Table of Contents

Against a non-process-based definition of competence

Baiba Vilks, Toronto

Chomskyan syntacticians have always insisted that the aim of transformational generative syntax is emphatically not to describe the causal processes that lie behind the production of a sentence, as originally prescribed by Chomsky himself. It is asked whether we ought not to abandon this arbitrary dictate. It is shown how the dictate has allowed the current instantiation of Chomskyan syntax, the Minimalist Program, to uphold a model that is incompatible with a realist, process-based understanding of how grammar works. Furthermore, the dictate has allowed generative syntacticians to ignore psycholinguistic evidence that should otherwise bear strongly on their theorizing. It is suggested that abandoning the Chomskyan non-process-based definition of competence could reopen the door to explorations in syntax that revert to the generative semantics of the 1960's and 1970's, rather than Chomskyan interpretive semantics.

Keywords*: competence, performance, transformational generative grammar, generative semantics, interpretive semantics, scientific realism*

1. General considerations

Ever since the middle of the twentieth century mainstream transformational syntax has subscribed to the dictate that in studying syntax the aim of investigators is not to describe the process by which language, say, a sentence, is produced, but rather to assess the competence of a speaker as it exists, as it were, in a snapshot in time:

> *A generative grammar is not a model for a speaker or a hearer…When we speak of a grammar as generating a sentence with a certain structural description, we mean simply that the grammar assigns this structural description to the sentence. When we say that a sentence has a certain derivation with respect to a particular generative grammar, we say nothing about how the speaker or hearer might proceed, in some practical or efficient way, to construct such a derivation. These questions belong to the theory of language use – the theory of performance. No doubt, a reasonable model of language use will incorporate, as a basic component, the generative grammar that expresses the speaker-hearer's knowledge of the language; but this generative grammar does not, in itself, provide the character or functioning of a perceptual model or a model of speech production* (Chomsky 1965: 9).

This dictate issued by Noam Chomsky in his view entails a key distinction between what is termed *competence* as opposed to *performance.* But note that this dictate is not, in fact, that which necessarily distinguishes between competence and performance, according to a common, contemporary understanding of these terms. Thus competence, the inner stored system of mental grammar can, in contrast to Chomsky's view of it, be seen to entail processual or causal aspects while still being distinct from performance, the more superficial cognitive mechanisms that execute the application of these inner grammatical representations.

Chomsky's dictate mandating a non-causal, non-processual system of inner grammatical representations has governed the party line of Chomskyan generative syntax from its beginnings through its evolution into *Government and Binding Theory* (GB) down to its current mainstream form as *Minimalism*. Thus, observe Chomsky's own comment in *The Minimalist Program* (1995: 380, note 3) coming about three decades after the above quotation:

> *Recall that the ordering of operations is abstract, expressing postulated properties of the language faculty of the brain, with no temporal interpretation implied. In this respect, the terms output and input have a metaphorical flavor, though they may reflect substantive properties, if the derivational approach is correct.*

This attribution of complete temporal abstraction to the notion of competence goes hand-in-hand with Chomsky's corollary that the purpose of a grammar is essentially to be able to generate all and only the grammatical strings of a language (Chomsky 1957: 49). Unfortunately, however, strict adherence to the dictate entailing only non-processual competence has resulted in theories of syntax that are notably *incompatible* and at odds with anything that could ever resemble a scientifically *realist*,[1] process-based model of how syntax and semantics function.

However, one can ask if we must continue to abide by this dictate that so severs the doing of syntax from a realist understanding of how grammar actually works. What binds us to this seemingly arbitrary notion that originates only with Chomsky and the independent scientific merit of which is dubious? Thus far this dictate has functioned as the loophole that has been used by syntacticians as an excuse for the fact that we don't yet have a transformational, generative syntactic theory that provides a model of how human language functions which can account for the causal relations involved. Thus many a syntax professor has been heard to explain that syntax does *not* consist in trying to construct a model of what we are actually doing when we utter a sentence – our apparent aim in syntax is emphatically *not* to describe how a sentence is produced in the brain (professors of the University of Toronto and York University Linguistics departments, e.g., Elizabeth Cowper, Alana Johns and Daniel Currie Hall, p.c.). But why should we continue to exalt this prescript? Perhaps if we were to abandon Noam Chomsky's dictate that a generative syntax is not supposed to model how language functions in a causal, processual manner we would be free at last to discover how we get from thoughts to words, that is, how human syntax truly works. While discovering how syntactic processes are computed in real time at the *micro* level of cells and chemicals in our brains – the less moderate form of realism referred to by Seuren as *hardware* realism (2004: 61) – does not have to be the goal of syntactic theory, at the very least the product of our investigation should be *compatible* with causal, processual models in the more general realist sense. Surely if any process model is of necessity *irreconcilable* with our syntax we are on the wrong track and in error.

---

[1] A *realist* scientific theory attempts to elucidate the surface, observable facts of some phenomenon with reference to an underlying reality, seeking to explain causally how the underlying reality leads to the coming about of the surface facts. An *interpretivist* theory, on the other hand, merely attempts to find regularities in the surface phenomenon without attempting to understand how an underlying reality (if one is assumed to exist) causes their manifestation (cf. Seuren 2004: 61-62).

The above characterization of competence is necessarily to be distinguished from a broader understanding of the notion and is only one of the diverse aspects of Chomsky's own definition of it. Thus, competence also entails the knowledge underlying language that is divorced from things such as memory limitations and speech errors and repairs, and which assumes an idealized speaker in a uniform speech community. Chomsky himself explicitly ascribes such characteristics to the notion of competence, and in addition distinguishes between competence as "the speaker-hearer's knowledge of his language" as opposed to performance as being "the actual use of language in concrete situations" (Chomsky 1965: 4). This much can be easily accepted; however, expunging in addition all causal processes from the notion of competence as Chomsky does is what has played such a large part in contributing to the aforementioned deficit in contemporary generative syntactic theory.

Note that as mentioned above there is a distinction to be made between the difference between competence and performance, on the one hand, and between a static, non-process understanding of competence proper as opposed to an understanding that allows for this inner stored system of mental grammar to display process-based aspects, on the other. Thus, to say that the definition of competence – "the formal properties of our proposed linguistic representations" according to Bresnan and Kaplan (1982: xxii) – should entail notions of process is not to argue against distinguishing between competence and performance. The latter is defined by Bresnan and Kaplan (1982: xxii) as "the cognitive processes that derive and interpret [competence] in actual language use and acquisition". As one anonymous reviewer has pointed out, numerous authors have attempted to explain apparently grammatical (competence-based) phenomena in terms that are rooted in performance.[2] The present paper, however, is not suggesting that competence phenomena ought necessarily to be accounted for or supplanted by performance mechanisms but rather that Chomsky's understanding of competence needs to be broadened so as to allow it to be defined in a causal, processual manner.

Furthermore, the precise intention behind Chomsky's non-process-based definition of competence remains unclear. Jackendoff (2002: 56) attempts to characterize it this way:

> *A hearer presumably constructs [a sentence] first by deriving a phonological structure from the auditory signal and then using that to arrive at the syntactic and conceptual structures. But a speaker presumably starts out with a meaning to express and develops a syntactic and phonological structure from it. So although the structure is the same, the hearer and speaker produce its parts in different orders. This is, I think, the essential*

[2] Hawkins for instance has suggested that performance considerations such as parsing efficiency might account for what might otherwise be attributed to grammatical parameters rooted in a competence system (see, e.g., Hawkins, 2009). Sag et al. (2009) provide evidence that supports a processing/performance-based account of subjacency effects over a grammar/competence-based one. And O'Grady (2001; 2005) attempts to account for syntactic phenomena such as phrase structure design and pronoun interpretation this way. Thus O'Grady explains their characteristics chiefly with reference to underlying performance-related pressures -- primarily the drive of the mental processor to reduce the burden on working memory -- rather than according to principles of grammar.

> *difference between competence theories and processing theories. Competence theories are concerned with what the total structure is for either speaker or hearer. Processing theories are concerned with how the structure is built in real time, so they naturally bifurcate into different theories for the speaker and the hearer*.[3,4]

But who is to say, without independent justification, that the "total structure for either speaker or hearer" can in fact be divorced from "how the structure is built in real time"? And as Seuren (2004: 64) also wonders, what exactly does Chomsky mean to say with a comment like that in the second quote above that "the terms output and input have a metaphorical flavor"? Moreover, Chomsky himself at times strays from his own strictly defined notion of competence, or at least, leaves one confused as to whether he really intends to pursue a model that is completely abstract and atemporal. Thus Lenerz (1998: 110) has this to say in regards to Chomsky's MP:

> *This (derivationally conceived) computation takes place as a logical structure outside time and space (cp. p. 380, note 3). The question remains to what extent one can speak here of a description of the biologically given human language faculty, which can only be a concrete system bound in time and space. The abbreviation C_{HL} ("computational system of human language"), occasionally used by Chomsky, is therefore at least problematic, as long as the unclarity persists regarding the connection between the logico-deductive descriptive system and the concrete biological-mental system. One also wonders what could possibly justify the imposition on a logical system outside time and space of both economy criteria and the specific criteria derived from the performance systems A-P [Articulatory-perceptual system] and C-I [Conceptual-intentional system] that are external to the computation system, which, in the end, only mediates between them.*
> (Translation by Seuren 2004: 63-4)

In spite of such confusion engendered by Chomsky's definition of competence, transformational syntacticians continue to allow his non-process competence dictate to hinder the acknowledgement of fundamental flaws underlying theories such as GB and the *Minimalist Program* (MP). This also holds them back from seeking to establish a model that is at least compatible with moderate scientific realism.

2. Empirical considerations

2.1 *The necessarily derivative nature of case conflicts with the notion of "numeration" in the MP*

It can easily be seen that Chomsky's latest syntactic theory, the MP, has not led to an understanding of grammar that is at all compatible with a realist, processual model, as an example can demonstrate. One aspect of the MP that makes such a claim particularly evident is how it treats the phenomenon of Case and how this relates to its notion of what is termed the

[3] Note that if, as Seuren (2004: 83-85) suggests, comprehension proceeds via a *reconstruction-by-hypothesis* method, the differentiation underlying such a portrayal would be moot.
[4] Note that Jackendoff, unlike the present author, is using the term 'processing' to essentially mean 'performance'.

numeration. In the MP the numeration is the starting point for each derivation and consists of a list of all the lexical items that enter into a derivation *fully inflected*, along with an index for each one signifying how many times each lexical item appears in the numeration (Chomsky 1995: 225-238; Hornstein et al. 2005: 68-71). Each potentially Case-bearing lexical item enters the numeration with its Case already specified; in fact, according to Chomsky (1995: 237), the Case for each Case-bearing item gets specified as it is selected from the lexicon and deposited into the numeration. This pre-specification of Case is in fact an indispensable aspect of the MP model. Itallows diverse surface forms of Case to manifest in the phonological representation while being able to dispense with the representational level of *Surface Structure* (SS). SS was the level at which at which in GB, the version of Chomskyan syntax that preceded the MP, Cases were necessarily *assigned* to non-Case-marked DPs. Thus, now in the MP these already pre-assigned Cases merely have to be *checked*, either overtly or covertly, by the level that is understood to interface with the semantics, that of *Logical Form* (LF) (Chomsky 1995: 197; Hornstein et al. 2005: 111-112).

Such pre-specification of Case entails that Case is in no way seen to be a *derivative* phenomenon in the MP, dependent on the syntactic arrangement of items that surrounds it. This is in fact stated explicitly by Chomsky (1995: 237): "It is hardly plausible that Case and phi-features of book are determined by its position in a clausal configuration…Case and phi-features are added *arbitrarily* as a noun is selected for the numeration [emphasis mine]".[5] However, though phi-features do not derive from a syntactic configuration and are inherent to a noun, it is clear from a look at Case as it functions in a context known as *exceptional Case marking* (ECM) in English that it is necessarily a derived phenomenon that depends on the grammatical geography it finds itself in and thus cannot adhere to a lexical item pre-specified. Thus, in sentences (1) and (3) below, the masculine pronoun *he*, which functions as the agent and subject of the verb *go*, displays the nominative Case. In sentence (2), however, the masculine pronoun with the same semantic role and once again occupying the subject position of the clause with the verb *go* – but this time in a slightly different configuration – surprisingly displays the accusative Case. As the Case of a DP here clearly differs depending upon its syntactic context and the syntactic constituents that abut it, it can be seen that Case is a derivative phenomenon as far as the syntactic derivation goes and cannot be prespecified. Thus, in real time Case must be selected at a point that follows whatever enters the derivation as its starting point. One can only conclude that Case must derive from and not directly feed the structural syntactic manipulation that goes on during a derivation. Thus, at least GB (Chomsky 1981) was more on the right track in positing Cases that are assigned as a result of syntactic computation, rather than pre-assigned and merely checked later either overtly or covertly.

[5] Chomsky justifies this position by pointing out that nouns can be used in isolation in which case their Case and phi-features will be "fixed one way or another, though there is no structure" (1995: 237). Note though that while the phi-features of a noun in isolation don't imply any particular clausal structure, a certain syntactic structure is implied by the Case of a noun in isolation. Thus, if one says "*him*", in the accusative Case, as opposed to "*he*", in the nominative, in isolation the default implication is that the accusative pronoun constitutes a syntactic object that is missing a subject; likewise, genitive "*his*" in isolation evokes an undetermined possessive construction of some sort.

(1) ***He*** is going to the movies on Saturday.
(2) I expect ***him*** to go to the movies on Saturday.
(3) I expect that ***he*** is going to the movies on Saturday.

It would be impossible for the computation to 'know' which Case to assign to which nominal *before* further syntactic manipulations take place -- the only way it could "know" this if there were another, entirely separate, computation that preceded the main syntactic computations from which the numeration itself could be derived. But such a scenario would amount to in essence having an entirely distinct, *pre-syntactic syntax* that took place before the "main" derivation – a far from efficient or minimalistic situation!

Another example demonstrating the necessarily derivative nature of Case is found in languages like Latvian, where the Case on a nominal occurring with a preposition differs depending on the preposition; moreover, in Latvian the same preposition will assign a different Case depending upon whether the nominal is singular or plural. This shows that Case is a derivative marking that occurs depending upon the structure of multiple words in concert, and couldn't be "known" by the syntax ahead of the concatenation of the entire prepositional phrase and hence couldn't be pre-assigned in a numeration.

It is obvious furthermore from split-ergative languages that nominals can't enter the derivation with their Cases prespecified, as any given nominal won't be able to 'know' which Case it is supposed to surface with until other aspects of the combined syntax have been determined. Thus, in a language for instance where the Case marking on verb arguments might follow an ergative-absolutive pattern if the verb is in the perfective aspect, but a nominative-accusative one if the verb is imperfective, the Case of certain nouns couldn't be determined until the aspect of the verb is established. Thus, once again, it is clearly seen that Case is a derivative phenomenon with respect to the syntax and that nouns cannot enter a derivation with their Case pre-specified. If Case is to be seen as being selected by the numeration in a realist sense there must necessarily be a separate syntax at work, not presently accounted for in the MP, that derives the numeration in the first place, which would essentially leave much of the rest of the syntactic computation described by the MP more or less redundant.

2.2 *Flaws with the 'random generator' model call for a return to the semantics-driven syntax of generative semantics*[6]

Of course, another approach, one not compatible with a realist causal, processual model of language production from thoughts to speech, would assume that nominals can enter a derivation pre-specified with their Cases because surface structures are *not*, in fact, to be seen as the output

[6] As one anonymous reviewer has pointed out, not all aspects of generative semantics are necessarily desirable; in particular, its derivations typically resulted in syntactic trees that were overly large and cumbersome (see for instance the (in)famous representation to the sentence *Floyd broke the glass* by George Lakoff and Haj Ross, moderately elided in Harris (1993: 214)). As the same reviewer points out however, syntactic derivations in the MP are also often inordinately complex. The one very desirable characteristic of generative semantics that is the present focus is its premise that semantics underlies and precedes the syntactic computation. This is in contrast to the Chomskyan interpretive semantics view, according to which the semantics is interpreted from the syntax only at the end of the syntactic derivation.

of syntactic computations upon semantic input. Rather surface speech is to be regarded as the end point of derivations that begin with what Seuren (2004: 46-49; 149-161) terms a *random generator*: syntax and semantics must operate on an output that already contains the basis for a sentence, the origins of which are entirely random and arbitrary. This is in fact the approach entailed by contemporary Chomskyan syntactic theory, including that of the MP, and ultimately made possible by the competence dictate that forbids competence theorists to be concerned with how a sentence is actually produced in terms of a process. Thus, in the MP, the numeration must necessarily be seen as the product of this kind of a *random generator*. This is, furthermore, of course the inescapable conclusion of Chomskyan *interpretive semantics* (IS), as opposed to the *generative semantics* (GS) view of the 1960's and 1970's. The IS position holds that a syntactic product needs to be *interpreted by* the semantics only by the end of a derivation. GS on the other hand holds that semantic information is what drives the computation which yields a surface syntactic output (cf. Newmeyer 1986: 81-138; Harris 1993). It is clear that a view in which semantics drives the syntax is better suited to model the actual causal mechanisms at play in generating language, which can only be ignored as long as Chomsky's non-process-based competence dictate stands.

As one anonymous reviewer has pointed out, it should be noted that an increasing number of works written from a Chomskyan Minimalist perspective actually diverge from their master with regard to the conception of the numeration as consisting of a collection of phonologically specified lexical items. Such works in fact display treatments that resemble those of GS in that they begin with an input of non-phonological features and adopt the late insertion of phonological exponents. A number of such treatments (see, for example, De Belder & van Craenenbroeck 2014; De Belder & van Craenenbroeck 2015; Kramer 2016), however, adopt a Distributed Morphology (DM) framework (Halle & Marantz 1993; Harley & Noyer 1998; Harley & Noyer 1999) in which the derivation starts off with only syntactic features, leaving all semantic features of roots, for example, to be filled in after the syntactic derivation has ended (Harley & Noyer 1998: 7-14; Harley & Noyer 1999: 3-5). This means that in this view the syntax generates a complete sentence – including the positions to be filled in by roots – first, before the entities that are denoted semantically by the roots are even considered. In fact, in many DM treatments the decision to insert any semantic root lexical items is essentially left to free choice (Harley & Noyer 1998: 7, 10; Harley & Noyer 1999: 5; De Belder & van Craenenbroeck 2015: 650).[7] While such a system is compatible with a Chomskyan IS random generator model, displaying, as Seuren says, the "production of thought through language" (Seuren, 2004: 27), it is utterly incompatible with a realist model that reflects the process whereby an input of thoughts is translated by the syntax into a linguistic output.

