemporary American (COCA, Davies 2008-), the most frequent manifestations of the pattern are *what the hell, what the fuck* and *what the heck*, boasting normalized frequencies of 32, 14 and 3 occurrences per million words, respectively.

In the next section, I will give a highly condensed survey of the basic ideas of the Entrenchment-and-Conventionalization Model. In the remaining sections of the paper, I will discuss the questions raised above from the perspective of this model and probe its potential for contributing to a better understanding of creativity in word-formation and phraseology.

3 A Survey of the Entrenchment-and-Conventionalization Model

The EC-Model has the following basic design: What it tries to explain is the very nature of language (or grammar) itself, including the questions why languages have structure, why they are subject to variation, and why they keep changing. What it offers as an explanation is a systematic description of the way in which three components interact under the influence of various forces:

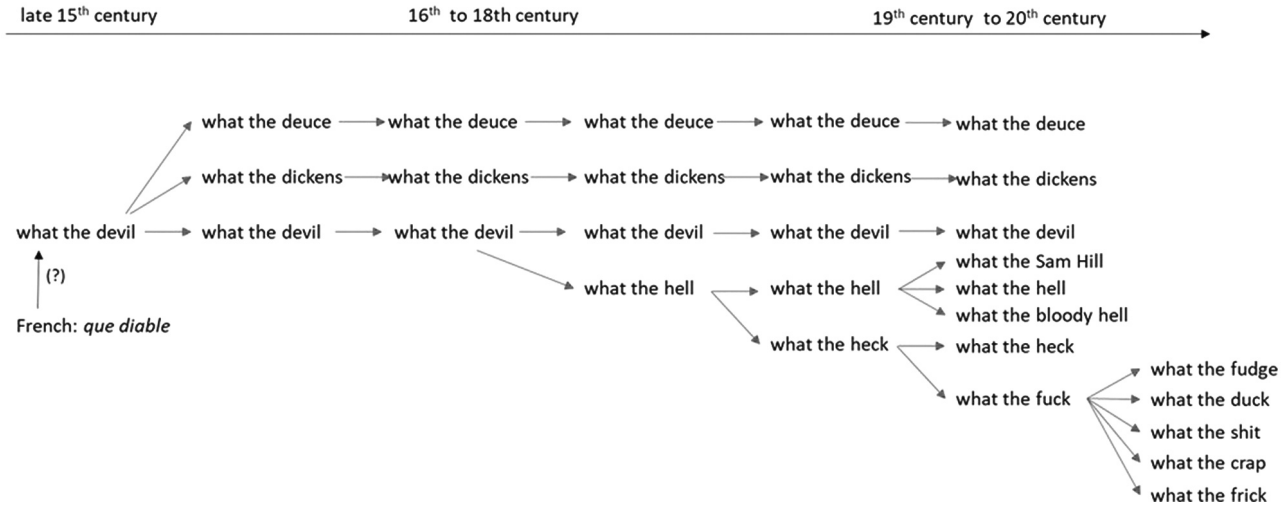

Figure 1: Rough sketch of the historical development of the pattern *what the N* (based on Stangl 2021).

– the use of language in concrete acts of communication,
– social conventionalization processes taking place in communities,
– and cognitive entrenchment processes taking place in the minds of speakers.

Put at its briefest, the model makes predictions regarding the question how usage, conventionalization and entrenchment interact to create and sustain linguistic structure, variation and change and how this interaction is influenced by internal and external forces.

In Schmid (2020), I explain the EC-Model with the help of the graph shown in Figure 2. The 'machine' shown in the figure is inspired by the work of the Swiss artist Jean Tinguely. Overall, it consists of an inner wheel, which is brought into motion by the communicative activities of speakers, and two outer flywheels driven by the inner wheel. The unorderly bundle of black circles in the centre of the figure represents the repetition of similar usage events, i.e. choices of linguistic elements and patterns instantiating conventional utterance types. Such repeated usage events can be token-repetitions of specific words such as *uncomfortable*, but also type-repetitions of a pattern underlying different words such as *unbelievable, unmarriageable* or *unskillable*. Even different tokens representing fully variable patterns such as N+N compounding count as repetitions of similar usage events from this perspective.

Whenever communication takes place, the energy produced in the inner wheel is transmitted to the two outer flywheels, which also begin to spin. This is meant to indicate that each act of communication has effects on the conventions shared by a community (represented on the left) and on patterns of associations in minds of the speakers present (represented on the right). Repetitions of words (*uncomfortable*) and patterns (*un* + V + *able*) will reinforce their conventionality and strengthen corresponding associations. In turn, the conventions shared by a community and the associations entrenched in the minds of speakers feed back into what happens when language is used, since conventions and routinized associations make communication possible in the first place. The whole system has the design of a double feedback loop centered in usage.

For the loop labelled "conventionalization feedback cycle", this means that the repetition of usage events feeds into and is in turn fed by the two conventionalization processes of usualization and diffusion. Conventions are not defined as linguistic units or rules, but instead as more flexible and probabilistic regularities of behaviour which are labelled *utterance types* (see Section 4 for more details).

The processes of usualization and diffusion establish and sustain these regularities in a feedback loop manner. Utterance types such as lexemes, patterns and schemas that are used in concrete acts of communication rely on the conventionality they already bring along due to their prior usage. With every repetition in

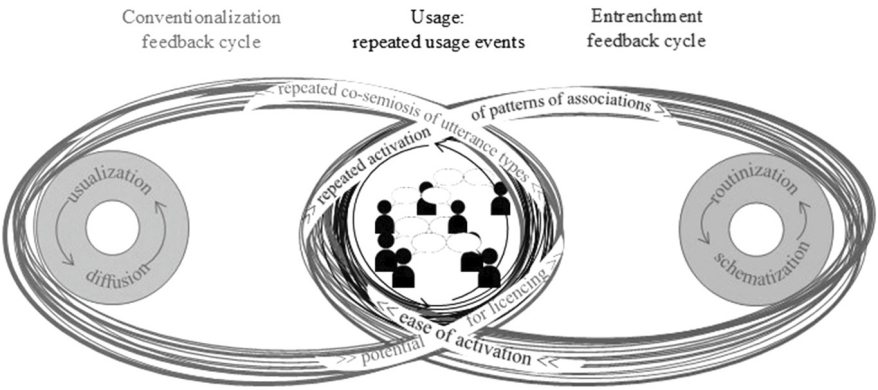

Figure 2: The 'Tinguely machine' model of language (adapted from Schmid 2020: 4).

an act of mutual belief of mutual understanding ("co-semiosis"), their conventionality is confirmed and increased, as their use provides yet another round of refreshment for their conventionality. The more frequently a given utterance type is repeated, the more firmly the underlying regularity becomes conventionalized by the processes of usualization (affecting the degree of regularity among speakers) and diffusion (affecting the spread of the regularity across speakers, groups and contexts). The frequent specific items listed in Table 1, e.g. *uncomfortable, unacceptable* and *unreasonable*, have become more conventional by repeated usage than the rare formations *unsingable, unshirkable* or *unskillable*. The frequent repetition of these and many more adjectives sharing formal and semantic properties sustains the conventionality of the productive semi-variable pattern *un-* + V + *-able*. The more conventionalized a given utterance type has become by past usage, the higher its potential for licensing future usage events. This is because conventional utterance types are more likely to be communicatively effective than unconventional ones. In this way, the conventionalization feedback cycle comes full circle.