[7] Note that some more recent work in DM differs slightly from this conception in positing that roots are in fact individuated prior to the syntactic computation, but not semantically. Thus in Harley 2014 (225-227; 242-247), for instance, roots are pre-identified with an abstract index. One author, Daniel Siddiqi, did briefly purport in 2009 (18-20) to embrace a DM model in which roots come semantically pre-individuated, which would in fact be compatible with a GS view. However, he has since abandoned this view and opted instead for a model that is compatible with the pre-individuation of roots by means of abstract indices (invited talk given at Morphology in Montreal-Ottawa-Toronto (MoMOT) 3, November 17th, 2018).

2.3 *Performance considerations incompatible with a non-process competence dictate invalidate the phase theory of the MP*

Holding tightly to Chomsky's dictate that the aim of syntax is not to discover how in a temporal, causal sense speech is produced also allows contemporary transformative syntacticians to ignore evidence that could otherwise be brought to bear on the subject from other disciplines such as psycholinguistics. Take, for instance, the notion of *phase* and concept of *derivation by phase* found in the MP. According to Chomsky (2000; 2001) (see also, Hornstein et al. 2005: 346-362), the product of a linguistic derivation is spelled out phase by phase, starting with the most embedded phase. A *phase* is taken to be a unit of syntactic structure that is inspected for convergence; if it converges the phase is spelled out and the derivation proceeds but subsequently as per Chomsky's *Phase Impenetrability Condition* (e.g., Chomsky 2000: 108; 2001: 13) the computation may not access whatever is inside the inner domain of the phase.

However, according to Labelle (2007: 6), psycholinguistic research has shown that syntactic processing takes place "left to right" and not "right to left" as in Chomsky's phase-based theory. She argues that short-term memory constraints simply would not allow language users to spell out a series of phases and store them in memory until the final one is reached and then articulate or process them in reverse order:

> *…whether from the point of view of the speaker or of the hearer, the computational system doesn't treat sentences starting from the most embedded phase in a language like English. Instead of saying: Who said that Mary gave a book to Paul? speakers don't spell-out something like the following (brackets added to make clear the approximative derivation): [v*P phase1 gave a book to Paul] [CP phase2 that Mary] [v*P phase3 said] [CP phase4 who]? It will not do to assume that speakers can keep in memory all the phases already planned, waiting for the most external phase to be completed, before spelling them out in the reverse order. The capacity of short term memory is simply too small for that.*

Once again, we see that the decades-old dictate enjoining syntacticians to abstain from describing causal relations underlying the production of sentences is at work. Unfortunately, here we see how it can lead adherents of the MP to ignore crucial performance-based considerations that would otherwise bear significantly on their model.

3. Conclusion

Perhaps we ought to jettison the Chomskyan dictate that it is not the aim of the syntactic enterprise to seek compatibility with an instantiable, process-based model. Such a move might necessarily lead to a realistic re-evaluation of the supposed merits of IS, as opposed to a GS view, as a next step. Seuren (2004: 169-190) revisits Chomsky's arguments dating as far back as the early 1970's that supposedly necessitate accepting that some aspects of meaning must be contributed in the syntax at the surface level rather than at the deep structure level and shows that

they are faulty.[8] Such arguments would have been the only thing that could possibly be used in defense of a random-generator syntax.[9]

As Pieter Seuren says, since it's obvious that a random language generator is not found in the brain, Chomskyan interpretivist syntax must necessarily be regarded as an example of scientific instrumentalism. However, at the end of the day, in order to come up with a realist model of how speech is actually produced that can causally explain the inner workings of language, a GS analysis, rather than an IS one, is required, with meaning driving speech. As Seuren puts it:

> *All [Chomsky's] theory of language and grammar does is provide for a purely algorithmic account of how a potentially infinite array of different sentences can be built up from a finite collection of primitive elements, but how sentences built up in this manner can be seen to express thoughts is left totally unexplained.*
> *In none of its successive varieties does the Chomskyan random-generator concept of grammar contain or imply anything like an account of the notion "expression of thought through language". On the contrary, it invites the utterly unintuitive notion "production of thought through language". (Seuren 2004: 4; 27)*

But if this is the case, of what benefit in the overall scheme of things is it to explore, in addition to a realist model, a non-causally based instrumentalist, IS analysis, like that Chomsky presents? It essentially only points us to an *alternative* way a sentence *might* be constructed if language was other than it was; thus, what positive contribution does it ultimately make towards the science of understanding how language works? To arrive at this goal while pursuing Chomsky's model one would have to necessarily construct two distinct versions of syntax. One of these would necessarily be *incompatible* with realism, therefore, not reflective of the actual mechanisms at play. This is ultimately the consequence if one holds to Chomsky's insistence that the goal of syntax is not to describe how we actually produce language in terms of a causal process.

In sum, perhaps there would be much merit in reconsidering the accepted dictate of Chomskyan linguistics that transformational syntacticians are not to attempt to discover the causal, processual aspects of how the syntactic system inside the mind functions. Otherwise, as Labelle (2007: 7) says, we are "pretending that what we are doing is describing the computational system for the human language…while what we are really doing is constructing a linguistic system independent of psychological and biological concerns."

---

[8] Seuren argues that there is no reason to accept that focusing strategies, for instance, have to be seen as deriving from the surface structure. He suggests that they are in fact better accounted for as originating from cleft structures present in the underlying semantic structure. He likewise shows that presuppositions and assignment of operator scope, two other phenomena Chomsky argues are surface-derived, are best seen as deriving from deeper semantic structure (2004: 169-190).

[9] Moreover, as Seuren (2004: 150-151; 169-170; 172-174) points out, in the MP there is no feeding line from the surface, Phonological Form (PF) of an utterance to its Logical Form (LF) (i.e., its semantic form) (unlike in the MP's predecessor, GB, where the surface structure directly fed LF). For this reason any aspects of meaning that are contributed at this surface level could not be fed into the LF regardless, so such arguments could not be used to defend the MP version of a random-generator model.

Upon reflection it becomes clear that considerations of performance ought to bear upon the construction of generative syntactic models of linguistic competence. This obvious fact is obscured by Chomsky's dictate eschewing the inclusion of causal factors in his stated definition of competence. As Seuren (1996: 5-6) has noted,

> *Relevant hard results of psycholinguistic experiments have to be respected by the linguist…An interesting and fruitful dialogue will then come about between linguists and psycholinguists against the double backdrop of psycholinguistic and linguistic theory formation.*

As one anonymous reviewer has pointed out, numerous psycholinguists have shown how performance-based criteria can be used to evaluate alternative linguistic theories (see, for example, Sedivy's account of the 'parsing wars' over *garden-path* versus *constraint-based* models of syntactic representation and processing (Sedivy 2018: 289-306). Moreover, serial models of speech production in psycholinguistics typically display compatibility with a GS-type conception of syntactic representation, according to which the semantics drives the computation and feeds the syntax, as opposed to an IS model of syntax and semantics where a syntactic input ends up 'interpreted' semantically at a later stage of the derivation (e.g., Bock and Levelt's model (Bock and Levelt, 1994: 946)). Such information shouldn't be disregarded by competence researchers. Furthermore, Jackendoff (2002: 57-59) identifies Chomsky's stance of rigidly separating his notion of competence from that of performance as seeming to have hardened over the years and points out that this is unfortunate. Surely, we should take our cue from such sentiments and revise our models accordingly; such action, however, might necessitate finally letting go of Chomsky's firmly anti-causal definition of competence.

At the end of the day, transformational generative syntacticians need to cease staunchly upholding the Chomskyan dictate that forbids studying the processual, causal mechanisms that necessarily underlie how a sentence is actually produced. Otherwise, we risk expending our efforts towards an enterprise that is ultimately redundant from the point of view of scientific realism and considerations of performance.

Abbreviations

DM – Distributed Morphology
ECM – Exceptional Case Marking
GB – Government and Binding Theory
GS – Generative Semantics
IS – Interpretive Semantics
LF – Logical Form
MP – Minimalist Program
SS – Surface Structure

References

Bresnan, Joan & Kaplan, Ronald M. 1982. Grammars as mental representations of language. In Bresnan, Joan (ed.), *The mental representation of grammatical relations*, xvii-lii. Cambridge, MA: MIT Press.

Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: the framework. In Martin, Roger & Michaels, David & Uriagereka, Juan & Keyser, Samuel Jay (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89-155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In Kenstowicz, Michael J. (ed.), *Ken Hale: A life in language*, 1-52. Cambridge, MA: MIT Press.

De Belder, Marijke & van Craenenbroeck, Jeroen. 2014. On vocabulary insertion. *Linguistic Analysis* 39. 173-210.

De Belder, Marijke & van Craenenbroeck, Jeroen. 2015. How to merge a root. *Linguistic Inquiry* 46. 625-655.

Halle, Morris & Marantz, Alec. 1993. Distributed morphology and the pieces of inflection. In Hale, Kenneth & Keyser, Samuel Jay (eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, 111-176. Cambridge, MA: MIT Press.

Harley, Heidi. 2014. On the identity of roots. *Theoretical linguistics* 40 (3/4). 225-276.

Harley, Heidi & Noyer, Rolf. 1998. Licensing in the non-lexicalist lexicon: nominalizations, vocabulary items and the encyclopaedia. (https://www.researchgate.net/publication/246685010_Licensing_in_the_NonLexicalist_Lexicon_Nominalizations_Vocabulary_Items_and_the_Encyclopaedia) (Accessed 2018-07-14)

Harley, Heidi & Noyer, Rolf. 1999. State-of-the-article: Distributed morphology. *Glot International* 4. 3-9.

Harris, Randy Allen. 1993. *The linguistics wars*. Oxford: Oxford University Press.

Hawkins, John A. 2009. Language universals and the performance – grammar correspondence hypothesis. (http://lingo.stanford.edu/sag/L204/Hawkins-LgUnivFinal.pdf). (Accessed 2018-06-18.)

Hornstein, Norbert, & Nunes, Jairo & Grohmann, Kleanthes K. 2005. *Understanding minimalism.* Cambridge: Cambridge University Press.

Jackendoff, Ray. 2002. *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.

Kramer, Ruth. 2016. A split analysis of plurality: Number in Amharic. *Linguistic Inquiry* 47. 527-559.

Labelle, Marie. 2007. Biolinguistics, the Minimalist Program, and psycholinguistic reality. *Snippets 14*. 6-7. (http://www.ledonline.it/snippets/allegati/snippets14002.pdf). (Accessed 2018-01-15.)

Lenerz, Jürgen. 1998. Review of Chomsky (1995a). *Beiträge zur Geschichte der deutschen Sprache und Literatur* 120.1. 103-111.

Newmeyer, Frederick J. 1986. *Linguistic theory in America: Second edition*. San Diego: Academic Press, Inc.

William O'Grady. 2010. An emergentist approach to syntax. In Narrog, H. & Heine, B (eds.), *The Oxford handbook of linguistic analysis, 257-283.* Oxford: Oxford University Press. (http://ling.hawaii.edu/wp-content/uploads/OGrady_EmergentistApproachtoSyntax.pdf). (Accessed 2018-06-19.)

O'Grady, William. 2005. *Syntactic carpentry: An emergentist approach to syntax.* Mahwah, N.J.: Erlbaum.

Sag, Ivan A. & Hofmeister, Philip & Snider, Neal. 2007. Processing complexity in subjacency violations: The complex noun phrase constraint. *Proceedings of the Annual Meeting of the Chicago Linguistic Society*, *43.1*. 215-229. (Available at: http://lingo.stanford.edu/sag/papers/cnpc.pdf). (Accessed 2018-06-18).

Sedivy, Julie. 2018. *Language in mind: An introduction to psycholinguistics*. Sunderland, Massachusetts: Sinauer Associates, Inc. (imprint of Oxford University Press).

Siddiqi, Daniel. 2009. *Syntax within the word: Economy, allomorphy, and argument selection in Distributed Morphology*. Philadelphia: John Benjamins Publishing Company.

Seuren, Pieter A. M. 1996. *Semantic syntax*. Oxford: Blackwell.

Seuren, Pieter A. M. 2004. *Chomsky's minimalism*. Oxford: Oxford University Press.

Baiba Vilks
St. George St.
Toronto
Canada
bvilks4@gmail.com

In SKASE Journal of Theoretical Linguistics [online]. 2018, vol. 15, no. 3 [cit. 2018-13-12]. Available on web page http://www.skase.sk/Volumes/JTL38/pdf_doc/01.pdf. ISSN 1336-782X.

Emergence of iambs in Eastern Arabic: Metrical iambicity dominating optimal nonfinality

Bassil Mashaqba & Anas Huneety
Hashemite University, Jordan

*This paper presents a relatively novel metrical and constraint-based analysis of stress patterns in Bedouin Arabic as spoken by Zalabiah and Zawaideh subtribes in southern Jordan. Different from Eastern Arabic variety, the dialect under investigation exhibits stress patterns which are basically controlled by an iambic foot. In terms of metrical parameters, final foot extrametricality is proposed where it does not violate non-exhaustivity and peripherality; extrasyllabicity is also proposed for the analysis of a final consonant in a superheavy syllable. Extrametricality is better analysed, in constraint-based terms, via the constraint NONFINALITY. Examining the domain hierarchy in Optimality Theory, the argumentation then explores how the metrical parameter 'iambicity' struggles with the optimal constraint 'nonfinality' as well as with other constraints to generate the more harmonic stress patterns. In terms of dominance hierarchy, the constraints IAMB and WEIGHT-BY-POSITION are ranked higher than the constraints NONFINALITY and *FINAL-C-μ, supporting that stress patterns in the dialect of Zalabiah and Zawaideh are mainly controlled by the iambic foot. Contrary to previous literature, this work provides evidence that not all Eastern Arabic dialects necessarily fit into trochaic foot-type.*

Keywords*: Arabic Word Stress, Metrical Model, Optimality Theory, Iamb*

1. Introduction

All modern Arabic dialects, including Jordanian Arabic, have a regular word stress which does not serve to distinguish meaning. Phonetically, stressed syllables in Arabic varieties are louder, longer and of noticeably higher/different pitch levels than unstressed syllables (Al-Ani 1992; Mashaqba 2015).

From the early stages of modern phonological theory, serious theoretical work has been devoted to Arabic word stress, within different theoretical approaches: the pregenerative approach, with reference to the nature of the syllable (Mitchell 1956, 1960; Harrell 1957; Wright 1971); the generative approach, where stress is distinctively represented by the feature [+/-stress] (Abdo 1969; Brame 1970, 1973, 1974; Broselow 1976; Johnson 1979; Weldon 1980); the metrical approach, where the syllable is weight based, rather than segment-based (McCarthy 1979a, 1980; Angoujard 1990; Hayes 1981, 1995; Watson 2002; Huneety & Mashaqba 2016); and the constraint-based approach of Optimality Theory (Kager 1999; McCarthy 2008; Rakheia 2009; Al-Jarrah 2002, 2011).

This great deal of attention to stress in modern Arabic dialects relates to a variety of factors including: (i) the significant variation among the dialects of stress patterns though they share some patterns in common;[1] (ii) the relevance and applicability of stress patterns to

[1] This variation is based on (i) the difference in the distribution of syllable types (Watson 2011), (ii) being 'differential' where only unstressed high vowels are subject to syncope interaction of stress or 'non-differential' where all unstressed vowels are subject to syncope (Cantineau 1939), (iii) long vowels shortening in open unstressed syllables (Younes 1995), (iv) interaction of stress, syncope and epenthesis, (v) and the lexical information (cf. Kiparsky 2003; Kager 2009; Watson 2011).

metrical theory parameters (cf. Hayes 1995), and (iii) the great interaction of stress, syllabification patterns, and phonological processes (especially epenthesis and syncope). (cf. Broselow 1982; Kenstowicz 1983; Kiparsky 2002); (for more details, see Kager 2009.)

Stress patterns in Eastern Arabic dialects are apparently indistinguishable from one another. Stress placement is subject to 'quantity-sensitivity', i.e., it is sensitive to syllable weight (light, heavy, or superheavy); it also has the *demarcative* property where the position of the main stress is limited to the last three syllables of the word and never exceeds the antepenult window. Eastern Arabic dialects (exclusively Levantine) share the following rules: (i) a superheavy ultimate syllable always triggers the main stress, as in *ʕæ:.la.'me:n* 'two worlds' (Younes 1995: 163); (ii) otherwise, assign the main stress to a heavy penult, as in *mos.'taʃ.fa* 'hospital' (Johnson 1979:154); (iii) otherwise, stress the antepenult, as in *ḥæ:.'ra.ba.to* 'she fought him' (Younes 1995:163). (iv). In disyllabic words ending in -CV or -CVC, stress falls on the first syllable, as in *'ḍarab* 'he hit', *'bana* 'he built' (Watson 2011: 3011).

Metrically, this analysis produces three main aspects to generate stress assignment: (i) parse the word from left to right into moraic trochees to assign a foot; (ii) a foot is rendered extrametrical at the right edge of the word; and (iii) feet are grouped into a right-headed word constituent. The antepenult (rule iii mentioned in the paragraph above) is active by rendering the final syllable extrametrical, and constructing a quantity-sensitive trochee at the right edge (cf. Kenstowicz 1983). On the other hand, Eastern Arabic dialects show a few internal differences (ambiguity) between a right-to-left analysis with syllable extrametricality, or a left-to-right analysis with moraic trochees and foot extrametricality. For example in LLLL structures, Palestinian variety assigns stress to the fourth syllable from the right-edge (i.e., 'LLLL),[2] but Beirut/Damascene assign stress to the third from the right edge (i.e., L'LLL). This variation is attributed to a difference of parsing direction (left-to-right in Palestinian but right-to-left in Beirut/Damascene), and a difference in extrametricality (with foot extrametricality in Palestinian but with syllable extrametricality in Beirut/Damascene) (cf. Halle & Kenstowicz 1991).

As for Jordanian Arabic (henceforth JA) dialects, stress has also been discussed in detail in a number of studies over the last few decades (e.g., Yasin 1980; Irshied 1984; Sakarna 1999, 2002; Palva 1976, 1980, 1986, 1991; Rakhieh 2009; Hérin 2010). Among the others, Abu-Abbas (2003), Rakhieh (2009), Al-Jarrah (2011), Mashaqba (2015), and Huneety (2015) have theoretically accounted for stress patterns of selected JA dialects primarily based on Optimality Theory and/or metrical approach. The core analysis of the earlier studies (except for Mashaqba 2015) leads the reader to conclude that (similar to the Levantine) all JA varieties exhibit (moraic) trochaic foot, and that JA dialects share the same metrification directionality and the level of extrametricality with other Levantine varieties.[3] However, we will see that the spoken dialect by Zalabiah and Zawaideh subtribes in Wæ:di Ramm southern Jordan, the target of the present work, has a different foot type, namely an iambic foot – a step that may promote the idea to have a serious revision of the literature on Eastern Arabic

[2] Palestinian dialect does not have words comprising a sequence of four light syllables (LLLL) due to processes of syncope and epenthesis. But the attested examples in the literature are actually selected Classical Arabic words that show how Palestinian speakers stress Classical Arabic words containing a sequence of four light syllables as in *'ʃaʤaratun* 'a tree' and *ʃaʤa'ratuhu* 'his tree' (Kenstowicz 1981).

[3] Based on examples from Hung (1995), Al-Jarrah (2002) makes specific reference to Bedouin Jordanian dialects being iambic but he does not expand this important criterion significantly.

dialects. For simplification, this dialect is referred to as Wæ:di Ramm Arabic (henceforth WRA).

2. Material and methodology

To examine stress patterns in WRA, the researchers recruited 30 participants, 15 males and 15 females. According to a metadata sheet, data were collected from participants of various ages ranging between 40–72 years covering the Zalabiah and the Zawaideh subtribes. The average age is 55 years old. None of the participants live outside the region (Ramm and Ad-Di:sih) nor did they travel outside the country. The spoken variety they produce is not affected by modern life aspects which change some of the linguistic features of the speech of young generations. Participants' level of education ranges from primary education (14 participants) to secondary education (16 participants). The participants have no speech impediments and were willing to answer all questions that serve this study. A consent form was filled in by each participant; they were additionally informed verbally about the purpose of the research and the way the material would be saved in order to guarantee the protection of the participants' confidentiality. At the end of the investigation, data relating to two participants were destroyed based on their wishes.

Data collection started in 2012 and the list of words collected was repeatedly updated until February 2013. A second fieldwork took place over two weeks in mid-December 2015, where many new participants and new words were retrieved and added to the corpus of this study. The second round revision leads to correcting few patterns that have been set in the first trip (as will be seen in data containing final closed syllables). The list included monosyllabic and polysyllabic words that comprise all possible syllable types in the dialect under investigation. The words pronounced were also double-checked with the participants themselves and with a language consultant, a researcher who is a native speaker of the variety under investigation. It has also been decided NOT to include rule-based analysis entirely since far fewer people are working in a rule-based framework these days.

3. Metrical stress-based model (moraic model)

This entry examines the stress patterns of WRA starting with a metrical model, followed by an Optimality-Theoretic model in Section 4. The Appendix includes a table with examples specific to WRA, organized according to syllable structure of all types and combinations attested (Monosyllabic, disyllabic, trisyllabic all combinations, etc.).

WRA is a quantity sensitive dialect which distinguishes three types of syllables: light (CV and CCV), heavy (CVV, CVC, and CCVV), and superheavy (CVVC and CVCC). To account for the aspects of stress in WRA, this section draws on metrical theory advanced in Liberman & Prince (1977), Selkirk (1980), McCarthy (1979a), and Hayes (1995) as follows: A CV syllable counts as a light syllable and is assigned one mora; CVV counts as a heavy syllable and is assigned two moras; CVC counts as a heavy syllable and is assigned two moras, with the final C assigned a mora through the WEIGHT-BY-POSITION constraint (Hayes 1995; Watson 2002). The word-final CVC in WRA is a heavy syllable, so the word-final C counts to the syllable weight and is assigned a mora. Geminate consonants are assigned one mora underlyingly (cf. Mashaqba 2015). More specifically, applying the

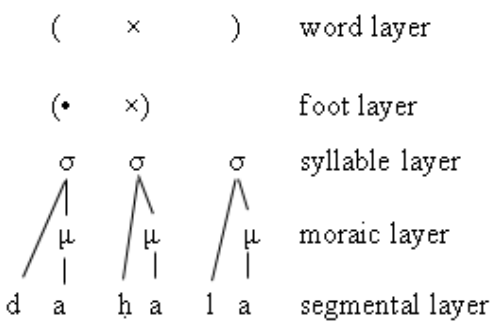
bimoraic parameter, light syllables are monomoraic; heavy and superheavy syllables are bimoraic. Superheavy syllables are rendered bimoraic (NOT trimoraic) by presuming the final C to be extrasyllabic in WRA as will be shown later on in this section.

Three major parameters can be captured in terms of the present theory to establish stress patterns of WRA and to differentiate between its stress patterns and the other Arabic dialects, namely: (i) foot type (trochee or iamb), (ii) metrification directionality (left-to-right or right-to-left), and (iii) type of extrametricality (consonant, mora, syllable, or foot). (See Rosenthall & Van Der Hulst 1999; Kager 2009; Hyde 2007, 2011.)

Foot in WRA is iambic, binary, bounded (of a fixed size) and right-headed, and feet are parsed from left-to-right (Mashaqba 2015). The basic foot inventory in WRA comprises two consecutive light syllables (L'L), a light syllable followed by a heavy syllable (L'H), or a single heavy syllable ('H). Feet follow right-headed constituents through application of End Rule Right (ERR) (cf. Hayes 1995).

The structure (HL) or an odd number of syllables (e.g., LLL) raises a problem to syllable parsing because it may produce a leftover light syllable unfooted. WRA belongs to those dialects which impose a ban on degenerate feet.[4] In a word like *da'ḥala* ‘dune’, the light syllables (da.ḥa) are successfully parsed into a foot, but the peripheral light syllable (la) fails to be parsed into a foot as it does not satisfy the bimoraicity condition, and thus is left stranded, as shown in the metrical grid in (1):[5] The foot parses left-to-right to produce a right-headed foot over the syllables (da) and (ḥa) in (1). The rightmost syllable (la) is monomoraic and left unfooted because it cannot constitute/support a foot. Dictated by the End Rule Right (ERR) principle, stress is assigned to the head of the rightmost visible foot, *ḥa* in (1).

(1)



However, WRA imposes weak prohibition on degenerate feet (see type b in footnote 4 below) as the final C in word-final CVC syllables counts to be moraic in monosyllabic words. This factor can be supported via a number of content words that must count as monomoraic if the final consonant is rendered extrametrical, as in *'xuð* ‘take!’ *'kul* ‘eat!’, *'gil* ‘say!’, *'tam*

[4] These sub-minimal elements may be stressed as in some languages degenerate feet (Hayes 1995: 87). In others, unparsed syllables remain unfooted and unstressed. Three different levels of prohibition on degenerate feet are proposed (Hayes 1995: 87 in Mashaqba 2015: 114):

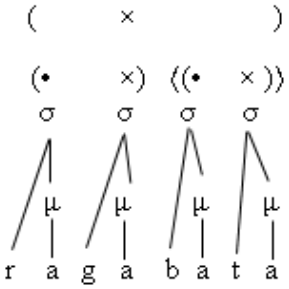
a) strong prohibition	absolutely disallowed
b) weak prohibition	allowed only in strong position, i.e., when dominated by another grid mark
c) non-prohibition	degenerate feet are freely allowed.

[5] The stressed head of a foot is marked by (×), and the weak/non-stressed element of the foot is marked by (•).

‘done’ *'ʔax* ‘brother’, *'ʔam* ‘mother’ *maj* ‘water’.[6] Such words surface as they are without resorting to a final consonant gemination or vowel lengthening, i.e., the final consonant is moraic. This may promote the idea of deactivation of extrametricality at the consonant level but its applicability at higher prosodic levels in this dialect.

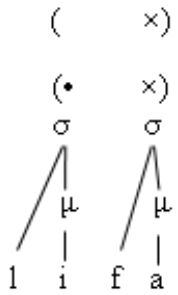
Extrametricality applies to a final unfooted light syllable, and to a final foot if it does not violate the non-exhaustivity condition. The inability of penultimate light syllables in four light-syllable words, such as *ra'gabata* ‘his neck’ in (2), presumes that the rightmost ‘peripheral’ foot counts extrametrical (not a degenerate foot, given that WRA imposes a weak prohibition on a degenerate foot) as it ‘does not exhaust the domain of the word’ (Hayes 1995: 58).[7, 8]

(2)



If the word comprises only one foot, foot extrametricality is blocked and the head of the only (peripheral) foot is stressed, as in *li.'fa* ‘to come’ in (3):

(3)



Motivated by the ban on trimoraic syllables, the behaviour of superheavy syllables CVVC and CVCC word-medially and finally poses challenges to the maximal bimoraic constraint of syllables as proposed in OT and pre-OT models that dealt with stress and syllabification patterns in Arabic varieties (e.g. Broselow 1992; Watson 2002; Kiparsky 2003; Al-Jarrah 2011). On the one hand, some argue that the final consonant falls outside the domain of the syllable and renders it *extrasyllabic* which is directly linked to a degenerate syllable (Hayes 1995: 106–107; Kager 1995: 376; Kenstowicz 1994: 274).[9] Hayes (1982: 229) proposes that

[6] For those familiar with Levantine Arabic, the claiming that CVC syllables do not undergo ‘final consonant gemination’ or ‘vowel lengthening’ is not convincing without some phonetic evidence. This is a long debated issue that cannot simply be brushed away with an unsupported statement.

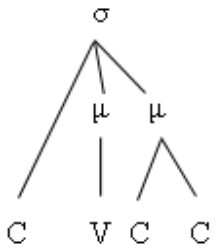
[7] Foot extrametricality has been demonstrated in many trochaic and iambic Arabic varieties (cf. Hayes 1995: 227 ff. for Negev Bedouin; Watson 2002 for San'ani).

[8] Angled brackets are used to denote extrasyllabic consonants or extrametrical foot.

[9] To avoid violating NONFINAL and FT-BIN, Crowhurst (1996: 416, cited in Al-Jarrah 2011), following McCarthy (1979b), ‘argues for treating final Cs in trimoraic’ sequences ‘as degenerate feet – as a repair

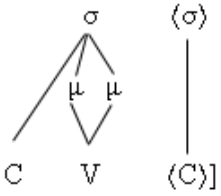
word-final superheavy syllables count as heavy, and heavy syllables count as light. On the other hand, Farwaneh (1995: 66–70) and McCarthy (2007a: 147–148) suggest that the final two consonants may share a mora, as in (4):

(4)

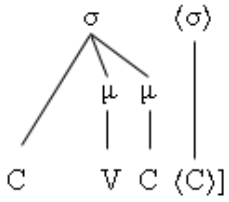


Drawing on (Hayes 1995: 106–107; Kager 1995: 376; Watson 2002: 92–94), the authors of the present paper assume that the final consonant in an ultimate superheavy syllable is extrasyllabic and is not integrated into the adjacent syllable in WRA, i.e., the ultimate superheavy syllable is mapped as a bimoraic syllable plus a degenerate syllabic consonant, as in (5a) and (5b).

(5) (a)

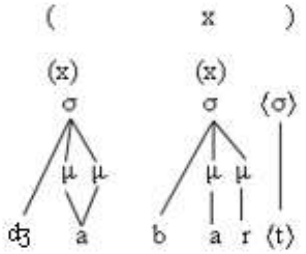


(b)



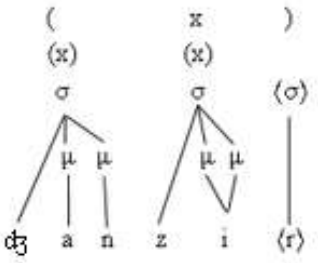
Peripherality condition also blocks foot extrametricality. That is, in words ending with a final superheavy syllable CVCC or CVVC, the final consonant is extrasyllabic and therefore intervenes between the syllable and the right edge of the word and therefore the syllable is no more in a peripheral position (Hayes 1995: 106–107). The extrasyllabic consonant will be immediately mapped to a separate syllable node. Take as an example, *ʤæ:.'bart* ‘I condoled’ and *ʤan'zi:r* ‘chain’ in (6) and (7):

(6)



mechanism’. Kiparsky (2003: 157) argues that ‘the final C is weightless, and thus NONFINAL(C) is undominated’.

(7)



Each of the leftmost bimoraic syllables *ʤæ:* in (6) and *ʤan.* in (7) constitutes a foot. Each of the rightmost syllables *.bart* and *.zi:r* is eligible to construct a second foot plus an extrasyllabic consonant. The rightmost foot escapes peripherality by the presence of the extrasyllabic ⟨t⟩ and ⟨r⟩ consonants which intervene between the rightmost foot and the end of the prosodic word. Stress is therefore assigned to the head of the rightmost visible foot by ERR.

Accordingly, the behaviour of stress system in WRA can be reformulated in (8) below: (cf. Mashaqba 2015: 116 after Hayes 1995; Watson 2002).

(8)	(i)	Syllable Weight	CVC, CVV = /–/, CV = /◡/. Word final C = syllable when after C or VV.[10]
	(ii)	Foot Construction	Form iambs from left to right. Degenerate feet are not permitted.
	(iii)	Foot Extrametricality	Foot → ⟨Foot⟩/___]word
	(iv)	Word Layer Construction	End Rule Right

4. Constraint-based model (OT)

Previous rule-based phonological theories have elaborated a variety of complex blocking and triggering constraints. A number of these works produce rule-less phonology because they only trace universal grammar. The rules of the widespread universal grammar, however, did not account for every language all of the time (cf. McCarthy 2007b). On the other hand, as a new approach to language, OT puts forward significant reliance on language-specific rules. The general premise of OT is that constraints are universal and the ranking is language-specific. The present section is intended to sketch constraint interaction on the output of the grammar of stress patterns in WRA. Data are examined to elaborate well-defined output constraints that block (ban) a syllable from receiving the main stress, and output constraints that trigger special syllables that satisfy main stress constraint, a necessary step to demonstrate the constraint hierarchy that accounts for all stress patterns in the dialect under investigation.

4.1 *Monosyllabic words*

For monosyllabic words, recall that a lexical word must be minimally bimoraic to be able to receive stress. The requirement of prosodic word-minimality can be achieved by the constraint *PrWdμ, as represented in (9) below.[11]

[10] This means that word-final C is syllabified in CVCC, CVVC.

(9) *PrWdμ: A prosodic word is minimally bimoraic (cf. Prince & Smolensky 1993/2004).

The constraint indicates that words must be footed and constraints on foot militate that words have at least two moras; this does not hold for CVC words in the other JA dialects where CVC counts as a light syllable word-finally (cf. § 3, data 2). Because a prosodic word must minimally have a single head foot and a foot must be bimoraic, FOOT BINARITY (which requires feet to be bimoraic) is a closely related constraint (cf. Selkirk 1984, 1995).

(10) FOOT-BINARITY (FT-BIN): Feet are binary at a moraic or syllabic level (Hayes 1995; Prince & Smolensky 1993/2004).

The metrical parameter that requires coda consonants to be moraic (Hyman 1985; Hayes 1989, 1995) is translated into OT's WEIGHT-BY-POSITION (WBP). Given that word-final consonants in absolute final domain are moraless, a monosyllabic word should end in a long vowel or in a cluster to be able to receive stress. This is translated into the constraint *FINAL-C-u which requires word-final consonants to be weightless. These two constraints are given in (11) and (12) below:

(11) WEIGHT-BY-POSITION (WBP): Coda consonants are moraic (cf. Hayes 1989).

(12) *FINAL-C-μ: A word-final consonant is weightless (Prince & Smolensky 1993/2004: 49; Kager 1999).

Recall that word-final consonants in CVC syllables are moraic in WRA as WRA imposes a weak prohibition on degenerate feet (cf. data in 2 above and footnote 4). Establishing the dominance hierarchy of these constraints, we can account for monosyllabic words of different forms. For example, the underlying form *xuð* 'take!' is rendered bimoraic by satisfying the higher ranked constraints *PrWdμ, FT-BIN, and WBP (see Table 1). Violating *FINAL-C-μ does not prevent the optimal candidate from being the most harmonic output, hence demotion of *FINAL-C-μ below the three constraints *PrWdμ, FT-BIN and WBP.[12]

Table 1: *PrWdμ, FTBIN, WBP >>*Final-C-u

/xuð/	*PrWdμ	FT-BIN	WBP	*FINAL-C-μ
(a) ☞ (xuð)				*
(b)(xu)<ð>	*	*	*	
(c)(xu)<ð>			*	
(d) xuð	*			

Table (1) shows that candidate (1b) satisfies *FINAL-C-μ but incurs a violation of the three higher ranked constraints *PrWdμ, FBIN, and WBP and thus ruled out. Candidates (1c) loses by violating WBP. Circumventing FTBIN by the null parse in (d) is not enough to win and

[11]The literature has well-established constraints that can replace the ad hoc *PrWdμ such as HEADNESS (a prosodic word contains a foot) (cf. Kager 1994), but we get stick to this constraint in harmony with the recent literature on Arabic dialects.

[12] Because *Final-C-μ and *PrWdμ are undominated in WRA, they will not be presented in the other tables.

thus it is penalised by *PrWdμ. This analysis shows crucial characteristics that distinguish WRA from other Eastern Arabic dialects: WBP outranks *FINAL-C-μ in the former; whereas *FINAL-C-μ outranks WBP in the latter.

To account for monosyllabic words of the form CVVC and CVCC (i.e., superheavy syllables), the final C is rendered extrasyllabic to satisfy FT-BIN. So FT-BIN and *PrWdμ will rank over WBP which ranks over *Final-C-u. Consider Tables 2 and 3 which establish the constraint ranking of these four constraints:

Table 2: *PrWdμ, FTBIN>> WBP >> *Final-C-u

/ṣæ:d/	*PrWdμ	FT-BIN	WBP	*FINAL-C-μ
(a) ☞ (ṣæ:)<d>			*	
(b) (ṣæ:d)		*		*
(c) ṣæ:d	*			

Table 3: *PrWdμ, FTBIN>> WBP >>*Final-C-u

/gamḥ/	*PrWdμ	FT-BIN	WBP	*FINAL-C-μ
(a) ☞ (gam)<ḥ>			*	
(b) (gamḥ)		*		*
(c) gamḥ	*			

Recall that WBP necessitates codas and semisyllables be moraic. Violating WBP does not deprive the winning candidates (2a) and (3a) from being the optimal output as long as they satisfy the higher ranked constraint FT-BIN. By way of contrast, satisfying WBP by (2b) and (3b) is not enough to win the competition. The null parse in (2c) and (3c) helps to satisfy FT-BIN but is penalised by *PrWdμ and thus ruled out.[13]

4.2 *Polysyllabic words*

The argumentation to be considered here is that extrametricality in the stress system of WRA can be better analysed, in constraint-based terms, via the constraint NONFINALITY. This constraint imposes a ban on a final mora from being parsed into the next higher prosodic level (syllable), a ban on parsing the final syllable into the foot structure of the word, followed by a ban on parsing the final foot into the prosodic word, and therefore banning final syllables from receiving the main stress. To deal with foot and syllable extrametricality principle, the constraint NONFINALITY is proposed against word-final stress unless other higher ranked constraints militate against that. For simplification, this constraint is presented as follows:

(13) NONFINALITY (NONFIN): No head of PW is final in PW (Prince & Smolensky 1993/2004: 56)

[13] For minimality, rendering FT-BIN higher than WBP in the dominance hierarchy is better than proposing a new constraint for word-final coda in CVCC/CVVC (say for example *SH-FINAL- μ) because this is tacitly indicated by the constraint *FINAL-C-μ.