The entrenchment feedback cycle works in an analogous way. It operates over patterns of linguistic associations which become active in the minds of those who are present in a usage event. For example, the form *uncomfortable* is symbolically associated with the meanings 'not having the quality of being comforting' and 'not at ease'. It is paradigmatically associated, formally and semantically, with its opposite *comfortable*, among others. Roughly speaking (see Section 8 below for more details), the pattern *un* + V + *able* is represented by a syntagmatic association between the three constituents *un-*, V and *-able* (with a variable slot in the middle) and a symbolic association to the meanings 'that can be V-ed, that has the quality of V-ing' (among other things, see Section 2). Patterns of associations

that are active in usage become routinized and strengthened by repetition, leaving increasingly strong traces and eventually serving as attractors in the associative network, i.e. as routines. The more frequently a given pattern of associations has been activated, the more strongly (or 'deeply') it becomes entrenched (i.e. represented) as a routine. The more strongly entrenched a routine has become, the more easily, effortlessly, automatically and thus also frequently it is activated in future usage events (Langacker 1987: 100; Schmid 2020: 43–72, 205–216).

Conventionalization in a community and entrenchment in the minds of individuals can only 'meet' and be tallied in usage. Conventionalization unfolds under the influence of social processes and forces (e.g. social order, prestige, power), but partly depends on degrees of entrenchment in the minds of the members of the community. Conversely, entrenchment is a cognitive process controlled by psychological factors such as salience, similarity or embodiment, but it is subject to the input supplied by the communities that a given speaker is a member of.[1]

4 What is the Base Line for Deciding what is or is not New or Original?

As mentioned above, the general notion of creativity crucially hinges upon the notions of novelty and originality. To decide whether or not, and to what degree, an utterance is original, we need a base line defining the nature of those utterances that are not original. The EC-Model suggests two candidates to serve as a base line: shared social conventions and individual cognitive routines. Creative utterances can either be considered original because they deviate from the conventions shared by the members of a speech community or because they are not part of the routinized repertoire of a given speaker. For instance, we can regard the word *unsexable* in the right column in Table 1 as creative, because it is not entered in the *OED*, suggesting that it is not a conventionalized solution for a communication task (if we take the dictionary as reflecting what is conventional). Alternatively, I could regard it as being creative because I believe I have never heard it before. Picking shared conventions as a base line has the advantage that we identify originality against the backdrop of language viewed as a socially distributed system of regularities of linguistic behaviour. The downside of taking

1 The interaction of the three main components of the model – usage, conventionalization and entrenchment – is subject to a range of forces which are neglected here.

large-scale, community-wide conventionality as a base line is that it is more difficult to explain that utterance types can be considered as perfectly conventional by speakers who use them frequently and at the same time as original by speakers who have not come across them.[2] Inter-individual variation of this type can be explained quite well if one takes routines as a base line for originality, but then we lose sight of the social dimension of language and have to deal with rampant subjectivity. Therefore, I would suggest to accept conventions as a base line, but keep routines in view to account for inter-individual and perhaps also social variation (see Section 8).

In Schmid (2020: 88), I define conventions as mutually known regularities of behaviour that the members of a community conform to because they expect each other to conform to them. Note that the concept of *regularity* is understood in a purely descriptive and empirical way here to identify observable similarities of utterances produced by different speakers. But what exactly is regular about speakers' linguistic behaviour in this sense? This question is usually answered by pointing to lexical entries, linguistic rules or formulaic expressions as prime examples of linguistic conventions (see, e.g., Clark 1996: 76–77). In order to provide a more fine-grained concept of conventions, I propose a multi-dimensional and probabilistic account and describe them in terms of five dimensions of regularity:

– *Onomasiological regularities*, i.e. repeated correspondences between communicative intentions and linguistic forms available for encoding them; e.g. to encode the communicative intention 'express the opposite of a quality', use the patterns *un-* + Adj (*unable, unlikely, unusual*) or *in-* + Adj (*independent, informal, inadequate*) or *non-* + Adj (*non-alcoholic*, *non-adjustable*, *non-existent*) (examples and semantic classification taken from Hansen et al. 1985).
– *Semasiological regularities*, i.e. repeated correspondences between linguistic forms and meanings; e.g. use the pattern *re-*V to encode the meanings 'V-again' (*rebuild, reopen, reconstruct*), 'V in a different way' (*reshuffle, reorganize, retell*), 'V . . . back' (*reimport, retranslate, repurchase*) (examples and semantic classification taken from Hansen et al. 1985).
– *Syntagmatic regularities*, i.e. regularities regarding the linear arrangement of linguistic elements; e.g. the patterns *un-* + V + *-able* and *what the* N, complementation patterns (*believe -> that/in*, *regard -> as, prone -> to*) or collocations (*make -> a difference/decision/point, take -> a picture/break/risk*).
– *Contextual regularities*, i.e. repeated correspondences between linguistic elements and situational factors, genres, registers, styles, etc.; e.g. the correspon-

2 This aspect is integrated in the EC-Model in the way that conventions are defined as being relative to communities of speakers (see Schmid 2020: 88–90).

dence of the routine formula *good morning* to the situational parameter 'time', the use of *POV* in social media posts or of the formula *I hereby* V in official statements with a declarative illocution.

– *Community-related regularities*, i.e. repeated correspondences between linguistic elements and social groups or variables captured by concepts such as *dialect*, *sociolect* or *genderlect*.

In principle, all five dimensions of regularity can be operationalized in terms of conditional probabilities derived from frequency counts in corpora. The most familiar one of these is the conditional probability in the syntagmatic dimension known as *transitional probability*, e.g. the forward transitional probability of the suffix *-able* given the prefix *un-*. Using frequency counts from the BNC to calculate this probability, we arrive at the value p(*-able*|*un-*) = c(un+able_[AJ0])/c(un*_[AJ0]) = 22,056/147,014 = 0.15. This means that 15% of the adjectives that begin with *un* end in *able* in the BNC, corresponding to a transitional probability of 0.15. The reciprocal backward transitional probability of the occurrence of the prefix *un-* given the suffix *-able* is slightly lower: p(*un-*|*-able*) = c(*un+able*_[AJ0]/c(+*able*_[AJ0]) = 22,056/170,115 = 0.13. Nevertheless, both types of transitional probabilities indicate strong degrees of syntagmatic regularity within the pattern *un-* + X + *-able*.

Transitional probabilities can also be calculated for each of the specific items found as bases in the variable slot of the pattern. One such measure is in fact listed in Table 1, where we can see that *comfort* accounts for 8.33% of all occurrences of the pattern *un-* + X + $able_{Adj}$, followed by *accept* (7.66%) and *reason* (6.14%). This measure has been called *attraction* (Schmid 2000: 54), since it reflects the degree to which the basis is – metaphorically speaking – attracted by the pattern. A reciprocal measure to be taken into consideration is *reliance*, i.e. the degree to which a word relies on the pattern for its usage. This measure can be calculated by taking into account the overall number of tokens of a given basis in the corpus, both in the pattern and elsewhere, and dividing the frequency in the pattern by this number. The three bases $comfort_V$, $accept_V$ and $reason_V$ behave very differently with regard to this measure: as verbs, *comfort* and *reason* occur 570 and 320 times respectively in the BNC outside the pattern *un-* + V + *-able*, the verb *accept* as many as 9,604 times. Relating this to their frequencies in the pattern as listed in Table 1, this means that *comfort* and *accept* rely on the pattern to 70% and 75% respectively, whereas *accept* does so only for 11% of its usage. Therefore, the degree of syntagmatic regularity for occurrence in the pattern is much higher for *comfort* and *reason* than for *accept*.