4.2.1 *Words including light syllables*

WRA allows a sequence of up to four light syllables. Recall that feet in WRA have final prominence, so the constraint IAMB (in 14) ensures that the final member of bimoraic feet is prominent. In disyllabic words, stressing the second member (the right-most syllable) proves that foot is iambic-oriented. To control 'directionality' which requires feet to have left-to-right alignment in the prosodic word, we need the constraint ALIGN-LEFT (15):[14]

(14) IAMB: Feet are iambic (cf. Hayes 1995).[15]

(15) ALIGN-LEFT (ALIGN-L): Every prosodic word begins with a foot (Kager 1999: 169)

The constraint in (15) does not guarantee LR footing. All it does is make sure a single foot, of indeterminate size, occurs on the left edge of a word. As dictated by FT-BIN, light syllables cannot construct feet on their own, but each two light syllables are grouped together to make a binary foot. FT-BIN is proposed to rule out candidates which include degenerate feet (cf. data in 1 above).

Where we have two light-syllable words, as in *saʕa* 'to intend', FT-BIN necessitates parsing both syllables violating NONFIN, in order not to avoid (L'L) structure. According to the IAMB constraint, stress falls on the second light syllable. So the constraint hierarchy will be: FT-BIN, IAMB, ALIGN-L >> NONFIN:

Table 4: FT-BIN, IAMB, ALIGN-L >> NONFIN

/saʕa/	FT-BIN	IAMB	ALIGN-L	NONFIN
(a)☞ (sa_{μ}'$ʕa_{\mu}$)				*
(b) ('$sa_{\mu}ʕa_{\mu}$)		*!		*
(c)sa_{μ} ('$ʕa_{\mu}$)	*!	*	*	*
(d) ('sa_{μ}) $ʕa_{\mu}$	*!	*		

Table (4) demonstrates how the alignment constraints interact with each other in order to obtain the optimal output. Stress falls on the second light syllable according to IAMB, and the first syllable is parsed to obey FT-BIN and ALIGN-L. The output (sa_{μ}'$ʕa_{\mu}$) wins because it satisfies the top three ranked constraints. Candidate (4b) is eliminated as it has a trochaic foot with stress on the first syllable, candidate (4c) loses by it violating FTB-IN and ALIGN-L, while candidate (4d) loses because of violating FT-BIN.

Light trisyllabic words, as in *ʔa'ʕama* 'blind' are accounted for in the same way. The first two light syllables are parsed together to construct a bimoraic iamb, and the rightmost syllable is left unfooted (stranded). This satisfies the constraint hierarchy: FT-BIN, IAMB, ALIGN-L >> NONFIN. Table (5) demonstrates the interaction of these constraints:

[14] ALIGN-L translates Ito's (1989) 'directionality' that requires syllables to be aligned with the left edge of the prosodic word in coda dialects to ensure that the epenthetic vowel lands to the left of the stranded consonant (e.g., the stranded .t. in /ka.tab.t.ha/ 'I wrote it' > *ka.ta.bit.ha* rather than *ka.tab. ti.ha*).

[15] IAMB replaces McCarthy & Prince's (1993) ALIGN (Hd-σ, Ft, R): Align the head syllable with the right edge of the foot.

Table 5: FT-BIN, IAMB, ALIGN-L >> NONFIN

/ʔaʕama/	FT-BIN	IAMB	ALIGN-L	NONFIN
(a)☞ (ʔa$_{\mu}$'ʕa$_{\mu}$) ma$_{\mu}$				
(b)('ʔa$_{\mu}$ʕa$_{\mu}$) ma$_{\mu}$		*		
(c)(ʔa$_{\mu}$ʕa$_{\mu}$) ('ma$_{\mu}$)	*!	*		*
(d)ʔa$_{\mu}$ (ʕa$_{\mu}$'ma$_{\mu}$)			*	*

In light syllable words, LL and LLL, FT-BIN dominates the NONFIN to optimize the desired output (L'L) and (L'L)L, respectively. Candidates (5c) and (5d) lose as mono-moraic (σμ) and light bi-moraic (σμμ) syllables are invisible to stress word-finally by violating NONFIN. Incorporating ALIGN-L into the constraint hierarchy eliminates candidate (5d) in favour of the three other candidates. Though candidate (5d) satisfies iambicity, the first syllable cannot construct a foot on its own as it would violate FT-BIN. This specification supports the argumentation that this syllable (leftmost syllable) needs to be footed. We need the constraint ALIGN-LEFT, which requires the left edge to be aligned with a foot in WRA. Candidate (5d) fails to align its left edge with a foot violating ALIGN-L. Four-light-syllable words will be argued for later on (cf. Table 14).

At this point, strictly speaking, a reader may impressionistically understand that what is needed to prove that the system is one word: (e.g., *sa'ʕa* 'to intend'). To this end, one might ask this legitimate question: in what sense do iambs emerge? i.e., WRA would be an iambic by fiat since it has a high-ranking IAMB constraint. Assuming the stress patterns are correct, we will see that this dialect (which belongs to Eastern Arabic) differs from all other eastern dialects (which are trochaic) insofar as it lacks final consonant extrametricality. Going a little further, data like *ʔa.'ʕa.ma* 'blind' show that another analysis (i.e. that stress is always final) is not necessarily true. Subsequently, as weight is forced in, adequate data is accounted for in the next sections to provide evidence for all the claims about weight relationships.

4.2.2 *Words including heavy syllables*

To account for words of heavy syllables, a new constraint should be proposed. Being a weight sensitive language entails WEIGHT-TO-STRESS Principle (WSP) which requires heavy syllables to be stressed (Prince 1990) (see data in 16). Two heavy syllables may occur within the stress window which entails that one of them receives the main stress. The pre-OT ERR, which requires the head of the prosodic word be aligned with the right edge of the word (Hayes 1995), is also replaced by the constraint RIGHTMOST (see 17), which forces the rightmost visible foot to bear main stress wherever more than one heavy syllable is attested:

(16) WEIGHT-TO-STRESS Principle (WSP): Heavy syllables are prominent in foot structure and on the grid (Prince & Smolensky 1993/2004).

(17) RIGTHMOST: the rightmost foot of the word is the head (Prince & Smolensky 1993/2004; Al-Jarrah 2002).

4.2.2.1 *Disyllable word stress*

The fact that a heavy syllable in final position receives stress violating NONFIN entails that WSP also dominates NONFIN. Recall that the final heavy syllable lacks final consonant extrametricality. With heavy disyllabic words, many possibilities are suggested, of which one optimal output is accepted: ('H)(H), (H)('H), ('H)H, H('H). The last two possibilities violate WSP; ('H)(H) satisfies WSP but violates RIGTHMOST; (H)('H), the optimal candidate, satisfies WSP and RIGTHMOST but violates NONFIN. RIGTHMOST and WSP are supposed to outrank NONFIN in WRA because violating NONFIN is superior to leaving the heavy syllable unfooted. To rule out candidates like ('H)(H), we should propose that RIGHTMOST outranks WSP as well. One aspect still needs to be solved: the optimal footing of words is to assign every heavy syllable a foot. In line with Rakhieh (2009) and Abu Guba (2016), heavy syllables always receive stress (being primary or secondary). This is evident from acoustic analysis and gemination (for details see Abu Guba 2016: 248). Seemingly, the existence of two or more adjacent prominent (stressed) syllables violates the constraint *CLASH which forces restrictions against two adjacent stressed (heavy) syllables (*CLASH: Adjacent prominent syllables are prohibited, cf. Kager 1999). There are potential problems by including this constraint. Specifically, *CLASH should dominate WSP since a heavy syllable adjacent to another heavy syllable needs to be stressless in violation of WSP. In line with many analyses of the Arabic dialects, we state that heavy syllables behave like secondary stresses, but are not marked as such; hence, *CLASH is irrelevant.

The dominance hierarchy is formulated as follows: RIGHTMOST, FT-BIN, ALIGN-L >> WSP >> NONFIN. Consider Table 6 of the word *mas'ʤid* ‘mosque’.

Table 6: RIGTHMOST >> WSP >> NONFIN

/masʤid/	RIGTHMOST	WSP	NONFIN
(a) ☞ ($ma_{\mu}s_{\mu}$)('$ʤi_{\mu}d_{\mu}$)			*
(b) ('$ma_{\mu}s_{\mu}$)($ʤi_{\mu}d_{\mu}$)	*!		
(c) ('$ma_{\mu}s_{\mu}$) $ʤi_{\mu}d_{\mu}$		*!	
(d) $ma_{\mu}s_{\mu}$('$ʤi_{\mu}d_{\mu}$)		*!	*

Heavy syllables surface stressed in final position by having WSP and RIGTHMOST dominate NONFIN. Violating NONFIN is more harmonic than violating WSP; hence WSP dominates NONFIN. Under evaluation, violating the dominated constraint NONFIN does not prevent the candidate ($ma_{\mu}s_{\mu}$)('$ʤi_{\mu}d_{\mu}$) from being the optimal output as long as it satisfies the undominated constraints WSP and RIGTHMOST. Satisfying WSP in (6b) is not enough to win the competition which confirms that RIGHTMOST >> WSP. The last heavy syllable receives the main stress violating the NONFIN constraint. Leaving a heavy syllable unfooted (as in 6d) would result in a fatal violation of the WSP constraint.

To account for stress assignment in disyllabic words comprising heavy and superheavy syllables, five main constraints are competing with each other in order to rule out the less optimal candidates. Recall that ALIGN -L and FT-BIN do not exhibit dominance ranking. Recall also that the word-final C in a superheavy syllable is extrasyllabic since the higher ranked FT-BIN requires final Cs in superheavy syllables to be weightless (extrasyllabic/extrametrical) in order to avoid trimoraicity (cf. Watson 2007), as Table 7.

Table 7: RIGHTMOST, FT-BIN, ALIGN-L >> WSP >> NONFIN

/mirjæ:ʕ/	RIGHTMOST	FT-BIN	ALIGN-L	WSP	NONFIN
(a) ☞ ($mi_{\mu}r_{\mu}$).('$ja_{\mu\mu}$)<$ʕ_{\mu}$>					*
(b) ($mi_{\mu}r_{\mu}$).('$ja_{\mu\mu}ʕ_{\mu}$)		*!			*
(c) ('$mi_{\mu}r_{\mu}$).($ja_{\mu\mu}$) ʕμ	*!				
(d) $mi_{\mu}r_{\mu}$.('$ja_{\mu\mu}ʕ_{\mu}$)		*!	*	*!	*

Both heavy syllables satisfy WSP, and thus EVAL is decided by other constraints. NONFIN prefers the candidate with stress on the initial syllable yielding ('mir)(jæ:)<ʕ>. However, the more harmonic winner is (mir)('jæ:)<ʕ>, where the final syllable is assigned the stress. This requires a constraint hierarchy that selects the main stressed syllable be the rightmost syllable, which proves that RIGHTMOST dominates NONFIN and WSP.

4.2.2.2 *Multisyllabic word stress*

With words comprising heavy and light syllables, PARSE-σ is required which indicates that syllables must be parsed into metrical feet.

(18) PARSE-σ: All syllables must be parsed into feet. (Prince & Smolensky 1993/2004)

Earlier, we demonstrated that FT-BIN dominates NONFIN in light syllable words of the shape LL and LLL, and WSP dominates NONFIN in heavy syllable words. On the other hand, NONFIN dominates PARSE-σ as NONFIN militates against footing final syllables especially when we consider words of the shape HLL, as in Table 8:

Table 8: WSP >>NONFIN >> Parse-σ

/jax.ti.ʃi/	WSP	NONFIN	PARSE-σ
(a) ☞ ('$ja_{\mu}x_{\mu}$) $ti_{\mu}ʃi_{\mu}$			**
(b) ($ja_{\mu}x_{\mu}$) ('$ti_{\mu}ʃi_{\mu}$)		*!	
(c) jax ('tiʃi)	*!	*	*

Table 8 shows that NONFIN dominates PARSE-σ; otherwise, candidate (b) would win the competition. Parsing the syllables (ti.ʃi) into a foot suggests that candidate (8b) should be the winner as it satisfies the highly ranked constraint RIGHTMOST. Since the last two syllables (ti.ʃi) were not parsed into a foot, (jax) is the rightmost visible syllable that bears the main stress.

In words comprising HHH structure, as in: *mitʕaʃji:n* 'they have eaten' (see Table 9), Candidate (9a) wins the competition to be the optimal output simply because it satisfies all the higher ranked constraints. Candidate (9b) fares worse on RIGHTMOST though it avoids violating NONFIN, WSP, and FT-BIN. Candidates (9c) and (9d) fall victim to RIGHTMOST and thus lose out to candidate (9a). Candidate (9e) falls victim to ALIGN-L, WSP, and NONFINAL.

Table 9: RIGHTMOST, FT-BIN, ALIGN-L >> WSP >> NONFIN

/mitʕaʃji:n/	RIGHTMOST	FT-BIN	ALIGN-L	WSP	NONFIN
(a) ☞ ($mi_{\mu}t_{\mu}$)($ʕa_{\mu}ʃ_{\mu}$)('$ji_{\mu\mu}$)n					
(b) ($mi_{\mu}t_{\mu}$) ('$ʕa_{\mu}ʃ_{\mu}$) ($ji_{\mu\mu}$)n	*!				
(c) ($mi_{\mu}t_{\mu}$) ('$ʕa_{\mu}ʃ_{\mu}$) ($ji_{\mu\mu}n_{u}$)	*!	*!			
(d) ('$mi_{\mu}t_{\mu}$) ($ʕa_{\mu}ʃ_{\mu}$) ($ji_{\mu\mu}$) n	*!				
(e) $mi_{\mu}t_{\mu}ʕa_{\mu}ʃ_{\mu}$ ('$ji_{\mu\mu}$)<n>			**	**!	*

In words comprising HLHH structure, as in: *jaʃtiri:hin* ‘he buys them’ (see Table 10), the optimal candidate is metrified as (jaʃ)(tiri:)('hin) with a LH iamb which insures that the dialect has a quantity-sensitive iamb.

Table 10: FTBIN, RIGHTMOST >> WSP >> NONFIN >> PARSE-σ

/jaʃ.ti.ri:.hin/	FT-BIN	RIGHTMOST	ALIGN-L	WSP	NONFIN	PARSE-σ
(a) ☞(jaʃ)(tiri:)('hin)					*	
(b) ('jaʃ)ti(ri:) (hin)		*				*
(c) (jaʃti) (ri:) ('hin)	*				*	
(d) (jaʃ) (ti) ('ri:) hin	*			*!	*	*
(e) jaʃ (ti) ('ri:) hin	*		*!	**!	*	**

In Table 10, candidate (10b) fails to parse the adjacent syllables into feet where the second syllable *.ti*. is left unfooted. The constraints hierarchy prefers (H)(LH)('H) over (HL)('H)(H). The result is that heads of binary feet are stressed and the rightmost visible foot receives the main stress.

In quadrisyllabic words comprising the HLLH structure, such as *maktabate:n* ‘two libraries’, parsing results as (mak)(taba)('te:)<n> forming three bimoraic feet. This means that among the possible parsing of syllables into feet as in: (H)(LL)(H), *H(LL)(H), *(H)LL(H), the optimal candidate should be (H)(LL)(H). The constraints hierarchy can be presented as follows: FT-BIN, RIGHTMOST, ALIGN-L >> WSP >> NONFIN >> PARSE-σ. This hierarchy can be manifested in examples like *maktabate:n* ‘two libraries’, as in Table 11:

Table 11: FTBIN, RIGHTMOST, ALIGN-L >> WSP >> NONFIN >> PARSE-σ

/maktabate:n/	FT-BIN	RIGHT MOST	ALIGN-L	WSP	NONFIN	PARSE-σ
(a) ☞($ma_{\mu}k_{\mu}$) ($ta_{\mu}ba_{\mu}$)('$te_{\mu\mu}$) <n>					*	
(b) $ma_{\mu}k_{\mu}ta_{\mu}ba_{\mu}$ ('$te_{\mu\mu}$) <n>			**!	*!	*	***
(c) ('$ma_{\mu}k_{\mu}$) $ta_{\mu}ba_{\mu}$ ($te_{\mu\mu}$) <n>		*!				**
(d) ($ma_{\mu}k_{\mu}$) ('$ta_{\mu}ba_{\mu}$) $te_{\mu\mu}n_{\mu}$		*!		*!		*
(e) $ma_{\mu}k_{\mu}(ta_{\mu}ba_{\mu})$ ('$te_{\mu\mu}$) <n>			*!	*!	*	*

In (maμkμ)($ta_{\mu}ba_{\mu}$)('$te_{\mu\mu}$) <n>, the ultimate receives stress by virtue of minimally violating the constraint hierarchy. Violating NONFINAL is more harmonic since it does not prevent candidate (11a) from winning the competition as long as it satisfies the higher ranked constraints FT-BIN, RIGHTMOST, and ALIGN-L.

Knowing that pre-antepenult syllables and heavy syllables do not always receive stress (cf. data in (2) and Tables 12 and 13 below), the present proposal is able to account for stress patterns in words comprising LLLL structure, as in: *na'ʕaʤata* 'his ewe' or HLLL structure, as in: *jista'hadi* 'to find (SG.M)' without the WINDOW constraint, which dictates that stress is confined to the last three syllables of a word (Kager 2012; Abu Guba 2016). Satisfying the higher ranked constraint RIGHTMOST is enough for candidate (12a) to win the competition over candidate (12b). Consider Table (12):

Table 12: FTBIN, RIGHTMOST, IAMB, ALIGN-L >> WSP >> NONFIN >> PARSE-σ

/yistahadi/	FT-BIN	RIGHT MOST	IAMB	ALIGN -L	WSP	NON FIN	PARSE -σ
(a) ☞(jis).(ta 'ha)di					*		*
(b) ('jis).tahadi							***
(c) jis.('ta ha) di			*	*	*!		**
(d) (jis). ta (ha 'di)						*!	*

The constraint hierarchy prefers (H)(LL)L over (H)LLL, H(LL)L, and (H)L(LL) and the other candidates. Stress falls on the penult light syllable violating WSP; the pre-antepenult *.jis.* does not trigger stress though it is a heavy syllable. A heavy syllable loses out to a light syllable in WRA where the constraint RIGHTMOST outranks WSP and guarantees that stress is not triggered by the heavy syllable in the pre-antepenult.