Similar measures can be applied in the domain of phraseology. According to the simple search `what the N` in the COCA, the attraction scores of the top-ranking

forms *hell* and *fuck* are 23% and 10% respectively. More extreme scores can be found for highly routinized patterns. For example, based on the BNC, the transitional probability of *mean* given the sequence *you know what I* is 0.76, reflecting the strong syntagmatic regularity and concomitant predictability of the phrase *you know what I mean*.[3]

In principle, conditional probabilities in the other dimensions of regularity work in an analogous way, but their application is encumbered by major practical obstacles: the frequency counts required for calculating conditional probabilities are much more difficult to come by. Schmid (2024) describes a method referred to as *Multi-Dimensional Regularity Analysis* and shows how it can be applied using simple descriptive statistics or more advanced vector space models. In theory, which is the focus of the present programmatic paper, onomasiological regularity is operationalized as the conditional probabilities of a range of linguistic forms suitable for encoding a given communicative intention. This requires counting communicative intentions – which remains a laborious manual task until we find an automatic way of annotating corpora for intentions. Semasiological regularity is operationalized as the conditional probability of one or more meanings given a linguistic form (see Schmid 2020: 90–92). In the idealized case that a given form has only one meaning, this probability would amount to 1. Perhaps the verb *procrastinate* with the sole meaning 'put off' is a case in point. For highly polysemous lexemes such as *get* or *run*, different meanings (e.g. 'obtain', 'become', 'turn' and many more for *get*) account for different proportions of their usage. According to Glynn's (2014) corpus study of the verb *run*, the meaning 'fast pedestrian motion' has a conditional semasiological probability of 0.32, followed by 'escape' (0.11), 'motion' (0.05) and 'fast motion', 'free motion' as well as 'execute' (0.04 each).

In this framework, the conventionality of the pattern *un-* + V + *-able* is described as a multi-dimensional probabilistic regularity configuration. The syntagmatic, onomasiological and semasiological dimensions of regularity dominate this configuration. Regularity in the syntagmatic dimension relates to the co-occurrence tendencies of the forms *un-*, V and *-able*. The onomasiological and semasiological regularity associated with this syntagmatic regularity was described in Section 2 above, where I listed typical structures and meanings of the pattern. The other two dimensions of regularity are less relevant in this case, because there are no obvious contextual or social regularities to be observed. In contrast, 'learned' suffixation patterns of Romance or Greek origin such as *-(a)tion*$_{N}$, *-ify*$_{V}$, *-ate*$_{V}$, *-ize*$_{V}$ or *-ism*$_{N}$ do

3 For the sake of transparency, I neglect the more sophisticated and realistic inferential statistical measures that can be used for assessing conditional probabilities, see Evert (2005), Stefanowitsch and Flach (2017: 115–116) and Schmid (2020: 52) for more information.

show contextual and social regularities, on top of syntagmatic, onomasiological and semasiological ones, as they are mainly used in formal genres and contexts (Schmid 2016: 181) and perhaps also more frequently by more highly educated speakers.

The inclusion of contextual and social regularities in the definition of conventions is also helpful for explaining words like *unsexable*, mentioned above and listed with a frequency of 1 in Table 1. Counter to my intuition and despite its not being entered in the *OED*, this word seems to be conventional in communities of practice revolving around fish and is used there to state that the sex of a given specimen cannot be determined.[4] This nicely illustrates that using the cognitive routines of a single speaker as a base line for creativity may be useful, since it helps to highlight differences in people's judgements of what is or is not creative (see Section 8 below).

The five dimensions of regularity chart a five-dimensional space. In theory, this space is populated by utterance types represented as five-dimensional vectors that reflect any type of regularity found in the complete historical record of the language or variety at hand.[5] In order to visualize this space for illustrative purposes (see Figure 3), we can conflate the onomasiological and the semasiological dimension under the label of *symbolic regularity* and the contextual and the social dimensions as *socio-pragmatic regularity*. The syntagmatic dimension can stay the way it is.

Highly regular utterances, i.e. utterances that have a usage history marked by high usage conformity, find their place near the origin (marked as "100% regular"), highly unconventional ones would be located towards the outer edges of the space. If we recall the prediction of the EC-Model that conventional utterance types are more effective than unconventional ones and assume that speakers value effectiveness highly, we can also assume that the historical record of utterance types is not distributed evenly in this space. Instead, we would expect a distribution of the type shown in Figure 3 (using toy, i.e. fictive data), which is intended to show a positive correlation between the density of vectors in the space and the degree of regularity (or conventionality) they indicate. The bubbles cluster in the bottom front left corner, indicating high conventionality, whereas the

4 The example in the BNC reads: "The stock consists of: one pair of Ps. lombardi, two Ps. socolofi (which I can't sex), one pair of Labidochromis ewarti, a pair of Labidochromis chizimulu and two Labidochromis caeruleus (unsexable)."

5 It should be noted that all five dimensions are internally multi-dimensional, in the sense that there is not really a single score or value for them, but instead a set of scores reflecting different aspects of regularity (see Schmid 2024 for details). The discussion of the syntagmatic dimension, for example, has provided an idea of the various types of transitional probabilities that are factored in. Likewise, for the social dimension, usage regularities with regard to numerous variables such as age, gender, education, social class and others can be taken into consideration.

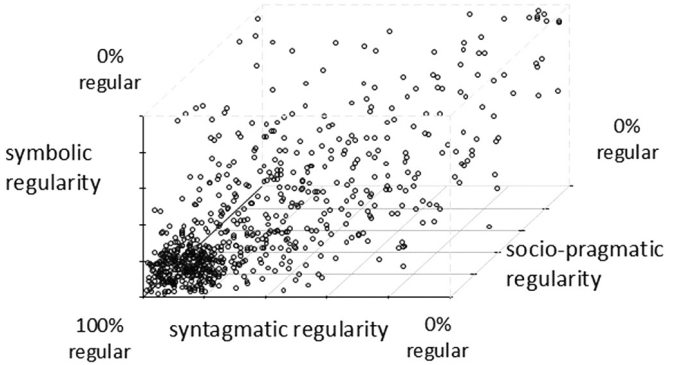

Figure 3: Fictive visualization of points in vector space representing utterance types in three-dimensional regularity space (bubbles stand for differently regular utterance types; the graph is not based on real data, but only serves illustrative purposes).

space towards the edges of the cube and the top right rear corner, containing utterance types marked by a low degree of regularity, is less densely populated. Obviously, it is in these remoter parts of the space that we have to start looking for creative uses of language.

5 What are the Ways in which Creative Utterances can Deviate from this Base Line?

5.1 Creativity in the Symbolic Dimension

Given the definition of conventions in terms of five probabilistic dimensions of regularity, we would expect that creative utterances can be non-conventional and thus original in one or more of these dimensions and to various degrees. Figure 4 visualizes what this means for symbolic (i.e. onomasiological and semasiological dimensions of) regularity. The bubbles located towards the ceiling of the cube representing multi-dimensional conventionality space indicate utterance types showing a low degree of symbolic regularity which remains, however, within the scope of existing usage variants. These utterance types correspond to cases of F-creativity in Sampson's (2016) terms. Cases of E-creativity are defined – in an idealized manner, i.e. taking into account the fuzzy boundary to F-creativity – by literally going 'out of the box', indicated by the four arrows, i.e. by going beyond the limits of what is conventional according to the historical record (see Bergs and Kompa 2020).

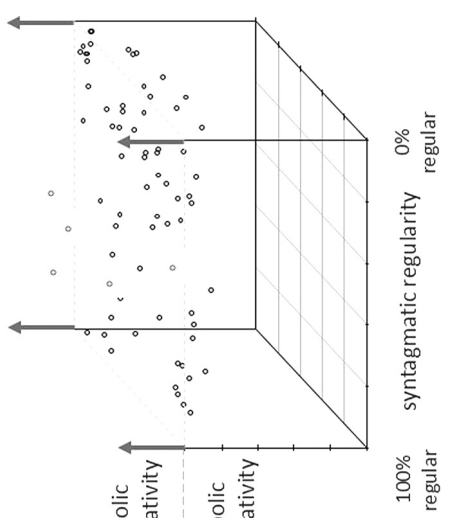

Figure 4: Symbolic F-creativity and E-creativity relative to three-dimensional regularity space (bubbles stand for differently regular utterance types; the graph is not based on real data, but only serves illustrative purposes).