In four light syllable words, different parsings of such forms are possible: (L'L)(LL), (LL)(L'L), (L'LLL), LL(L'L), (L'L)LL, L(L'L)L, and (LL'L)L. The failure of stress of penultimate light syllables in LLLL structures, as in: *naʕaʤata* 'his ewe',[16] presumes that the rightmost foot is rendered extrametrical if peripherality does not exhaust the domain of the word (Hayes 1995: 58, *see* footnote *5*). The pre-antepenult syllable is invisible to the main stress; consider Table (13).

Table 13: FT-BIN, RIGHTMOST, IAMB, ALIGN-L >> NONFIN

/naʕaʤata/	FT-BIN	RIGHTMOST	IAMB	ALIGN-L	NONFIN
(a) ☞ (na_{μ}'$ʕa_{\mu}$) $ʤa_{\mu}$ ta_{μ}					*
(b) ('naʕa)(ʤata)		*	*		
(c) na('ʕaʤa)ta			*	*	*

We argue that the parsing (L'L)LL[17] is the most harmonic structure in WRA. Recall that FT-BIN, RIGHTMOST, and IAMB dominate NONFIN in the hierarchy. To yield such a parsing, RIGHTMOST is not violated by the optimal output (13a) whenever the peripheral foot <*ʤa.ta*> counts as extrametrical because it does not exhaust the domain of the word and thus is invisible to the main stress. Candidate (13b) does not satisfy the undominated constraints RIGHTMOST and IAMB so that the pre-antepenult fails to win; ALIGN-L eliminates candidate (13c) as it ensures that the word starts with a foot erected at the left edge. RIGHTMOST and IAMB assign stress to the rightmost visible foot yielding (L'L) LL.

[16] Such forms are always subject to epenthesis in WRA, e.g., *naʕja* 'ewe' > *naʕaʤa*.

[17] Here extrametrical foot is not marked in the candidates nor in the input since it is a consequence of the constraint ranking.

5. Concluding remarks

The present work has examined stress patterns in WRA adopting moraic analysis within an OT framework. The analysis confirms that stress in WRA is fairly regular. As expected, foot extrametricality is blocked when it is the only foot in the word or when the rightmost foot is non-peripheral. For final superheavy syllables (CVVC, CVCC), the final C is unsyllabified until the stress rules apply, and thus prevents the rightmost foot from being peripheral, i.e., the unsyllabified final C blocks the rightmost foot from being extrametrical, which allows it to receive the main stress.

By contrast to Eastern Arabic dialects, WRA, a JA variety has an iambic foot in which the right head is assigned the main stress. Applying metrical parameters, WRA shows accentual parallels with Negev Bedouin (Blanc 1970); they share foot type, foot extrametricality, and stress assignment according to ERR. This entails that a future serious revision of the literature of JA dialects should be carried out.

Recognizing that the foot is iambic in WRA, the stress rules outlined in this study are able to account for all stress patterns in WRA. The study has concluded with a limited and economic number of universal constraints that cover all the possible patterns of stress in WRA. Heavy syllables in quantity sensitive languages (including WRA) attract stress assignment. WRA imposes a weak prohibition on degenerate feet as the final C in word-final CVC syllables counts to be moraic supported by a number of CVC content words that surface as they are, without resorting to a final consonant gemination or vowel lengthening.

Examining words of different syllable types (CV.CVC, CV.CV.CV, CVC.CVVC, CVC.CVV.CVVC, CVC.CV.CV.CV), the study has come up with this simplified minimal dominance hierarchy: *PrWdμ, IAMB, RIGHTMOST, ALIGN-L, FT-BIN >> WBP, WSP >> *FINAL-C-μ, >> NONFIN >> PARSE-σ. The proposed dominance hierarchy has introduced a new constraint namely ALIGN-L which was ignored in the literature of JA varieties.[18] This constraint necessitates a prosodic word to be left-aligned with a foot. This constraint hierarchy proposed its ability to account for all possible structures without resorting to WINDOW constraint which ensures that stress is assigned within the three syllable window. The study has shown how metrical *iambicity* parameter struggles with optimal *nonfinality* constraint as well as with other parameters/constraints to generate the optimal stress output. WBP also competes with *FINAL-C-μ. Contrary to Eastern Arabic varieties, WBP is ranked higher than*FINAL-C-μ in the dominance hierarchy.

Further research would be interesting as well as promising to answer why stress patterns in WRA are different from the other neighbouring dialects and most of the Eastern Arabic varieties such as Bani Hassan Bedouin and Ahl Al-Jabal Bedouin (Mashaqba & Huneety 2017). Would it be possible that it once had the stress patterns found in other Eastern Arabic dialects (moraic trochaic), and underwent a shift to an iamb? If so, what parameters and processes that are involved to capture such outcomes? Is it by Move X (cf. Hayes 1995: 34-37)? We here leave the answer for future serious research. Additionally, future research may be suggested to investigate opaque stress patterns in JA, if any, which involve interaction with syncope, epenthesis or vowel quality.

[18] This constraint is new in the way it is incorporated in the constraint hierarchy of Arabic stress patterns.

References

Abdo, Daud A. 1969. *On stress and Arabic phonology*. Beirut: Khayats.

Abu-Abbas, Khalid. H. 2003. *Topics in the Phonology of Jordanian Arabic.* University of Kansas. (PhD dissertation.)

Abu Guba, Mohammad. 2016. *Phonological adaptation of English loanwords in Ammani Arabic*. University of Salford. (PhD dissertation.)

Al-Ani, Salman. 1992. Stress variation of the construct phrase in Arabic: A spectrographic analysis. *Anthropological Linguistics* 34. 256–276.

Al-Jarrah, Rasheed. 2002. *An optimality-theoretic analysis of stress in the English of native Arabic speakers.* Ball State University. (PhD dissertation.)

Al-Jarrah, Rasheed. 2011. Extrametricality Revisited. *Linguistic and Literary Broad Research and Innovation* 2, (1). 3–35.

Angoujard, Jean-Pierre. 1990. *Metrical structure of Arabic*. Dordrecht: Foris.

Blanc, Haim. 1970. *The Arabic dialect of the Negev Bedouins*. Jerusalem: Jerusalem Academic Press.

Brame, Michael K. 1970. *Arabic phonology: Implications for phonological theory and historical Semitic*. Massachusetts Institute of Technology. (PhD dissertation.)

Brame, Michael K. 1973. On stress assignment in two Arabic dialects. *A Festschrift for Morris Halle*, ed. Stephen R. Anderson and Paul Kiparsky. New York: Holt, Rinehart and Winston. 14–25.

Brame, Michael K. 1974. The cycle in phonology: Stress in Palestinian, Maltese and Spanish. *Linguistic Inquiry* 5. 39–60.

Broselow, Ellen. 1976. *The Phonology of Egyptian Arabic*. University of Massachusetts. (PhD dissertation.)

Broselow, Ellen. 1982. On predicting the interaction of stress and epenthesis. *Glossa* 16. 115–132.

Broselow, Ellen. 1992. Parametric variation in Arabic dialect phonology. *Perspectives on Arabic Linguistics, John Benjamins, Amsterdam* 4. 7–46.

Broselow, Ellen, Marie Huffman, Su-I Chen & Routhmei Hsieh. 1995. The timing structure of CVVC syllables. In Eid, Mushira (ed.), *Perspectives in Arabic Linguistics,* 7, 119–140. Amsterdam, Philadelphia: John Benjamins.

Broselow, Ellen, Su-I Chen & Marie Huffman. 1997. Syllable weight: Convergence of phonology and phonetics. *Phonology* 14. 47–82.

Cantineau, Jean. 1939. Remarques sur les parlers de sédentaires syro-libano-palesintinens. *Bulletin de la Société de Linguistique de Paris* 40. 80–88.

Crowhurst, Megan. J. 1996. An Optimal Alternative to Conflation. *Phonology*13. 409–24.

Farwaneh, Samira. 1995. *Directionality effects in Arabic dialect syllable structure*. University of Utah. (PhD dissertation.)

Halle, Morris & Michael Kenstowicz. 1991. The free element condition and cyclic versus noncyclic stress. *Linguistic Inquiry* 22. 457–501.

Harrell, Richard. 1957. *The phonology of Colloquial Egyptian Arabic*. New York: American Council of Learned Societies.

Hayes, Bruce. 1981. *A metrical theory of stress rules*. Massachusetts Institute of Technology. (PhD dissertation.)

Hayes, Bruce. 1982. Extrametricality and English Stress. *Linguistic Inquiry* 13. 227–276.

Hayes, Bruce. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20. 253–306.

Hayes, Bruce. 1995. *Metrical stress theory: Principles and case studies.* Chicago: University of Chicago Press.

Hérin, Bruno. 2010. *Le parler Arabe de Salt (Jordanie). Phonologie, morphologie et éléments de syntaxe*. [The Arabic Dialect of aṣ-Ṣalṭ (Jordan). Phonology, morphology and elements of syntax]. Université Libre de Bruxelles. (PhD dissertation.).

Huneety, Anas. 2015. The Phonology and Morphology of Wadi Mousa Arabic. University of Salford. (PhD dissertation.).

Huneety, Anas, & Bassil Mashaqba. 2016. Stress rules in loan words in Bedouin Jordanian Arabic in the north of Jordan: a metrical account. *SKASE Journal of Theoretical Linguistics* [online]. 13, 3: 1–13. Available on web page http://www.skase.sk/Volumes/JTL33/pdf_doc/01.pdf. ISSN 1336– 782X.

Hyde, Brett. 2007. Non-finality and weight-sensitivity. *Phonology* 24(2). 287–334.

Hyde, Brett. 2011. Extrametricality and nonfinality. In Marc van Oostendorp, Colin Ewen, Beth Hume & Keren Rice (eds.) *A Companion to Phonology*, 1027–1051. Blackwell.

Hyman, Larry. 1985. A *theory of phonological weight.* Dordrecht: Foris.

Irshied, Omar. 1984. *The phonology of Arabic: Bani Ḥassan, a Bedouin Jordanian dialect*. University of Illinois. (PhD dissertation.).

Itô, Junko. 1989. A prosodic theory of epenthesis. *Natural Language and Linguistic Theory* 7, 217–259.Johnson, C. Douglas. 1979. Opaque stress in Palestinian. *Lingua* 49.153–168.

Kager, René.1995. *The metrical theory of word stress*. In Goldsmith, John. (ed.), The handbook of phonological theory, 376–402. Cambridge, Mass. Blackwell.

Kager, René. 1999. Optimality Theory. Cambridge: Cambridge University Press.

Kager, René. 2009. Stress. In Versteegh, Kees, Mushira Eid, Alaa Elgibali, Manfrid Woidich & Andrzej Zaborski (eds.), Encyclopedia of Arabic Language and Linguistics 4. 344–353. Leiden – Boston – Cologne: Brill.

Kager, René. 2012. Stress in windows: Language typology and factorial typology. *Lingua* 122(13). 1454–1493.

Kenstowicz, Michael. 1981. A note on cyclic stress in Levantine Arabic. In Hagit Borer & Joseph Aoun (ed.), *Theoretical issues in the grammar of Semitic languages*, 21–31. (= *MIT Working Papers in Linguistics*, 3.) Cambridge, Mass.: Department of Linguistics and Philosophy, Massachusetts Institute of Technology.

Kenstowicz, Michael. 1983. Parametric variation and accent in the Arabic dialects. *Proceedings of the Chicago Linguistic Society* 19. 205–213.

Kenstowicz, Michael. 1994. Phonology in generative grammar. Oxford: Blackwell.

Kiparsky, Paul. 2002. Syllables and moras in Arabic. In Caroline Féry & Ruben van de Vijver (eds.), 147–182. *The syllable in Optimality Theory*, Cambridge: Cambridge University Press.

Kiparsky, Paul. 2003. Syllables and moras in Arabic. In C. Féry and R. van de Vijver (eds.), 147–182. *The syllable in Optimality Theory*. Cambridge: Cambridge University Press.

Liberman, Mark & Alan Prince. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8. 249–336.

Mashaqba, Bassil. 2015. The Phonology and Morphology of Wadi Ramm Arabic. University of Salford. (PhD dissertation).

Mashaqba, Bassil & Anas Huneety. 2017. Morpho-phonological structure of sound feminine plural: Revisited. *International Journal of Applied Linguistics & English Literature* 6 (6). 115–122.

McCarthy, John. 1979a. On stress and syllabification, *Linguistic Inquiry* 10. 443–465.

McCarthy, John. 1979b. *Formal Problems in Semitic Phonology and Morphology*. MIT. (PhD dissertation.).

McCarthy, John. 1980. A note on the accentuation of Damascene Arabic. *Studies in the Linguistic Sciences* 10, 77–98.

McCarthy, John. 2007a. Hidden generalizations. Phonological opacity in OT. London: Equinox.

McCarthy, John. 2007b. What is Optimality Theory? *Language and Linguistics Compass* 1. 260–291.

McCarthy, John. 2008. Doing Optimality Theory: Applying theory to data. Oxford: Blackwell.

McCarthy, John & Alan Prince.1993. Generalized alignment. In *Yearbook of morphology* 79-153. Springer, Dordrecht.

McCarthy, John & Alan Prince. 1995. Faithfulness and reduplicative identity. In J. Beckman, L. W. Dickey & S. Urbanczyk (Eds.), *Papers in Optimality Theory*. *University of Massachusetts Occasional Papers* 18. (pp. 249–384). Amherst, Mass.: Graduate Linguistic Student Association. [Rutgers Optimality Archive 60, http://roa.rutgers.edu].

Mitchell, Terence F. 1956. *An introduction to Egyptian Colloquial Arabic*. London: Oxford University Press.

Mitchell, Terence F. 1960. Prominence and syllabification in Arabic. *Bulletin of the School of Oriental and African Studies* 23. 369–389.

Palva, Heikki. 1976. *Studies in the Arabic dialect of the semi-nomadic al Aégæ:rme tribe (al-Balqæ:' District, Jordan).* (Göteborg): Acta Universitatis Gothoburgensia, distr.

Palva, Heikki. 1980. Characteristics of the Arabic dialect of the Bani Ṣaxar tribe. *Orientalia Suecana Uppsala* 29. 112–139.

Palva, Heikki. 1986. Characteristics of the Arabic dialect of the Ḥwe:ṭæ:t tribe. *Orientalia Suecana* 33–35. 296–312.

Palva, Heikki. 1991. Is there a North West Arabian dialect group? In: M. Forstner (ed.) *Festgabe für Hans-Rudolf Singer, zum 65. Geburtstag am 6. April 1990 überreicht von seinen Freunden und Kollegen*, Part 1. Frankfurt: Peter Lang. 151–166.

Prince, Alan. 1990. Quantitative consequences of rhythmic organization. In K. Deaton, M. Noske, & M Ziolkowski (Eds.), CLS 26(2): Papers from the Parasession on the Syllable in Phonetics and Phonology. (pp. 355-398). Chicago: Chicago Linguistic Society.

Prince, Alan & Paul Smolensky. 1993 [2004]. Optimality Theory: Constraint Interaction in Generative Grammar. Oxford: Blackwell.

Rakhieh, Belal. 2009. *The phonology of Ma'ani Arabic*: *Stratal or parallel OT*. University of Essex. (PhD dissertation.).

Rosenthall, Sam & Harry Van Der Hulst. 1999. Weight-by-position by position. *Natural Language & Linguistic Theory* 17. 499–540.

Sakarna, Ahmad. Kjalaf. 1999. *Phonological aspects of 9Abady Arabic, a Bedouin Jordanian dialect*. University of Wisconsin-Madison. (PhD dissertation.).

Sakarna, Ahmad. Kjalaf. 2002. The Bedouin dialect of Al-Zawaida tribe, southern Jordan. *Al-Arabiyya* 35. 61–86.

Selkirk, Elisabeth. 1980. On prosodic structure and its relation to syntactic structure. In: T. Fretheim (ed.). *Nordic Prosody* 2. Trondheim: TAPIR. 111–140.

Selkirk, Elisabeth. 1984. P*honology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.

Selkirk, Elisabeth. 1995. The prosodic structure of function words. In Jill Beckman, Laura Walsh Dickey and Suzanne Urbanczyk (eds.) University of Massachusetts Occasional Papers in Linguistics 19: Papers in Optimality Theory. Amherst, MA: GLSA. 439–469.

Watson, Janet. C. E. 2002. *The phonology and morphology of Arabic*. Oxford: Oxford University Press.

Watson, Janet. C. E. 2007. Syllabification patterns in Arabic dialects: Long segments and mora sharing. *Phonology* 24. 335–356.

Watson, Janet. C. E. 2011. Word stress in Arabic. In: M. van Oostendorp et al (eds.). *The Blackwell companion to phonology*. Wiley-Blackwell, Oxford. 2990–3019.

Welden, Ann. 1980. Stress in Cairo Arabic. *Studies in the Linguistic Sciences* 10, 99–120.

Wright, William. 1859. *A grammar of the Arabic language*. Cambridge: Cambridge University Press. (Repr., 1971).

Yasin, Raslan. 1980. *A critical and comparative study of the dialectal speech of the Ghawarna community in the Jordan Valley in Jordan.* University of Leeds. (PhD dissertation.).

Younes, Munther. 1995. On vowel shortening in Palestinian Arabic. In Mushira Eid (ed.), *Perspectives on Arabic linguistics 7*, 157–172. Amsterdam & Philadelphia: John Benjamins.

Appendix

Example words in WRA with special focus on syllable types and stress patterns

Monosyllabic words		
CV	['hu]	‘he’
CVC	['lak]	‘yours m.s.’
CVVC	['læ:ʃ]	‘not a good man’
CVCC	['ʃadd]	‘to tie’
CCVCC	['bṣall]	‘onions’
CCVVC	['ʃdæ:d]	‘saddle and blanket on camel’
Bisyllabic words		
CV.CV	[li.'fa]	‘to come’
CV.CVC	[la.'ḥam]	‘meat’
CVV.CV	['kæ:.fi]	‘enough’
CV.CVCC	[ʃi.'ribt]	‘I drank’
CVC.CVVC	[mir.'jæ:ʕ]	‘ram leading the herd’
CVC.CV	['ḥaf.li]	‘party’
CVV.CVC	[næ:.'gil]	‘newly pregnant’
CVV.CVCC	[ḥæ:.'babt]	‘I kissed’
CCVCC.CVVC	[mʕazz.'bi:n]	‘guests (M)’
Trisyllabic words		
CV.CV.CV	[fa.'ru.na]	‘our fur’
CV.CV.CVC	[ʔa.ḥa.'mar]	‘red’

CVC.CV.CVCC	[ʔaʃ.ta.'ʁalt]	'I worked'
CVC.CVCC.CVVC	[mit.waʤʤ.'hi:n]	'heading to (PL, M)'
CVC.CVV.CVVC	[ḥaṭ.ṭæ:.'bæ:t]	'women preparing firewood'
CVC.CV.CVC	[jax.ti.'riʕ]	'to frightened'
CV.CVV.CV	[ʔa.'xu:.na]	'our brother'
CV.CVC.CV	[ʤi.'mal.ha]	'her camel'
Polysyllabic words		
CV.CV.CVC.CVC	[ji.ti.gaṭ.'ṭam]	'to break down'
CVC.CV.CVCC.CV	[miʕ.ni.'gij.ji]	'pure-bred horse'
CVC.CVC.CVV.CV	[ʔim. ʕaʃ.'ʃæ:.ha]	'its pasture'
CV.CVC.CVC.CVC	[ʔin.daʤ.'ʤin.ha]	'we domesticate it (F)'
CV.CVC.CV.CVVC	[mu.ʤam.ma.'ʕæ:t]	'bus stations'
CV.CV.CVV.CVVC	[ta.la.fo:.'næ:t]	'telephones'
CVC.CV.CVC.CVC.CVC	[miʕ.na.gij.jit.'hum]	'their (M) horse'

Bassil Mashaqba
Department of English Language and Literature
The Hashemite University
Zarqa, Jordan
almashaqba@live.com or b_mashaqba@hu.edu.jo

Anas Huneety
Department of English Language and Literature
The Hashemite University
Zarqa, Jordan
anasi@hu.edu.jo

In SKASE Journal of Theoretical Linguistics [online]. 2018, vol. 15, no. 3 [cit. 2018-13-12].

Available on web page http://www.skase.sk/Volumes/JTL38/pdf_doc/02.pdf. ISSN 1336-782X.

The linking morpheme in Afrikaans: a Cognitive Grammar description*

Eddie Benito Trollip & Gerhard B. van Huyssteen, North-West University, South Africa

In Germanic languages the linking morpheme, like the ·s· *in Afrikaans* seun·s·naam *'boy's name', or* ·en· *in Dutch* pann·en·koek *'pancake' is quite common. This word element has been the topic of discussion in the past, with no definite consensus about its origin or possible semantic input. There has been a renewed interest in this phenomenon, especially during the last few years, and not exclusively for Germanic languages. The objective of this paper is to categorise the linking morpheme in Afrikaans in terms of principles from Cognitive Grammar culminating in the postulation of the linking morpheme in two categorisation networks. The goal to construct categorisation networks are met in the conclusion to the paper, and it is concluded that the function of the linking morpheme is semantically highly schematic, but not functionally negligible.*

Keywords: *Afrikaans, Cognitive Grammar, linking element, linking morpheme, morphology*

1. Introduction

Linking morphemes (most often called a linking element, but also known as an interfix, link phoneme, phonomorpheme, connecting morpheme, linker, stem extender, and valence morpheme, amongst many others) are found in many languages of the world. In this article we consider Afrikaans linking morphemes, such as the ·*e*· in *hond*·***e***·*hok* (dog·LK·cage; 'kennel'), and the ·*s*· in *seun*·***s***·*naam* (boy·LK·name; 'boy's name'). For reasons that will become apparent, we use the term 'linking morpheme', instead of the more widely used "linking element".

In the past few years linking morphemes have been the subject of a number of large-scale linguistic enquiries, including Fuhrhop & Kürschner (2015), Krott et al. (2007), Van Tiel et al. (2011), and Wegener (2008), to name but a few. The questions raised in these projects ranged from the theoretical (e.g. the possible morphemic status of this word element), to the descriptive (e.g. historical origins, current uses, and productivity). Specifically in Dutch there has been a decades long investigation into the possible meaning of linking morphemes, from Mattens (1970), to most recently Hanssen (2011), and Banga et al. (2012; 2013). Similarly German has profited from studies especially highlighting the phonological value of these morphemes, like Krott et al. (2007), and Nübling & Szczepaniak (2013). Research on linking morphemes continues to this day, as is evident from the recent investigation by Schäfer & Pankratz (2018) into the plural interpretability of linking morphemes in German.

In contrast to this body of work, the status of linking morphemes in Afrikaans still remains largely unexplored. Apart from some remarks made in passing by a handful of Afrikaans linguists, writing exclusively in Afrikaans (i.e. Combrink 1990; Kempen 1969), no substantive, comprehensive and unifying description of Afrikaans constructions with linking morphemes exist – written in either in Afrikaans, or English. The main aim of this article is therefore to fill this gap in the international descriptive literature on linking morphemes.

* We would like to express our gratitude for insightful and constructive comments and suggestions made by the reviewers of this article.

A secondary aim also relates to the descriptive nature of this article, albeit on a more meta-level, namely to demonstrate how Cognitive Grammar (hereafter CG) can be used as a descriptive framework for morphological constructions. CG (see the two-volume *Foundations of Cognitive Grammar*; Langacker 1987, 1991) is one of the earliest sub-theories of what would become known as the Cognitive Linguistics enterprise (Evans & Green 2006). As such, CG is also one of the oldest construction grammar theories, and has been used widely in the description of numerous grammatical constructions in various languages. However, compared to especially lexical, syntactic, and discourse studies, the use of CG in morphological descriptions has been rather scant. In addition to some writings by Langacker (e.g. 1990) and Taylor (e.g. 2002, 2015), two of the main proponents of CG, and an overview by Evan & Green (2006), the only other significant body of morphological research within this framework is by Tuggy (e.g. 2003, 2005) and Hamawand (2011). Van Huyssteen (2010) mentions several other morphological studies that have been done within the broader Cognitive Linguistics paradigm, though not specifically using CG as descriptive framework (e.g. Janda 2011; Manova 2011). This article therefore strives to contribute to this relatively small body of literature employing CG.

Of course, one would immediately ask why there is only such a small body of literature. Is CG perhaps not appropriate for morphological descriptions? There might be two main reasons why CG has not caught on as a popular morphological theory. Firstly, mainstream CG specialists have tended up to now to focus more on 'larger' constructions, such phrase, sentence and discourse constructions; 'smaller' constructions (like morphological constructions) have been mentioned in passing, or were described in isolated publications. This, however, does not imply that CG is not appropriate for morphological descriptions. On the contrary, Tuggy (2005) makes a convincing case for CG and other constructionist approaches to describe and explain various morphological phenomena that are difficult to account for in other theories. As will be illustrated in this article, we believe that the linking morpheme is another such a phenomenon that will benefit from a CG/constructionist treatment.

Secondly, Booij's theory of Construction Morphology (hereafter CM; 2010) became the de facto flavour of constructionist approaches to morphology, thereby overshadowing other sub-theories like CG, Cognitive Construction Grammar (Goldberg 2006), or Radical Construction Grammar (Croft 2001). If this creates the impression that CM and CG are at odds, nothing could be further from the truth. In our opinion, these two theories are both sub-theories of the general theory of Construction Grammar (CxG), within the broader Cognitive Linguistics paradigm. As such, CM and CG could and should be used in tandem, as was illustrated already in Van Huyssteen (2018). Continuing along these lines, we will demonstrate in this article that we generally work within and subscribe to the tenets of CM, while using specific tools and constructs from CG mainly for two purposes, namely:

(a) to give fine-grained descriptions of the realisational (i.e. phonological or orthographical) and conceptual (i.e. semantic details) of morphemes (component constructions) and complexes (composition constructions); and
(b) to construct (visual) categorisation networks, providing an overview of a specific morphological construction, its schemas and instantiations, while also explicating the interrelationships among constructions and their allomorphic variants.

In addition to the aforementioned primary and secondary aim, this article has another secondary aim, namely to provide a CG perspective on the age-old question of whether the linking morpheme is indeed a morpheme, i.e. a form-function pairing serving as a component

construction in morphological composition constructions. This question has been considered from different approaches, each highlighting different aspects of the linking morpheme. For example, Neef (2015) criticises linguists that follow a functional approach to linking morphemes, and holds that they are searching for meaning/content for this "morpheme", while it is nothing more than a form of stem allomorphy.

Contrary to Neef's (2015) viewpoint, the linking morpheme in Afrikaans will be characterised in this article as a morpheme, albeit with minimal form (prototypically consisting of only one grapheme/phoneme), and highly schematic conceptual content (i.e. highly abstract or vague meaning, to the extent that it is mostly meaningless from a synchronic viewpoint). In addition, it will be illustrated that one of the linking morpheme's functions is to create allomorphs with the purpose to increase the valence of component constructions to combine with other components. This opinion is grounded in the constructionist (and specifically CG) view that a morpheme is the smallest/minimal symbolic manifestation in language, which cannot be analysed into smaller meaningful parts (Langacker 2013: 16). Despite its 'size' and schematicity, the linking morpheme still contributes to the overall construal of a composite construction, since:

> […] the meaning of many linguistic elements – especially those considered "grammatical" – consists primarily in the construal they impose, rather than any specific content. Yet every element evokes some content (however schematic it might be), and conversely, any content evoked is construed in some fashion (Langacker 2008: 43).

The discussion will begin with an account of constructions in CG (§2), followed by the specific application of linking morphemes in Afrikaans in terms of composition (§3.1) and entrenchment (§3.2). In §4 Afrikaans corpus data will be analysed to ultimately postulate a categorisation network for the linking morpheme in compound and non-compound words. The origin, nature, and structure of the corpora form part of the discussion in §4. The article concludes with a discussion of the two categorisation networks for the linking morpheme in Afrikaans.

2. Constructions in CG

A construction is any symbolic form-function pairing in a language (Langacker 2013: 15). In constructions such as (1), which is a representation of the word *eend* 'duck' in Afrikaans, the uppercase letters symbolise the conceptualised idea (i.e. meaning) of a water bird with webbed feet (on the pole of conceptualisation, also known as the semantic pole), while the lower case letters represent the realisation (i.e. form) of the idea on the realisation pole (also known as the phonological pole). Square brackets are used to show that the concept is already an accepted word in the language; normal/rounded brackets are used for unknown examples such as neologisms and newly constructed compounds.

(1) [[DUCK]/[eend]]
'duck'

Importantly, note that in accordance with Langacker (2013: 15) any formal realisation, whether phonological or orthographical, is taken into account in the construal of constructions. When

referring to the realisation/phonological/orthographical pole in this article, the actual orthographical realisation will be used as representation rather than the sounds that the construction consists of. Van Huyssteen (2018:405) reiterates the role orthographical representation plays in the overt symbolisation of meaning in CG, and therefore motivates why orthographical elements like hyphens could also be considered as linguistic elements. As will be indicated in this article, the hyphen sometimes fulfils a valence function (e.g. to avoid potential readability problems), and at other times as semantic function (i.e. to indicate a coordinative relation between constituents in compounds). As such, the hyphen complies to the general definition of what a morpheme is, namely a form-function pairing, despite the fact that the form is not realised in traditional letters or sounds.

Constructions need not be fully specified: words like the example in (1), or the representation of the plural construction in (2a), are both constructions, seeing as both have a semantic and phonological pole. The absence of a middle dot (·)[1] in the case of (1) denotes phonological independency which is not the case with (2a), seeing as (2a) is a suffix that must combine with other constructions to be able to function (Van Huyssteen 2017: 186). When these constructions combine, they form complex constructions (Langacker 2013: 15), which are represented as in (2b), and which can be simplified notationally as in (2c). Note that 'THING' is used in CG in a rather technical sense, referring to entities that are profiled by nouns.

(2) a. [[PL]/[·s]]
b. [[[THING]/[X]]/[[PL]/[·s]]]
c. [THING·PL/X·s]

When two or more constructions are combined, phonological and semantic dependency comes into play (to be discussed as part of composition in §3.1). An example of a composition structure is given in (3a) where two independent component structures ([EEND/eend] and [HOK/hok]) combine *with a highly dependent component [LK/·e·]*.

(3) a. [DUCK·LK·CAGE/eend·e·hok]
'duck's cage'
b. [DUCK·LK·FARMER/eend·e·boer]
'duck farmer'
c. [DUCK·LK·LIVER/eend·e·lewer]
'duck's liver'

Other composition structures, like those in (3b) and (3c), clearly share commonalities with (3a). These commonalities can be represented as a constructional schema (Langacker 2013: 219). Constructional schemas are the way in which knowledge of linguistic patterns are expressed (Evans et al. 2007: 25) – schemas in CG fulfil a similar role as rules in generative grammar, namely to model our knowledge of patterns of commonalities in language use (Langacker 2013: 23). Similar to constructions, constructional schemas are not limited in their level of specificity (Langacker 2013: 24), and because of this characteristic, (4a) (together with Figure 1), (4b) (together with Figure 2) and (4c) (together with Figure 3) all serve as increasingly

[1] Seeing as hyphens are analysed as graphemic linking morphemes, it would be confusing to use them to indicate morpheme boundaries. For this reason, middle dots (·) are consistently used in this paper to indicate morpheme boundaries, following the tradition of Bauer (2003).

schematic constructional schemas of the *eend* composition structures. Ellipses are used to indicate a high level of schematicity, or non-specificity.

(4) a. [DUCK·LK·CAGE/eend·e·hok]
‘duck’s cage’
b. [DUCK·LK·THING/eend·e·...]
‘duck·LK·THING’ (any compound with left-hand constituent *eend*, followed by a linking morpheme and another noun, e.g. *eend·e·plaas* ‘duck farm’, *eend·e·boerdery* ‘duck farming’, and *eend·e·dam* ‘dam for ducks’.)
c. [THING·LK·THING/...$_1$·e·...$_2$]
‘THING·LK·THING’ (any noun-noun compound with a linking morpheme)

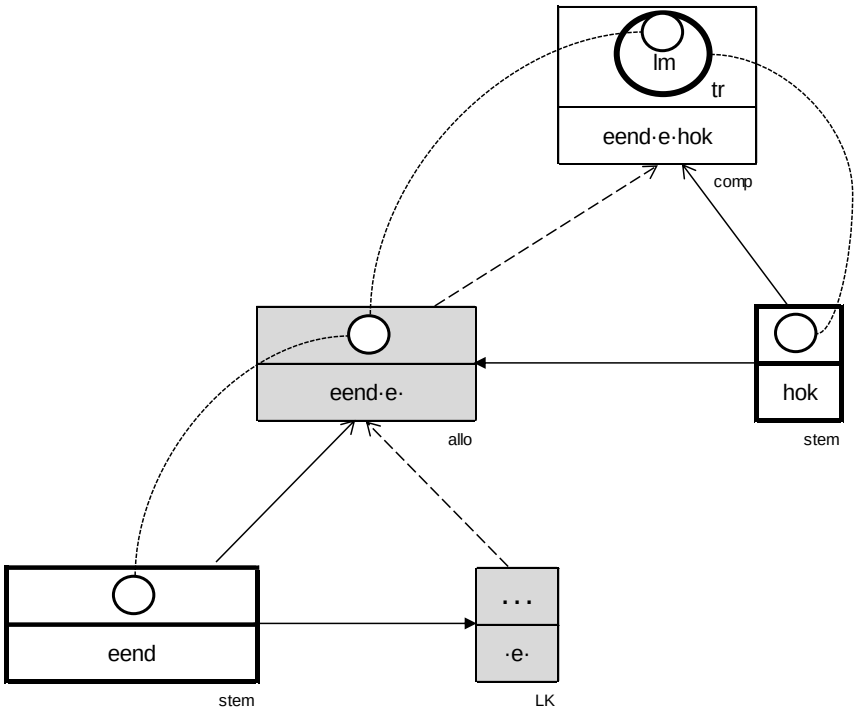


Figure 1: Schema of [EEND·E·HOK/eend·e·hok] ‘duck’s cage’

These schemas are illustrative of the way in which morphological constructions are described in CG. Langacker (2013: 10) emphasises that any visual representation serves a heuristic function; it is merely an aid to explain a certain linguistic phenomenon. In visual representations, halved rectangles are component structures. Things (nouns), irrespective of their status as component or composition structures, are represented as circles. The dark border around some rectangles represents the important role of that specific component structure as the profile determinant in the formation of a more complex construction. In §3.1.1 constructional profiling and the representation of constructional schemas with the use of, among others, correspondence lines and ‘tr’ and ‘lm’ will be discussed in detail. On the realisation pole, structures are represented with lower case lettering. The abbreviation ‘allo’ is used to indicate allomorphic structures; ‘comp’ is used to indicate compounds; ‘LK’ for linking morphemes; and ‘stem’ labels an independent component (either words or stems). In subsequent figures (e.g. Figure 6) the abbreviation ‘aff’ is used for affixed forms.

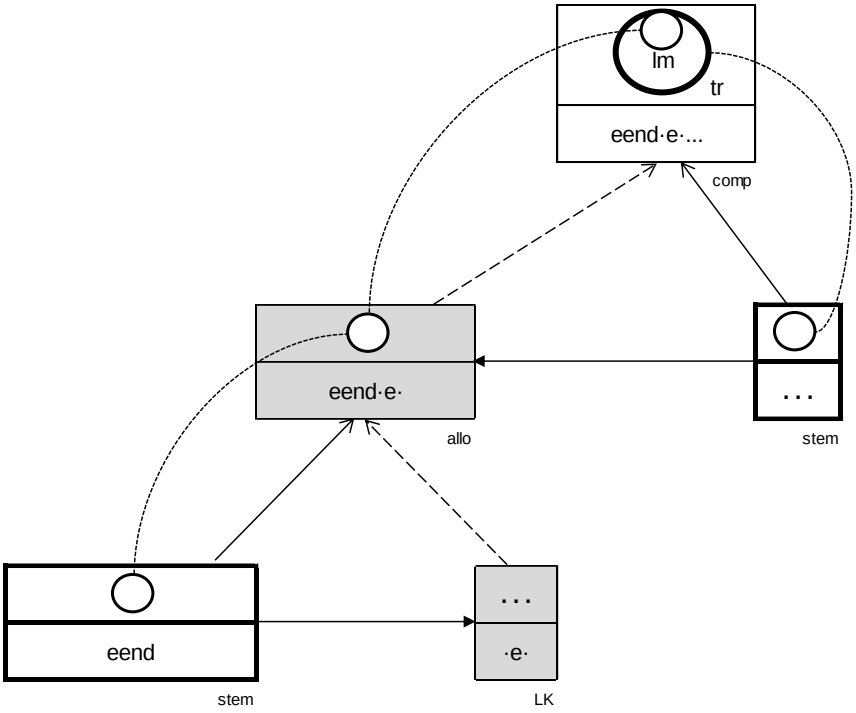


Figure 2: Schema of [EEND·LK·THING/eend·e·...]