5.1.1 Creativity in the Onomasiological Dimension

Speakers can be creative in the *onomasiological dimension* by selecting a novel linguistic form for encoding a communicative intention or concept (see Štekauer 1998, 2005). What will perhaps first come to mind here are those word-formation patterns that are sometimes labelled as *creative*, first and foremost blending. Coining novel blends such as *glampsite* or *hamdog* can be regarded as onomasiologically creative, since the coiner comes up with a new form for expressing an existing or new concept. Another relevant pattern is secretion, i.e. the use of non-morphemic parts of words such as *-gate, -athon, -aholic* or *-erati* as productive suffix-like formatives (Warren 1990; Schmid 2016: 130, 167; see also Szymanek 2005; Callies 2016). These, too, are onomasiologically creative, since they offer new ways of encoding ideas. The same goes for innovative compounds, i.e. compounds that combine existing lexemes in new ways. While established compounds show many regularities that have been described in terms of formal, structural and semantic patterns (see, e.g. Bauer, Lieber and Plag 2013; Schmid 2016; Warren 1978), root-compounding in particular following the patterns N + N and Adj + N is extremely flexible. As a consequence, on the one hand, there is enormous scope for F-creativity, meaning that speakers can create novel compounds whose originality ranges from hardly noticed to fairly salient. On the other hand, in order to exploit compounds for E-creativity, speakers have to come up with highly unconventional formations that are difficult to interpret even in context. The study by Ryder (1994) provides interesting invented examples such as *cow tree* or *hamburger shrub*. These were used as stimuli for an interpretation task designed to find out whether participants applied recurrent semantic schemas when confronted with novel N+N compounds. The situation is similar for the highly flexible process of conversion (see Clark and Clark 1979).

Regarding the use of derivational patterns such as *un-* + V + *-able*, onomasiological creativity can take place in the way that novel fillers are inserted in the variable slot of the pattern. As long as these fillers are in line with the regular characteristics of the pattern, such innovations would be considered as cases of F-creativity. Many of the hapaxes listed on the right of Table 1 can be seen as examples, e.g. *unrhymable, unscalable, unscrewable* or *unshavable*. Formations that deviate from the specifications more strongly and perhaps in a more salient way mark the gradual transition to E-creativity. The forms *unsleepable, unliveable* and *unhaveable* come to mind as examples, since they violate the regularity – usually formulated as a productivity restriction or constraint – that the verbal slot is more conventionally filled by a transitive verb than by other types of verbs (Bauer, Lieber and Plag 2013: 307). The intransitive verbs *sleep* and *live* and the weakly transitive verb *have* are less frequently found in the pattern and therefore

less regular and conventional. More extreme cases of deviation found in the BNC include, for example, the noun-based forms *unmarriageable, unknowledgeable*, *unstereotypeable* and *uncommonable* (used to encode the intention 'cattle which cannot be sent to the commons for grazing'). To extend the envelope of originality even further towards E-creativity, one could coin the adjective-based form *unhappyable* to convey the intention 'impossible to be, become or make happy'. As is shown in Figure 4, the pattern *what the N* is extended in an F-creative manner by formal derivatives of *fuck* such as *fudge* or *frick*. Omission of the *wh*-pronoun (e.g. *The fuck are right wing conservative Republicans doing on a gossip site*, COCA) and adjectival insertions (*What the fuckety fuck?*, COCA) can count as being more E-creative, as would the invented variant *what the Nancy Pelosi*. Readers should be reminded that due to the absence of clear-cut boundaries, all these examples find their place in a continuum of onomasiological regularity ranging from highly frequent via rare tokens to novel variants that show little connection to existing exemplars of the pattern.

From a mathematical point of view, the way in which I conceptualize and visualize E-creativity may well be considered slightly misleading. Figure 4 suggests that there could be regularity scores below 0, i.e. negative ones, which is mathematically weird. Conceptually, the visualization is also somewhat shaky, as it would indicate that a form such as *unhappyable* does not tap into any known symbolic regularity, which does not make an awful lot of sense either. One could indeed argue that any linguistic innovation that makes use of conventional linguistic raw material such as phonemes, graphemes or morphemes inevitably stays within the regularity space delimited by previous utterances (for a discussion of this issue in a rule-based framework, see Bergs and Kompa 2020). Keeping this in mind, the point of the illustration is indeed to suggest that at the time when they are uttered, E-creative uses relying on an existing pattern are characterized by the fact that they extend the realm that can be considered regular in the symbolic dimension. However, by virtue of the very fact that they have been uttered, they shift the boundaries of the regularity space and become part of the available regularity potential. If the word *unhappyable* caught on, i.e. if it got used by speakers and gradually became usualized and diffused, the bubble representing its vector would embark on a trajectory leading from the top-right to the bottom-left corner of the cube.

5.1.2 Creativity in the Semasiological Dimension

When speakers use existing forms with new meanings and in order to denote new referents, they stretch conventionality in the *semasiological* dimension of

regularity. Consider the lexeme *uncomfortable*, which is conventionally associated with the meanings 'not able to give comfort' (e.g. *He was on a very uncomfortable bed*, BNC) and 'not able to feel comfort, feel uneasy' (e.g. *He looked faintly uncomfortable*, BNC). A semasiological extension of these regularities could be to use the adjective in an utterance like *Little Anna is totally uncomfortable today* with the novel meaning 'it is impossible to comfort little Anna' rather than 'little Anna feels uneasy'. Speakers using this type of extension in an F-creative manner run a greater risk of compromising effectiveness, because it is uncertain whether the hearer actually gets the new meaning. In the long run, such acts of F-creativity can lead to a shift or extension of the semantic scope of a pattern, which we can also see in formations of the type *unsleepable, unhaveable* and also *unmarriageable*. These also extend existing semasiological regularities, as they loosen constraints on what the pattern *un-* + X + *-able* can mean. Semasiological creativity can be exploited in an E-creative manner with the intention to be recognized as a salient metaphor or deliberate pun or other form of humour.

5.2 Creativity in the Syntagmatic Dimension

All the examples of symbolic creativity provided so far – except for conversion and the semantic extension of *uncomfortable* – also involve creativity in the syntagmatic dimension, because existing morphological and lexical material is combined in a novel way. Regarding the pattern *un-* + X + *-able*, the most striking and isolated example is the invented *unhappyable* (see Figure 5). As a highly innovative combination of base, prefix and suffix, it lies outside the existing pool of syntagmatic regularities. Other interesting examples of syntagmatic E-creativity that are attested in the BNC include formations based on prepositional verbs such as *unputdownable* or *unturndownable*, as well as formations using complex nouns such as *unmarriageable, unknowledgeable, unpigeonholeable* and *unstereotypeable*. Such formations stand out because the existing record of syntagmatic regularity includes few examples of multi-word verbs and complex nouns as fillers of the variable slot.

5.3 Creativity in the Contextual Dimension

Essentially, what I have discussed so far under the label of symbolic and syntagmatic creativity has been part and parcel of research on productivity and creativity in word-formation. In contrast, creativity in the remaining two types of regularity has received less attention. Consider first creativity in the contextual

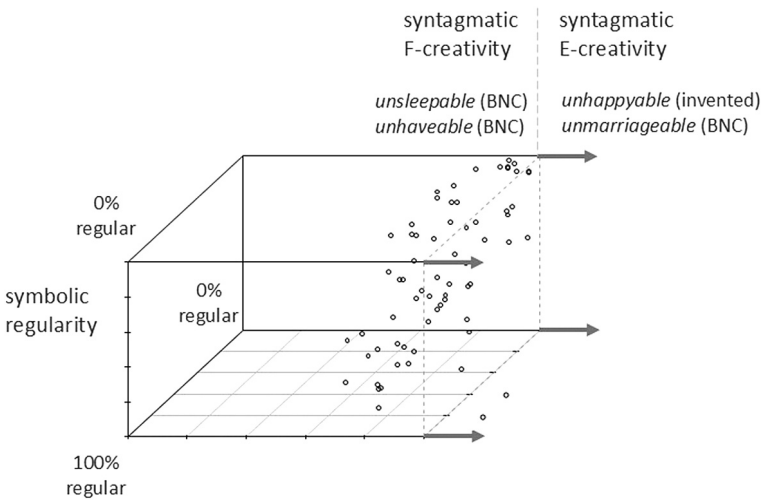

Figure 5: Syntagmatic F-creativity and E-creativity relative to three-dimensional regularity space (bubbles stand for differently regular utterance types; the graph is not based on real data, but only serves illustrative purposes).

dimension. As the pattern *un-* + V + *-able* is not sensitive to contextual and socio-pragmatic regularities, it cannot provide good examples here. The phrase *what the fuck* can, however. In the large, multi-million-word news section of COCA, this taboo phrase does not occur a single time, indicating a strong regularity in the contextual dimension. If a news anchor or reporter actually decided to use it, or if it was inadvertently recorded and broadcast, this would clearly be innovative and thus E-creative, since it would considerably extend the contextual regularity profile of this phrase. To add some more examples from the field of phraseology, when I use the time-sensitive routine formula *good morning* to greet a student who has just resurfaced from a deep slumber in an afternoon lecture, then I also perform a breach of a contextual regularity. While this would not be very funny, it can count as being creative. Another example is the – by now conventional – recontextualization of the phrase *thanks for coming to my TED talk* from the original TED talk situation to its use as an ironic meme on Tumblr and other social media platforms. Or when someone transfers the phrase *And the winner is* from its usual habitat, the announcement of the winner of an Oscar Academy award, to announce the winner of some private contest, then we can also argue that it is a case of creativity in the contextual dimension.

5.4 Creativity in the Community-Related Dimensions

Finally, if an underlying pattern is sensitive to the social or community-related dimensions of regularity, these can also be exploited for the full range from F-creativity to E-creativity. The news anchor example mentioned above is in fact a case in point, as it not only goes against contextual conventions regarding formality, register and style, but also conventions regarding social roles and relations. Imagine further an 85-year-old who alludes to a meme originating in youth slang by saying something like *one does not simply beat grandpa at ping pong*, or a speaker of American English mimicking a posh upper-class British accent. Although there is no novelty in form or meaning, such utterances can be considered E-creative, as they extend the contextual and/or socio-pragmatic regularity profiles of the given utterance types (see Figure 6).

As we have seen, utterances are often creative in several dimensions at the same time. The most extreme cases of F-creativity would combine low degrees of regularity in all dimensions, as is indicated in Figure 7.

Mathematically, these cases could be identified by looking for vectors representing very low scores for conditional probabilities in all dimensions. However, such cases are very unlikely to exist, because novelty in all dimensions of regularity would cause a major challenge for mutual understanding and thus run the risk of being ineffective. Multi-dimensional E-creativity is even less likely to occur, because it would essentially mean that someone produces an utterance that is not at all regular in any dimension of conventionality. Such formations do exist, however: they are referred to as *ex-nihilo creations* in the relevant literature and illustrated by examples such as *Kodak* or *Google*. And they are known, or at least said, to be extremely rare. What keeps speakers from indulging in such high degrees of creativity and what may encourage them to do so once in a while will be discussed in the next two sections.

6 How do Effectiveness and Appropriateness Constrain Creativity, and what other Forces Working Against Creativity can be Identified?

According to the definitions provided in the introduction, creativity is mainly constrained by communicative effectiveness and appropriateness. This is shown in Figure 8, where effectiveness is represented as a counter-force to creativity. The more speakers venture to leave the current regularity space, here illustrated for

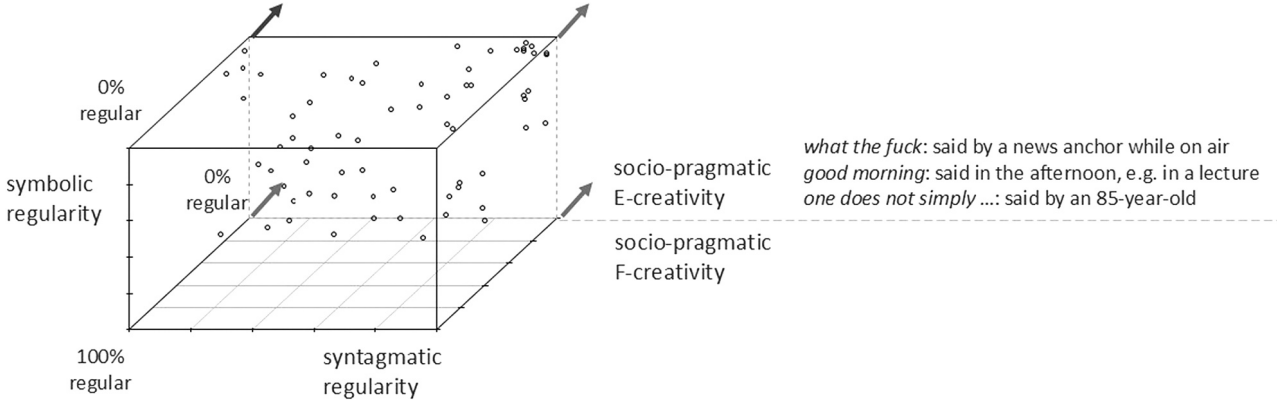

Figure 6: Socio-pragmatic E-creativity relative to three-dimensional regularity space (bubbles stand for differently regular utterance types; the graph is not based on real data, but only serves illustrative purposes).

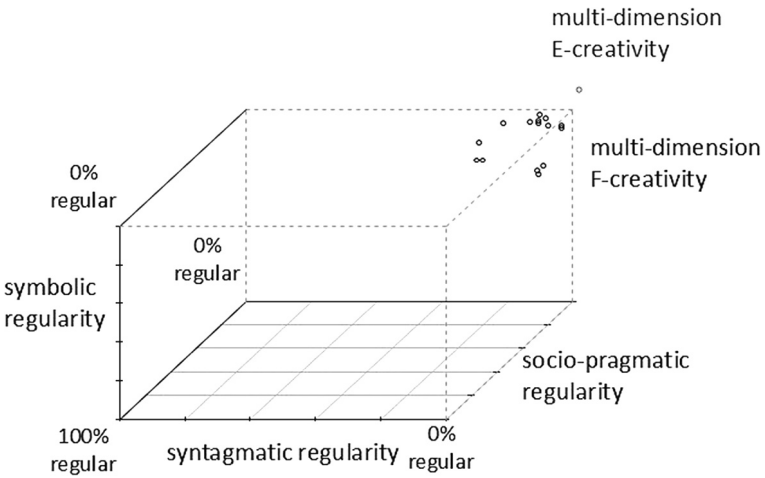

Figure 7: Extreme cases of multi-dimensional F-creativity and E-creativity relative to three-dimensional regularity space (bubbles stand for differently regular utterance types; the graph is not based on real data, but only serves illustrative purposes).

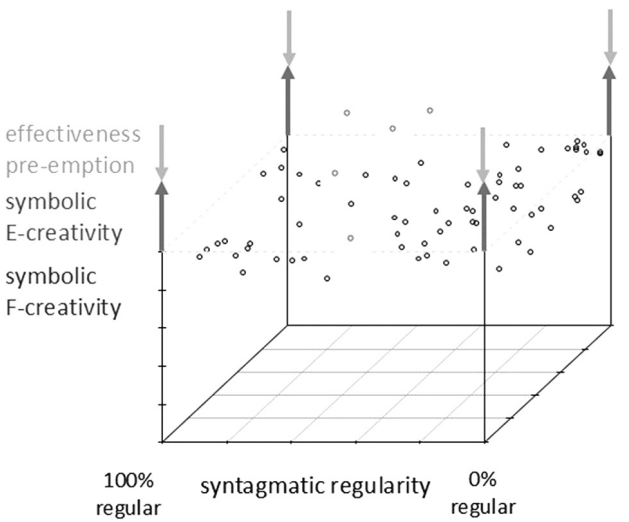

Figure 8: Effectiveness and pre-emption/blocking as forces constraining creativity.

the symbolic dimension, the more they risk compromising co-semiosis, i.e. the mutual belief of mutual understanding, and thus communicative effectiveness.