A crucial difference between the representation in (4a) and Figure 1 is the arrows. The arrows with the black heads in Figure 1 represent the order in which component structures combine. This arrow always points in the direction of the component structure that elaborates the composition structure on the next level of constituency. The arrow originates from the profile determinant or the component that determines the *nature* of the composition structure on the next level of constituency. The broken line arrow and normal arrow (with open heads) that feature between the levels, join the profile determinant (normal arrow) with the component that serves to elaborate the profile determinant (broken line arrow) (in §3.1.1 it will be made clear what is meant with profile determinant.) Another difference between the representation in (4a) and Figure 1, is the grey colouring of the constructions that are semantically and/or phonologically dependent.[2] The shading indicates the dependency of the component structures, and that it is necessary for the dependent structures to combine with at least one other, more independent component structure (like an affix or independent word) to function.

Figure 3 is the most abstract of the constructional schemas and makes use of two wholly unspecified component structures on the realisation pole. It is important to remember that the construal of more general, higher-level schemas (abstract schemas) is not preferred to the construal of lower level schemas (concrete schemas). It is a characteristic of CG that distinguishes it from other approaches: more general schemas are important to show commonalities, but an extensive network of abstract and specified/concrete schemas is necessary to postulate a complete description of any linguistic phenomenon (Tuggy 2003: 28). The implication here is that lower level schemas have to be included in the description of the linking morpheme.

[2] Langacker uses hatching to represent non-specifity/dependency. See Langacker (2013: 198) for an example.

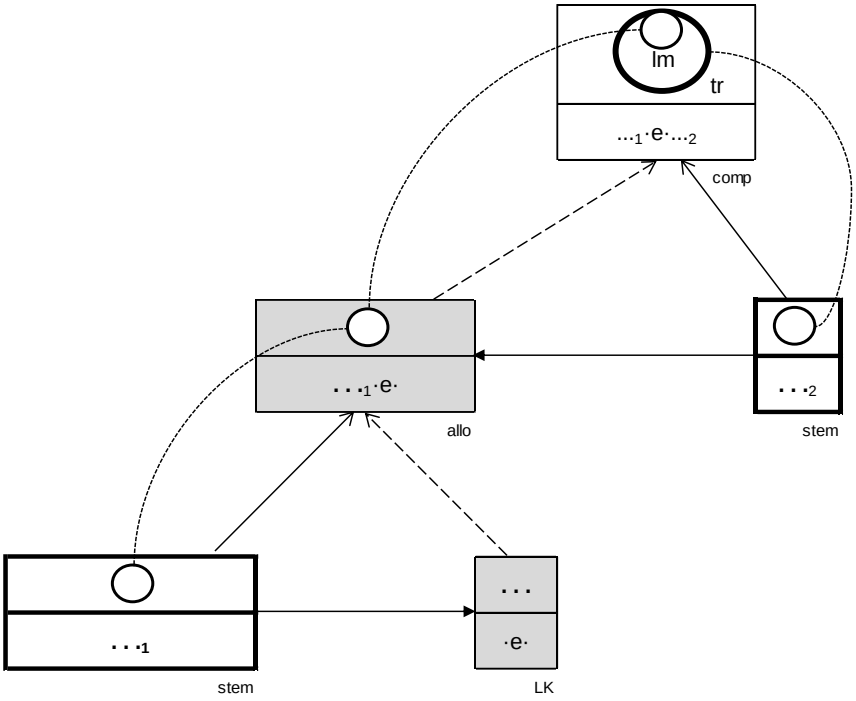


Figure 3: Schema of [THING·LK·THING]/...$_1$·e·...$_2$]

The ease with which a schema is activated by users, depends on the level of *entrenchment* that the specific schema enjoys (Van Huyssteen 2005: 134). The more entrenched a constructional schema is, the more easily it will be activated by users in the process of making new constructions. (In §3.2 entrenchment will be discussed and as components thereof, *generality* and *productivity*.) In the current example, Figure 3 functions as the most productive of the three schemas seeing as it accommodates a massive range of possibilities regarding the first and third component structures. It is unlikely that users will use the schema in Figure 3 when new complex structures with *eend* as the first component must be constructed, as it is formulated very generally/abstractly. Even though Figure 2 is less generally formulated than Figure 3, it is entrenched by/familiar to users when forming complex structures that use *eend* as its first component.

The schemas are themselves generalisations of language use, but they also share characteristics with each other. Related schemas form a categorisation network: a collection of two or more schemas that share certain aspects. The composition structures in Figures 1, 2, and 3 are conflated into a simplistic categorisation network in Figure 4. The solid arrows between schemas represent *elaborations*, as opposed to dashed arrows indicating *extensions* (Langacker 2013: 17–18). Each node on the left-hand side of the categorisation is an elaboration of the one above it, without deviating from it. An extension is a schema that shows commonalities, but also differences, such as the right-hand node in Figure 4. Because an eagle is a bird and the word *eagle* systematically occurs with a linking morpheme when used in complex constructions as first component, it still has something in common with the rest of Figure 4. The use of the ·*s*· linking morpheme is grounds enough to see it as an extension rather than an elaboration. This article aims to construct such a categorisation network for the linking morpheme in Afrikaans that would serve as a synchronic description of this particular word element. Before an extensive categorisation network can be postulated,

it would be prudent to first discuss the processes of *composition* and *entrenchment* with regard to linking morphemes.

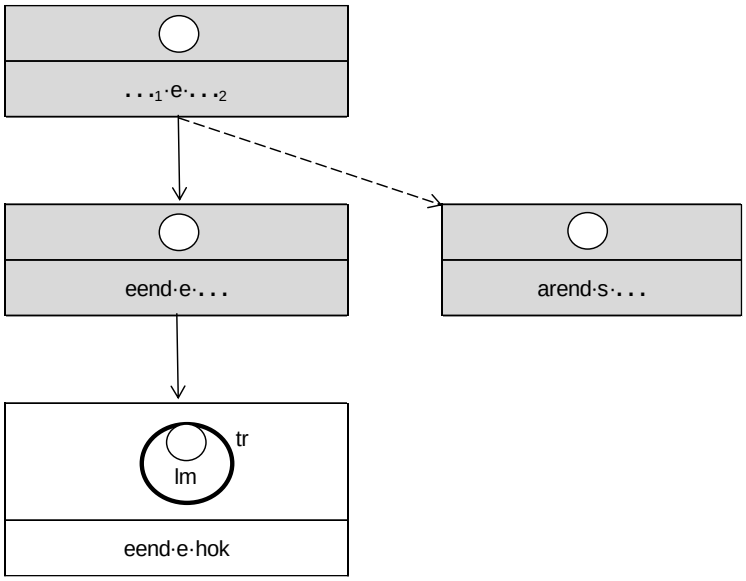


Figure 4: Categorisation network of composition structures with [EEND] 'duck', with [arend·s·...] 'eagle's THING...' as extension

3. Characterisation of the linking morpheme

When describing a linguistic element from a CG perspective, the following cognitive processes need to be taken into account (Van Huyssteen 2005: 128): *symbolisation* (the construction of form-meaning pairs), *composition* (how composition structures are constructed), *categorisation* (where a linguistic element fits into the network of linguistic elements), *schematisation* (the abstraction of commonalities in language use), and *entrenchment* (how easily a constructional schema is activated to construct new, similar constructions). The linking morpheme is, with regard to symbolisation, a form-function pair where the form is specified, but the meaning is highly schematic. The description of the systematic behaviour of the linking morpheme can be seen as a description of the schematisation thereof – the main aim of this article. In this article, composition and entrenchment will be discussed in detail seeing that these processes are at the core of the characterisation of the linking morpheme in Afrikaans.

3.1 *Composition*

Composition entails the combination of component structures to construct more complex composition structures. The manner of combination and the relationship between the component structures are central to this cognitive process. Composition also has to do with the commonalities between the component structures' substructures, the layout of the compositional route that the component structures follow to ultimately form a composition structure, and the distinction between dependent and independent structures (Butler 2014: 55;

Van Huyssteen 2005: 128–131). Constructional profiling, semantic and phonological dependency, and constituency are concepts that form part of composition.

3.1.1 *Constructional profiling*

Constructional profiling entails drawing attention to a specific substructure within a conceptual base structure (Langacker 2013: 66). In terms of complex constructions, it refers to such a structure's *profile determinant* (Van Huyssteen 2005: 129). Compare (5a) versus (5b), for example: both of these structures invoke the image of an eagle (base structure), but different aspects (substructures) of the eagle are being brought to the forefront, namely the claw and the eye respectively. The same goes for other complex constructions like (6) where the emphasis is placed on the person swimming by the use of the nominalising affix ([NR/·er]) that functions as a profile determinant. The profile determinant is also called the *trajector* and the base component structure the *landmark*. [AREND/arend] in (5a), as the base structure of the complex, functions as the landmark, and [KLOU/klou] as the core and specific substructure, functions as the trajector. The trajector and landmark are distinguished from each other in the schemas by the use of 'tr' for trajector and 'lm' for landmark; the trajector is also encircled with a darker line than the landmark; the landmark is connected to the component structured with a dotted arrow; the trajector is connected to the component structure with a solid arrow. Figure 5 serves as an example of these sketch conventions.

(5) a. [EAGLE·LK·CLAW/arend·s·klou]
'eagle's claw'
b. [EAGLE·LK·EYE/arend·s·oog]
'eagle's eye'

(6) [SWIMM·NR/swemme·er]
'swimmer'

In this article the constructional profiling of compounds is characterised using the semantic relationships described in Verhoeven et al. (2014: 28–50). Figure 5 is an example of a part-whole relationship (Verhoeven et al. 2014: 38). In (5a), [CLAW/klou] is a part/component of the whole [EAGLE/arend] and it is sketched as a smaller circle inside a bigger circle. The dotted arches serve as correspondence lines that show [CLAW/klou] (small circle) is the part of the [EAGLE/arend] (big circle) that is being focused on. A more detailed discussion of the semantic categories will be given in §4.2.1.

Especially important when dealing with constructional profiling is to determine which component structure adds to the content of the composition structure on a conceptual level. Complex constructions must be seen as unique constructions, not only a sum total of the component structures that are combined to construct them (Van Huyssteen 2005: 128). The opposite should also be true, namely that any additional component structures of which the objective/separate conceptual content is restricted/seems meaningless, does *not necessarily* have no effect on the ultimate compositional structure.

Langacker (2013: 187–189) discusses *redundancy* as a way in which extra clues are given about the meaning of an expression – it boils down to the fact that different component structures in a composition structure carries the same or very similar information. There are, for example, in both (5a) and (7) instances of part-whole relationships between the

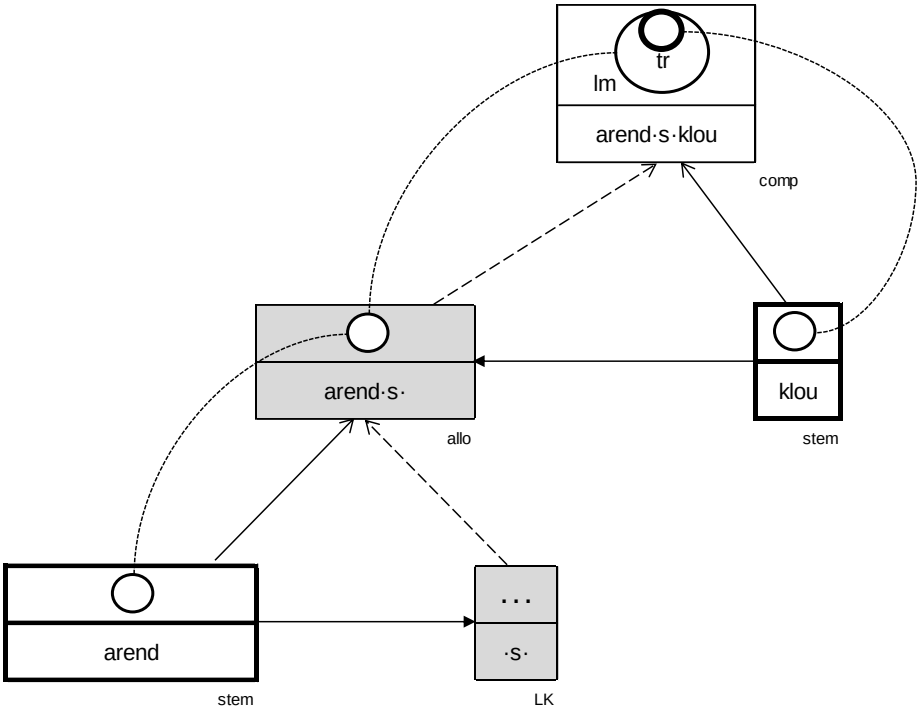


Figure 5: Schema of [AREND·LK·KLOU/arend·s·klou] ‘eagle’s claw’

component structures despite the presence of a linking morpheme in the former and the absence thereof in the latter. The semantic structures of the compounds are similar, but the linking morpheme in the first example is a historical (albeit redundant) clue as to the semantic relationship between the constituents – see Van Tiel *et al*. (2011) for a discussion on the interaction between the loss of the genitive and the use of the *·s·* linking morpheme in Dutch. If seen in this light the linking morpheme does (indirectly) add to the content of the composition structure on a conceptual level.

(7) [OWL·CLAW/uil·klou]
‘owl’s claw’

3.1.2 *Semantic and phonological dependency*
Linking morphemes are highly dependent on both the level of conceptualisation and realisation (Van Huyssteen 2010: 12). They are dependent to such a degree that two independent component structures or one independent and one dependent component structure must combine with the linking morpheme to make it possible to function. In this regard linking morphemes differ from other affixes which combine with only one independent component structure (Langacker 2013: 199–202). Figure 6 and 7 show the schemas of two comparable suffixed forms, (8a) and (8b). In (8a) the independent component structure [WORK/werk] and the dependent component structure [NR/·er] combine with each other; in (8b) two similar components combine, but the linking morpheme is an extra dependent component structure that first combines with the suffix before it combines with the independent component structure (Combrink 1990: 172). This illustrates the general dependency of the linking morpheme and specifically how the linking morpheme first combines with another component structure (in

the case of non-compounds the other component is an affix) before it can be part of the final compositional structure.

(8) a. [WORK·NR/werk·er]
‘worker’
b. [LEARN·LK·NR/leer·d·er]
‘pupil’

The semantic and phonological dependency of the linking morpheme is crucial when formulating a description thereof. The stem allomorph ([EAGLE·LK/arend·s]) in Figure 5 and suffix allomorph ([·LK·NR/·d·er]) in Figure 7 serve as illustrations. In Figure 5 the linking morpheme initially combines with [EAGLE/arend], an independent word/component structure, and it leads to a semantically and phonologically dependent component structure (i.e. a stem allomorph). An additional component structure must combine with this stem allomorph to make it sufficiently independent to function as a complex word. In the case of Figure 7, where [NR/·er] combines with the *·d·* linking morpheme, the effect of the linking morpheme differs slightly. The difference is that the linking morpheme combines with a component structure that was itself dependent and makes it phonologically more complex. The content of the affix is not affected semantically because of the lack of a second independent component structure that would have served as elaboration.

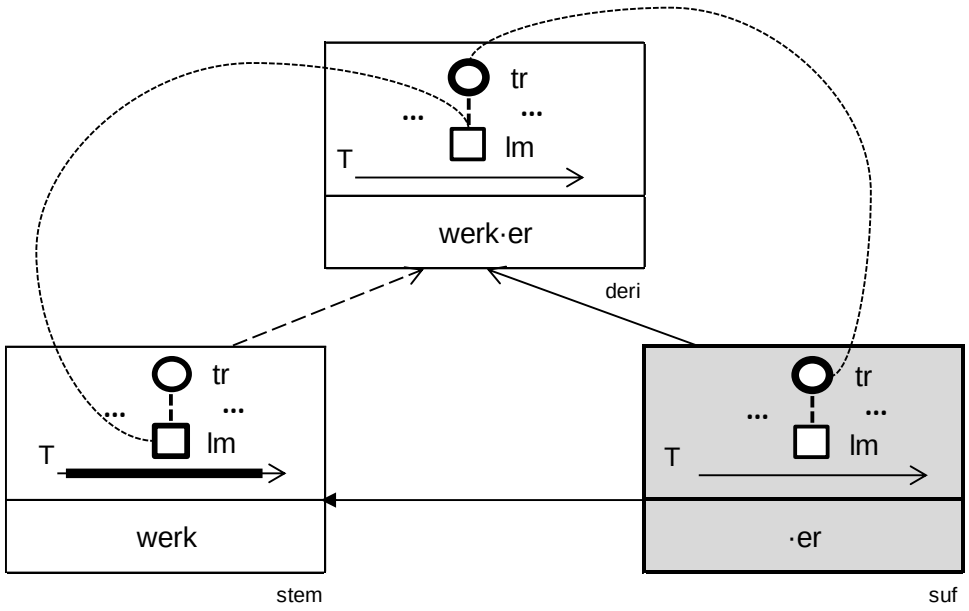


Figure 6: Schema of [WORK·NR/werk·er] ‘worker’

With regard to phonological dependency, a simplified manner in which the realisation pole can be represented, is sketched in Figures 8 and 9. The ‘T’ label (and accompanying arrow) represents a temporal order in which the component structures are realised by a language user (Langacker 2013: 163). The other arrows, like those in Figure 5, point in the direction of the

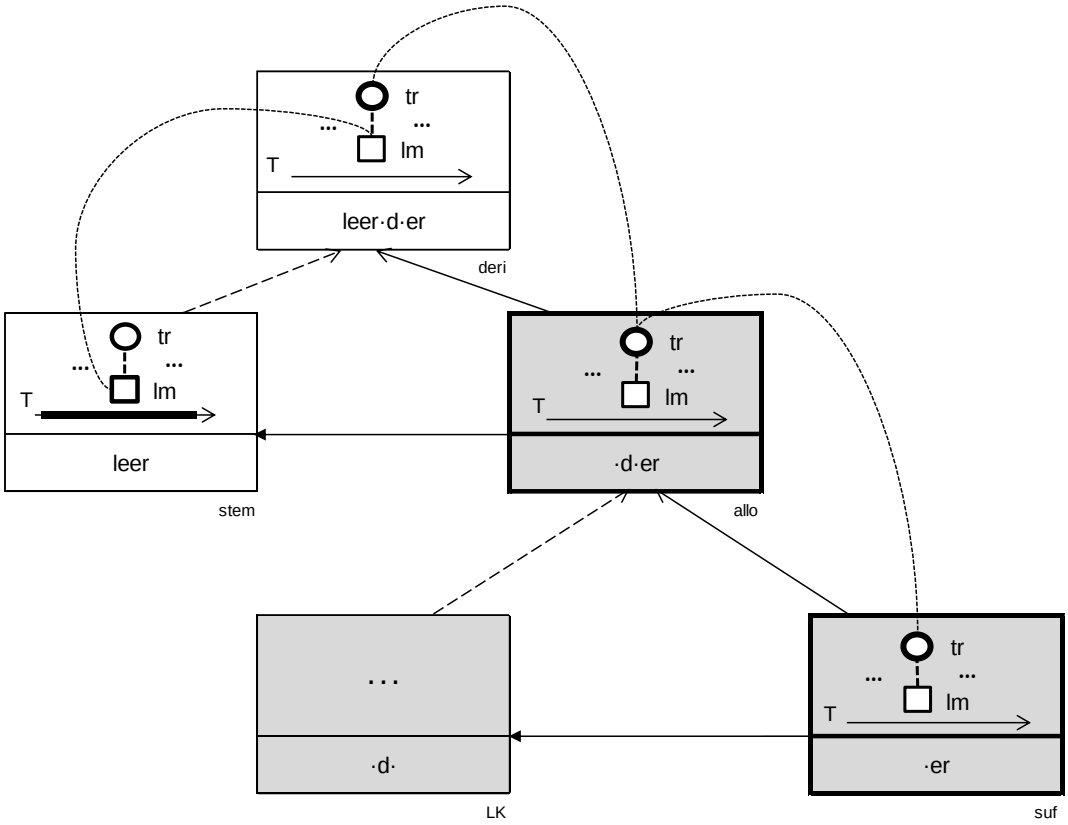


Figure 7: Schema of [LEARN·LK·NR/leer·d·er] ‘pupil’

components that elaborate the profile determinant. The block around [EAGLE/arend] and [LK/·s·] is used to indicate the phonological unit formed by the linking morpheme and the morpheme that it is preceded by (Kempen 1969: 94).