The notion of *appropriateness*, which was also mentioned in Section 1, is actually built into the system of multi-dimensional regularity by means of the contextual and socio-pragmatic dimensions. These dimensions capture the degree to which an utterance is situationally and socially appropriate in the sense of being correlated with and indexical of certain situational or social factors.

Effectiveness turns out to be an ambiguous notion when we take into account another factor constraining creativity, viz. the degree of unconventionality or distance to the conventional. So far, I have talked about effectiveness in terms of *communicative effectiveness*, associated with the goal of effortless and smooth mutual understanding. It is in this sense that effectiveness is a creativity-constraining force. However, intentional E-creativity tends to target the opposite type of effectiveness, one associated with extravagance, humour, fun or originality for originality's sake. Let's call this type of effectiveness *creative effectiveness*. As these two types of effectiveness produce opposite effects, speakers are forced to balance them out depending on their goals. On the one hand, when speakers aim to maximize mutual understanding in spite of their wish to be creative, they tend to minimize the degree of unconventionality. The data that we have looked at in this paper indicate that speakers do not venture far from the trodden paths. Rather, they tend to rely on strong formal, semantic or larger conceptual, frame-like similarities to reduce the distance to what is highly conventional and make themselves understood.[6] This is not only true of clearly F-creative formations, but also of those that reach out into the domain of E-creativity. On the other hand, when the explicit aim is to produce a strong creative effect, then the distance to the conventional must be larger and salient enough to be recognized. How speakers balance this out depends on their goals, but also on the context and other factors which may actually give more leeway to their creative inclination.

A further factor that constrains creativity is the prior existence of a highly conventional way of solving a given communicative task. This constraint is known as *blocking by synonym* or *pre-emption* (Clark and Clark 1979: 798, Schmid 2016: 116–117). Essentially, both notions capture the onomasiological phenomenon that speakers are much less likely to coin a novel composite, derived or converted form, e.g. *stealer*, if one or more conventionalized simple (*thief*) or complex forms (*robber, shoplifter, pickpocket*) are already available for expressing the target concept. In actual practice, it seems that speakers do not always adhere to the princi-

6 In case the reader may have wondered about *What the Sam Hill* when looking at Figure 1, *Sam Hill* is a conventional euphemism for *devil* based on formal similarity.

ples of blocking or pre-emption. This can have various reasons. First, the speaker may not be aware of the conventionalized solutions, which is frequently the case in first and second language acquisition. Second, the speaker may currently be unable to access and produce this solution, for various reasons. Third, and closest to the present concerns, the speaker may decide to overrule blocking/pre-emption in order to achieve a creative effect.

7 What are the Factors that can Loosen the Constraints on Creativity Imposed by Effectiveness, Appropriateness and Other Constraining Forces?

Speakers can indulge in creativity to achieve communicative effectiveness or creative effectiveness, whichever they strive for. The leeway they have is controlled by various factors. First of all, it is increased by the linguistic, social and cultural experience speakers share with their interlocutors. It is a general principle of communication that I can expect to communicate more effectively and with less effort when I talk to someone I have known for a long time. Therefore, the tolerable distance from 'the cube' increases in correlation with shared experience, common ground and mutual familiarity among the interlocutors (Figure 9). For instance, while perhaps 'over the top' in most everyday situations, a daringly creative utterance like *what the Nancy Pelosi are you doing here* might work perfectly well in a close-knit group of US politicians, in spite of the very low degree of similarity to existing variants of the pattern. In a group of colleagues working together in an administrative unit, even the variant *what the paperwork do you want me to do with this here* may work well enough.

Turning to the effects of context, I should begin by clarifying how the notions of *context* and *contextual regularity* are connected. *Contextual regularity* is a characteristic of conventions. This entails that it emerges cumulatively over repeated usage events which share certain contextual features. *Good morning* is regularly used in the morning, nouns ending in *-ism* are typically used in academic prose and more formal contexts, the word *unsexable* is typically used in contexts revolving around the topic 'fish'. These are contextual regularities. In contrast, *context* is a characteristic of situated usage events.

The situational, social and cultural context of utterances functions as a modulating factor for acts of creativity (see Figure 10). Contexts that deliver strong cues for understanding extend the scope for creativity. The same is true of informal

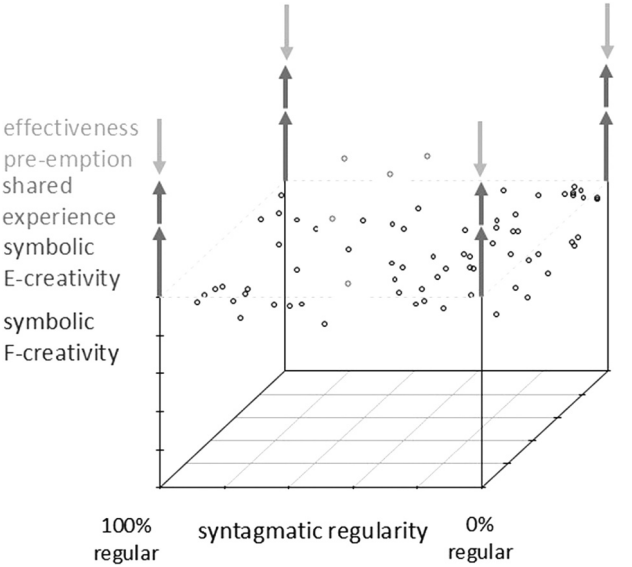

Figure 9: Shared experience as a force extending the scope of creativity.

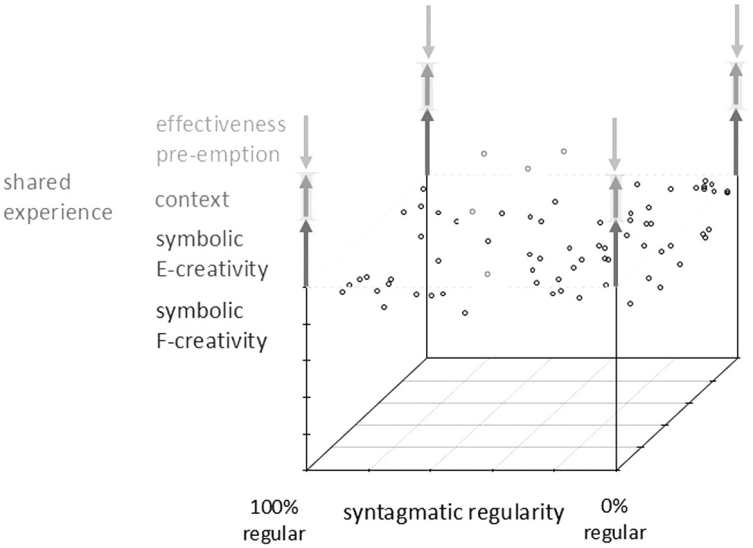

Figure 10: Current context as a factor modulating creativity.

contexts, where precise mutual understanding may not be the top-ranking goal of interlocutors, compared to other interpersonal and social functions of communication. Less constrained and formal contexts that require more precision reduce the leeway that speakers have. A casual conversation among friends in a bar is a better breeding-ground for audacious acts of creativity than a meeting of lawyers who discuss the details of a contract or a meeting of linguists who put the finishing touches on a co-authored paper.

8 What is the Role of the Cognitive Routines of Individual Speakers in the Production and Perception of Creative Utterances?

The notion of *routine* has surfaced in several places in the course of this paper. In the EC-Model, it is assumed that linguistic routines – and indeed linguistic knowledge in general – are subserved by patterns of associations which are extracted from the commonalities of usage events and routinized by repetition (see Section 3 above). Usage events are rife with commonalities – for the simple reason that they rely on the regularities discussed above. Communication requires conventions, conventions are regularities of behaviour manifested in commonalities of utterances, associations become routinized by repeated processing of such commonalities.

I distinguish between four types of associations, which largely correspond to the dimensions of regularity. *Symbolic associations* are activated and become routinized when processing symbolic (onomasiological and semasiological) regularities. *Syntagmatic associations* correspond to syntagmatic regularities. *Socio-pragmatic associations* are activated and become routinized while processing situational and social regularities. The fourth type of association are *paradigmatic associations*, which are active and become routinized as links to semantic and/or formal alternatives that are co-activated as competitors during processing.

All four types of associations become active in cooperation and competition during processing, and eventually routinized by repetition. This is what I mean when I talk about *patterns of associations*. Figure 11 provides illustrations of such patterns for the specific lexemes *uncomfortable* and *unstereotypeable*. Starting with the former, we see strong syntagmatic associations (marked by bold horizontal arrows) linking the constituents *un-*, *comfort* and *-able*. We can assume that these associations have become strengthened by the frequent repetition of the lexeme *uncomfortable*. This syntagmatic strengthening has a number of effects

which are summarized in the *syntagmatic-strengthening principle* (Schmid 2020: 236): As the syntagmatic links between the constituents are strengthened, the symbolic associations connecting the forms of constituents to meanings are weakened (indicated by thin broken vertical arrows), while the symbolic association connecting the whole sequence to its holistic meaning is strengthened (represented by the bold vertical arrow connecting the composite form *uncomfortable* with the meaning 'not at ease'). At the same time, and also as a consequence of syntagmatic strengthening, paradigmatic associations to semantic and formal competitors of the constituents are weakened (thin broken diagonal arrows), while paradigmatic associations of the composite form are strengthened (solid diagonal arrow pointing to 'unhappy', 'comfortable, 'sad', etc.). This means, e.g., that speakers do not have strong associations connecting the word *uncomfortable* with formal and semantic competitors of its base *comfort*, reflecting the high degree of lexicalization of the composite form. Socio-pragmatic associations are not rendered in the figure, as the word is not subject to any specific regularities here.

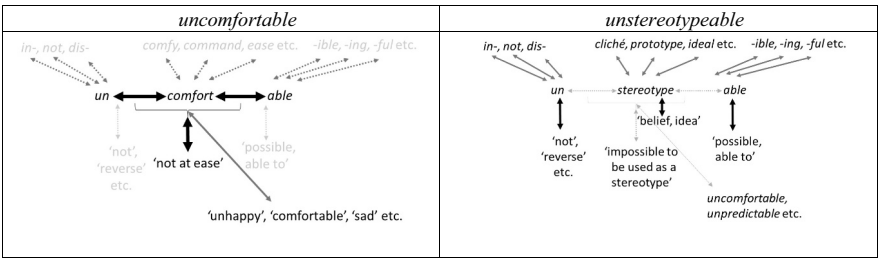

Figure 11: Hypothetical patterns of associations subserving the lexemes *uncomfortable* and *unstereotypeable.*

Shifting our attention to the representation for the lexeme *unstereotypeable*, we can appreciate the contrast between repetition-driven syntagmatic strengthening (on the left, *uncomfortable*) and a novel, or at least highly unconventional formation, which is therefore fully compositional (*unstereotypeable*, on the right). In the graph on the right, syntagmatic associations are weak, as they have not been strengthened by repetition. Symbolic associations of component parts are strong, since these parts are familiar, but the symbolic association of the composite form is weak, too. Paradigmatic associations can be activated to other, more familiar words sharing the internal makeup with *unstereotypeable*; paradigmatic associations that are specific to this word itself are unlikely to become active. Instead, and in contrast to *uncomfortable*, paradigmatic associations of the component parts remain strong.

The processing of the innovative, or at least very rare, word *unstereotypeable* can also rely upon a pattern of associations that has been gleaned from the commonality of all adjectives of the *un-* + V + *-able* variety and their meanings (see the left part of Figure 12). This commonality lies in the formal identity and semantic similarity of the forms *un-* and *-able* and the relational similarity (or analogy) of the verbal bases vis-à-vis this environment (shown on the left, vertical arrows representing paradigmatic associations here). Repetition of different, but in these respects similar forms and meanings leads to strengthening of the pattern of associations depicted on the right in Figure 12: syntagmatic associations involving a variable component; symbolic associations to the meanings shared by the tokens of this pattern; relatively strong paradigmatic associations to competing patterns and also to elements competing with constituents.

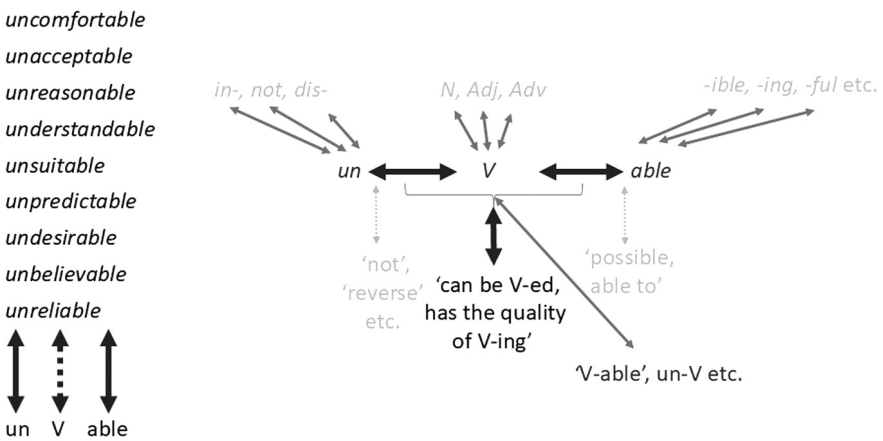

Figure 12: Hypothetical patterns of associations subserving the pattern *un-* + V + *-ing*.

The mental representation that is available in such a form is in many ways and for most practical purposes equivalent to a word-formation schema or morphological construction of the type suggested in network models (Bybee 2010), construction grammar (Diessel 2019, Goldberg 2019), or construction morphology (Booij 2010). One crucial difference, however, is that the associationist account is more flexible and hence less surprised, metaphorically speaking, to see that there are also some formations, like *unstereotypeable*, which use nouns as bases. This account of routines can also explain why *unstereotypeable* will be experienced as being creative by some speakers and can still be communicatively effective: we can assume that it is not in tune with most speakers' routines as a specific item, and it is not densely connected to other forms in the associative network, but still

similar enough to familiar forms and meanings and to a more general pattern of associations to be processed successfully.

Given their very nature as linguistic habits, routines generally stand in the way of creativity. Under normal circumstances, we tend to do what we always do, and what we always do tends to be in line with existing conventions. Therefore, we can treat routines as factors that constrain creativity. This is not to say that speakers are reluctant to stray from the trodden paths in their minds. In fact, they do so for many reasons, not only when they want to achieve creative effectiveness, but also, for instance, when they are facing a new communicative task, when they want to get across an original idea or simply because they enjoy novelty for novelty's sake. As routines are speaker-specific by definition, the perspective on routines also entails that one speaker or a group of speakers, e.g. the members of a family or closely knit community of practice, can have a highly routinized representation of a linguistic element or pattern that they encounter and use frequently. In contrast, another speaker, e.g. an outsider, has no specific representation, but must rely on similarities to other items and more general entrenched routines – which comes down to experiencing a given processing event as being 'creative'. In this way, routines can be considered as sources of speaker-specific assessments of what is creative, both in production and comprehension. If I go out of my entrenched ways of expression and coin something novel, only to find out that others consider my novel formation as perfectly conventional, then I can still feel justified in claiming that I was creative.

That said, it is not the case that speakers only accept as conventional what they themselves have represented as routines. We are perfectly able and happy to accept ways of speaking that lie outside our own box, because we are used to the fact that others speak differently (depending, for example, on their social background and corresponding differences regarding linguistic biographies). Furthermore, entrenchment is not as solipsistic as it often seems (Schmid 2022). Speakers do not entrench routines in a social vacuum, but instead extract all their linguistic knowledge from interpersonal communication, which inevitably comes complete with an interpersonal encounter (if only between writer and reader) and a social event. Part of the repetition-driven knowledge we entrench is information about what other speakers regularly do, what these speakers are like and how they behave in different types of situations. As a result, we are used to not considering utterances as creative, even if they lie outside the scope of what we have routinized.

9 Conclusion

I started out by asking the question what is creative to whom and why. Given the discussion in the preceding pages, the answer to this question might read as follows:

Linguistic creativity is linguistic novelty, defined and operationalized as a low degree of regularity in conventional regularity space charted by multi-dimensional conditional probabilities,

- which is constrained by
  - communicative effectiveness, defined as degrees of mutual understanding,
  - distance from what is highly conventional, defined by location in multi-dimensional regularity space,
  - pre-emption, defined as onomasiological competition within multi-dimensional regularity space,
  - and speakers' routines, defined as differentially entrenched patterns of associations,
- afforded and promoted by
  - the availability of variable patterns of regularity, defined by multi-dimensional conditional probabilities,
  - creative effectiveness, defined as distant, intentional and ostentatious novelty,
  - long-term shared experience,
- and modulated by current context.

This definition of linguistic creativity is very general. As noted in the introduction, the domains of word-formation and phraseology provide a special breeding-ground for creativity by offering mutual competition and support between lexically specific complex items and expressions, on the one hand, and partly or fully schematic patterns, on the other. Creativity, especially strong forms of creativity labelled as *E-creative,* are achieved and recognized more easily, the more morphologically and lexically specific and fixed its conventional source is. Modifications of highly fixed proverbs or idioms are cases in point. A large part of the raison d'être of partly specific morphological patterns such as derivations and phraseological templates lies in the F-creative potential that speakers can exploit to react to changing communicative demands in a fairly innocuous way. At the same time, these highly productive patterns open up ways of being E-creative in the form of more daring and extreme new formations while keeping up comprehensibility. Due to their open form, the lexically unspecific word-formation patterns of

conversion and N + N compounding or Adj + N compounding offer enormous scope for F-creativity. Speakers who plan to exploit these patterns for E-creativity have to come up with fairly extreme variants, simply because the space of possibly conventional variants is so large.

References

Bauer, Laurie. 2001. *Morphological productivity*, Cambridge: Cambridge University Press.

Bauer, Laurie, Rochelle Lieber & Ingo Plag. 2013. *The Oxford reference guide to English morphology*. Oxford: Oxford University Press.

Bergs, Alexander. 2019. What, if anything, is linguistic creativity? *Gestalt Theory* 41(2). 173–183.

Bergs, Alexander & Nikola Anna Kompa. 2020. Creativity within and outside the linguistic system. *Cognitive Semiotics* 13 (1).

Booij, Geert. 2010. *Construction morphology*. Oxford: Oxford University Press.

Bybee, Joan L. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.

Callies, Marcus. 2016. Of soundscapes, talkathons and shopaholics: On the status of a new type of formative in English (and beyond). *STUF – Language Typology and Universals* 69 (4). 495–516.

Clark, Eve V. & Herbert H. Clark. 1979. When nouns surface as verbs. *Language* 55 (4). 767–811.

Clark, Herbert H. 1996. *Using language*. Cambridge: Cambridge University Press.

Davies, Mark. 2008-. *The corpus of contemporary American English (COCA)*. Available online at https://www.english-corpora.org/coca/.

Diessel, Holger. 2019. *The grammar network: how linguistic structure is shaped by language use*. Cambridge: Cambridge University Press.

Evert, Stefan. 2005. The statistics of word cooccurrences: word pairs and collocations. Stuttgart: PhD thesis Universität Stuttgart.

Glynn, Dylan. 2014. The many uses of *run*. in Dylan Glynn & Justyna A. Robinson (eds.), *Corpus methods for semantics: quantitative studies in polysemy and synonymy*, 117–144. Amsterdam & Philadelphia: Benjamins.

Goldberg, Adele E. 2019. *Explain me this: Creativity, competition, and the partial productivity of constructions*, Princeton, NJ: Princeton University Press.

Hansen, Barbara, Klaus Hansen, Albrecht Neubert & Manfred Schentke. 1985. *Englische Lexikologie. Einführung in Wortbildung und lexikalische Semantik*. 2. Aufl. Leipzig: VEB Verlag Enzyklopädie.

Hoffmann, Thomas. 2022. Constructionist approaches to creativity. *Yearbook of the German Cognitive Linguistics Association* 10 (1). 259–284.

Plag, Ingo. 1999. *Morphological productivity: Structural constraints in English derivation*. Berlin & New York: Mouton de Gruyter.

Plag, Ingo. 2004. Syntactic category information and the semantics of derivational morphological rules. *Folia Linguistica* 38 (3–4). 193–225.

Runco, Mark A. & Garrett J. Jaeger. 2012. The standard definition of creativity. *Creativity Research Journal* 24 (1). 92–96.

Ryder, Mary Ellen. 1994. *Ordered chaos. The interpretation of English noun-noun compounds*. Berkeley etc.: University of California Press.

Sampson, Geoffrey. 2016. Two ideas about creativity. In Martin Hinton (ed.), *Evidence. Experiment and argument in linguistics and philosophy of language*, 15–26. Bern: Peter Lang.

Schmid, Hans-Jörg. 2015. A blueprint of the Entrenchment-and-Conventionalization Model. *Yearbook of the German Cognitive Linguistics Association* 3 (1). 3–25.
Schmid, Hans-Jörg. 2016. *English morphology and word-formation: An Introduction*. 3rd edn. Berlin: Erich Schmidt.
Schmid, Hans-Jörg. 2020. *The dynamics of the linguistic system. Usage, conventionalization, and entrenchment*. Oxford: Oxford University Press.
Schmid, Hans-Jörg. 2022. Social and cognitive foundations of linguistic variation: towards a unified complex-adaptive framework. Talk presented on *Abralin Ao Vivo*, available at https://aovivo.abralin.org/lives/hans-jorg-schmid/
Schmid, Hans-Jörg. 2024. Multi-dimensional regularity analysis: How the Entrenchment-and-Conventionalization Model can be applied to corpus data. *Yearbook of the German Cognitive Linguistics Association* 11 (1). 23–56.
Stangl, Joshua. 2021. Investigating the *wh-word the X* construction. Munich: LMU BA Thesis.
Stefanowitsch, Anatol & Susanne Flach. 2017. The corpus-based perspective on entrenchment. In Hans-Jörg Schmid (ed.), *Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge*, 101–127. Boston & Berlin: APA and Walter de Gruyter.
Sternberg, Robert J. & Todd I. Lubart. 1999. The concept of creativity: Prospects and paradigms. In Robert J. Sternberg (ed.), *Handbook of creativity*, 3–15. Cambridge: Cambridge University Press.
Štekauer, Pavol. 1998. *An onomasiological theory of word-formation*, Amsterdam & Philadelphia: John Benjamins.
Štekauer, Pavol. 2005. Onomasiological approach to word-formation. In Pavol Štekauer & Rochelle Lieber (eds.), *Handbook of word-formation*, 207–232. Dordrecht: Springer.
Szymanek, Bogdan. 2005. The latest trends in English word-formation. In Pavol Štekauer & Rochelle Lieber (eds.), *Handbook of word-formation*, 429–448. Dordrecht: Springer.
Tin, Tan Bee 2022. *Unpacking creativity for language teaching*. London & New York: Routledge, Taylor & Francis Group.
Uhrig, Peter. 2018. I don't want to go all Yoko Ono on you. *Zeitschrift für Anglistik und Amerikanistik* 66 (3). 295–308.
Warren, Beatrice. 1978. *Semantic patterns of noun-noun compounds*. Gothenburg: Acta Universitatis Gothoburgensis.
Warren, Beatrice. 1990. The importance of combining forms. In Wolfgang U. Dressler, Hans Christian Luschützky, Oskar E. Pfeiffer & John R. Rennison (eds.), *Contemporary Morphology*, 111–132. Berlin & New York: Mouton de Gruyter.