In the case of affixed constructions like [LEARN·LK·NR/leer·d·er], as depicted in Figure 9, the phonological composition is in the opposite direction, since the linking morpheme causes an otherwise cohering morpheme to be non-cohering (Booij 2018). Such a suffix allomorph of the suffix [NR/·er] is sketched in Figure 9. The dotted block around the left-hand constituent and the linking morpheme represents the syllabic relationship between the two components. Labrune (2014) argues that allomorphy and the use of linking morphemes are different processes when it comes to the study of compounds and that they should be studied separately; hence, in this article, they are represented by two separate categorisation networks (Figures 17 and 18).

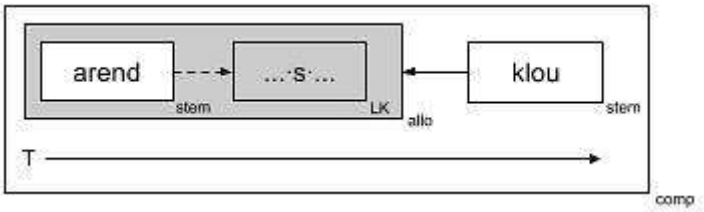


Figure 8: Realisation schema of [EAGLE·LK·CLAW/arend·s·klou] ‘eagle’s claw’

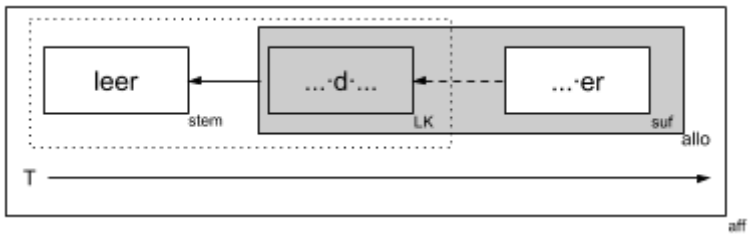


Figure 9: Realisation schema of [LEARN·LK·NR/leer·d·er] ‘pupil’

3.1.3 *Constituency*

Constituency involves the way and order in which composition structures’ components combine. According to Tuggy (2005: 257–258) constituency is mainly about how component structures combine, rather than the order in which they combine. Tree structures/hierarchies can be used to depict different levels of constituency (Langacker 2013: 205–207). In contrast to generative grammar that sees tree structures as ‘autonomous grammatical primitives’ (Langacker 2003: 55–57), CG uses it exclusively as a heuristic aid to help explain the symbolic connections between concepts.

As we have illustrated thus far, linking morphemes in Afrikaans combine with either an independent component structure (like [EAGLE·arend] in [EAGLE·LK·CLAW/arend·s·klou]), or a dependent component structure (like [NR/·er] in [LEARN·LK·NR/leer·d·er]). However, there is considerable uncertainty pertaining to the manner of combination of these two cases.

The *binary branching hypothesis* (BBH) is seen as the solution for the constituency question. The hypothesis was developed in generative grammar to restrict the complexity of grammar and the underlying relationships between components and to keep grammar as unambiguous as possible (Guevara 2007: 1–2). Guevara (2007) compares the BBH with the alternative *simpler syntax hypothesis* (SSH). The SSH, suggested by Culicover and Jackendoff (2005) as a simpler way of analysis, seems to be just as vague and unverified as the BBH. The adoption of one of the hypotheses necessarily excludes the other (Geuvara 2007:3–4, 7). Langacker (2003: 58) states that both methods of constituency are valid, and that the crux of the matter is to clearly show the relationship between the constituents. It is nevertheless clear that more research should be done on this subject. For the purposes of this article the BBH is assumed in accordance with the current tradition in Afrikaans morphology (Combrink 1990; Kempen 1969), where both left and right branching are accepted.

3.2 *Entrenchment*

Entrenchment pertains to the ease with which a certain constructional schema is activated, whether as an entrenched constructional schema with unit status, or for the formation of new constructional schemas. Generality and productivity are the two components of entrenchment that warrant discussion.

3.2.1 *Generality*

Generality can have two meanings: firstly, with regard to the nature of a constructional schema; secondly, with regard to the distribution thereof. In terms of the nature of a constructional schema, (4c) (repeated here for convenience of reference) is more general than (4b), which is more general than (4a), since each successive constructional schema accommodates more possibilities.

(4) a. [DUCK·LK·CAGE/eend·e·hok]
‘duck’s cage’
b. [DUCK·LK·THING/eend·e·...]
‘duck·LK·THING’ (any compound with left-hand constituent *eend*, followed by a linking morpheme and another noun, e.g. *eend·e·plaas* ‘duck farm’, *eend·e·boerdery* ‘duck farming’, and *eend·e·dam* ‘dam for ducks’.)
c. [THING·LK·THING/...$_1$·e·...$_2$]
‘THING·LK·THING’ (any noun-noun compound with a linking morpheme)

However, in terms of the distribution of the constructional schema, (9a) will be more general than (4c) because the latter schema is applicable to a more restricted part of Afrikaans complex words. The schema in (9a) is the default schema for the formation of noun-noun compounds in Afrikaans, whereas (4c) applies to a smaller section of noun-noun compounds i.e. those that also take linking morphemes; hence, (9a) is more general than (4c). Van Huyssteen (2005: 133) points to the fact that generality is not a precondition for a constructional schema – constructional schemas that are not very general could still be productive.[3] In this article, the use of generality will be used specifically in this sense.

(9) a. [THING·THING/x·y]
‘THING·THING’

3.2.2 *Productivity*
Productivity becomes relevant when one asks to what extent a constructional schema is available for fashioning new constructions. (10a) is an example of a constructional schema that is highly productive, even though not general in the first sense of the meaning of ‘general’. Complex constructions like (10b) and (10c) are formed on the basis of this productive constructional schema. With (10b) and (10c) as analogues examples, one can similarly construe (10d) as a novel compound. Generality and productivity are characteristics that aid in the expansion of schemas, and ultimately add to the expansion of categorisation networks. A constructional schema like that of the linking morpheme, is similarly not as general, but still productive, as will be illustrated subsequently when we postulate categorisation networks for the linking morpheme.

(10) a. [PROCESS·NR·LK·THING/x·er·s·y]
b. [WORK·NR·LK·UNION/werk·er·s·bond]
‘worker’s union’
c. [NOMAD·NR·LK·LIFE/swerw·er·s·lewe]
‘nomad’s life’
d. (SKATE·NR·LK·TOURNAMENT/skaats·er·s·toernooi)
‘skater’s tournament’

[3] This hypothesis should be investigated further specifically with reference to the discussion on rule-based and analogy-based morphology.

4. Categorisation networks for the linking morpheme

To construct a categorisation network for the linking morpheme in Afrikaans, constructional schemas on the realisation as well as conceptualisation pole are required. These schemas should be based preferably on usage-based data, since natural, real-world data forms an integral part of CG research (Van Huyssteen 2005: 135). For this purpose, a variety of data sources were used for a usage-based representation of the linking morpheme in Afrikaans.

For compounds, we use the *AuCoPro*[4] data set, comprising of 25,266 split Afrikaans compounds and 3,828 semantically annotated compounds. Only annotated compounds that contained linking morphemes were considered, for a total of 288 compounds. For affixed words, the *NCHLT*[5] corpus was used, comprising a total of 64,257 tokens; all affixed forms containing the 'MLG' tag (*Morpheme>Linking>Germanic*) were extracted, for a total of 72 words as our non-compound data set. The quantitative summary of the data sets can be seen in Table 1.

Because of the relatively small size of the data sets, we supplemented our data with three other sources, namely the data-based studies of Kempen (1969) and Combrink (1990), and the rules for the use of the linking morpheme (specifically for the *·s·* and hyphen) in Afrikaans prescribed in the official Afrikaans orthography (Taalkommissie 2017).

Table 1: Summary of the data sets

Data set	Number of complex words
Compound data set	288
Non-compound data set	72
	360

4.1 *Description: Realisation pole*

4.1.1 *Linking morphemes in compounds*

The distribution of linking morphemes in this study is summarised in Table 2. In the data set consisting of compounds, the *·s·* linking morpheme and the hyphen are clearly the linking morphemes that are used the most (with a combined total of 268 (93%) out of the possible 288 compounds). This confirms Combrink's (1990: 272) assertion that the *·s·* is the most frequently used linking morpheme in Afrikaans.

[4] For more information on the project, see www.sourceforge.net/projects/aucopro/ and www.gerhard.pro/aucopro/.

[5] For more information on the project, see rma.nwu.ac.za.

Based on the data in the compound data set, it is possible to construe a categorisation network of the linking morpheme in Afrikaans compounds. The darker the outline of a node, the more compounds appeared with the appropriate linking morpheme and the more prototypical they are, as illustrated for the *·s·* linking morpheme and the hyphen. Figure 10 provides the complete realisation pole of the linking morpheme in Afrikaans compounds that featured in the sources used.

Table 2: Quantitative analysis of the linking morpheme distribution in the compound data set

LK	Example	Count
·s·	*mag·s·balans* (‘power balance’)	164
·-·	*rune·-·inskripsie* (‘runic inscription’)	104
·e·	*neut·e·dop* (‘nut shell’)	14
·ns·	*lewe·ns·probleme* (‘life problems’)	4
·er·	*kind·er·naam* (‘child’s name’)	1
·n·	*lewe·n·styl* (‘life style’)	1
		288

The prominence of *·s·* linking morpheme and the hyphen is confirmed by the fact that specific spelling rules[6] exist for them in the official Afrikaans orthography (Taalkommissie 2017), but not for any other linking morphemes. The effect of these rules is that they afford these two linking morphemes a prime position when new complex words are formed in the language. Specifically, the hyphen as an orthographical morpheme plays an important role with regard to readability and semantics (Taalkommissie 2017: 77), which makes it compulsory in certain contexts; occasionally without consideration of semantic factors, and occasionally without consideration of readability.

The *·er·* linking morpheme, *·n·* linking morpheme, and *·ns·* linking morpheme (occurring together in only six compounds in the data set), seem highly exceptional. The *·n·* linking morpheme is related to the Dutch infinitive/gerund form of a verb. The *·ns·* linking morpheme is seen as a combination of the infinitive/gerund form and the Afrikaans *·s·* linking morpheme; from the data (see Table 3), it seems that *·ns·* combines exclusively with *lewe* ‘life’ as left-hand component in compound, e.g. *lewe·ns·vreugde* ‘life’s joy = joy of life’. The *·er·* linking morpheme (that only features in (11a)) seems to attach to a select few constituents, which include (11b), (11c), and (11d) (Combrink 1990: 249).

[6] It is accepted a priori that the orthographical rules of a language and its morphology influences one another – see Schäfer & Pankratz (2018:333) where reference is made to the influence of Dutch spelling reforms on language users’ interpretation of compounds.

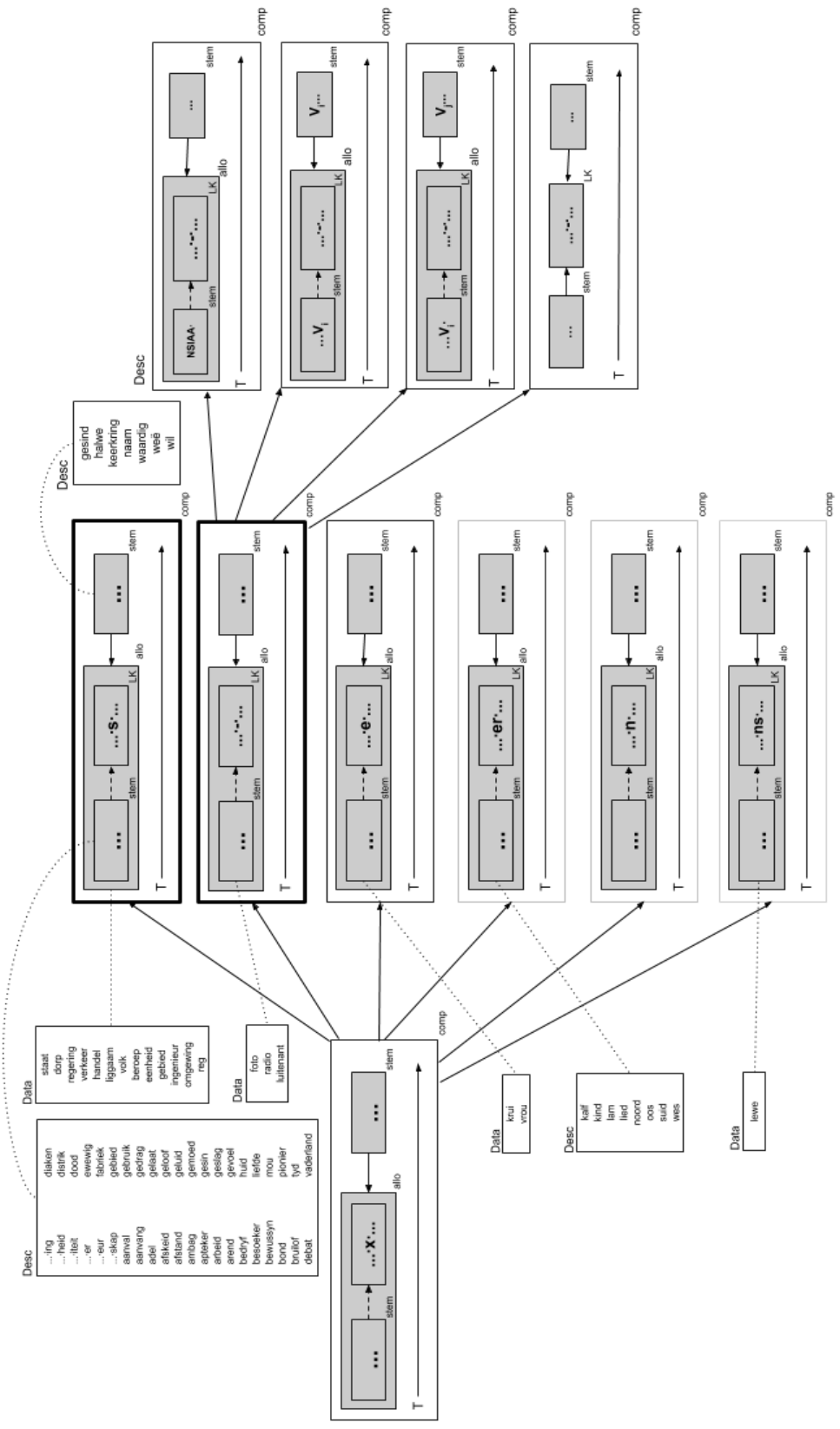


Figure 10: Categorisation network for the realisation pole for the linking morpheme in Afrikaans compounds (Full-scale version available at http://gerhard.pro/publications.)

(11) a. [CHILD·LK·NAME/kinder·er·naam]
‘child’s name’
b. [CHILD/kind]
‘child’
c. [CALF/kalf]
‘calf’
d. [SOUTH/suid]
‘south’

When postulating a categorisation network, one must include both lower-level (less general) schemas, and more general schemas. The left-hand constituents that appeared three times or more in the compound data set (three times is sufficient in the light of the size of the data set) are listed in Table 3.

Table 3: Left-hand constituents that appeared three or more times in the compound data sets (with the specific linking morpheme they appear with)

Constituent	LK	Count	Constituent	LK	Count
staat (‘state’)	*·s·*	8	*beroep* (‘occupation’)	*·s·*	3
dorp (‘town’)	*·s·*	5	*eenheid* (‘unit’)	*·s·*	3
foto (‘photo’)	·-·	5	*gebied* (‘area’)	*·s·*	3
regering (‘government’)	*·s·*	5	*ingenieur* (‘engineer’)	*·s·*	3
verkeer (‘traffic’)	*·s·*	5	*krui* (‘herb’)	*·e·*	3
handel (‘commerce’)	*·s·*	4	*luitenant* (‘lieutenant’)	·-·	3
lewe (‘life’)	*·ns·*	4	*omgewing* (‘environment’)	*·s·*	3
liggaam (‘body’)	*·s·*	4	*reg* (‘law’)	*·s·*	3
radio (‘radio’)	·-·	4	*vrou* (‘woman’)	*·e·*	3

Twelve of the eighteen left-hand constituents in the table combined with the *·s·* linking morpheme, three with a hyphen, two with the *·e·* linking morpheme, and one with *·ns·*. With a

frequency of eight, it is clear that [STATE/staat] calls for the use of the *·s·* linking morpheme when used in complex words; Combrink (1990) refers to such constituents as 'linking morpheme attracting stems', i.e. constituents that usually require a linking morpheme when functioning as left-hand constituent of a compound. The same applies to [LIFE/lewe], which clearly requires the *·ns·* linking morpheme. The three component structures that seem to demand hyphens when they function as the left-hand components in compositions structures will be discussed later in this section.

Instead of making individual constructional schemas for each complex word, the constituents that feature the most in the data set are included as a node, labelled 'Data', in Figure 10. Every constituent is connected to the relevant linking morpheme constructional schema by a dotted line. The blocks marked with the label 'Desc' in Figure 10 (and all subsequent figures) contain specific component structures that co-occur with the specific linking morpheme, as identified in linguistic descriptions (hence 'Desc'; specifically, Kempen 1969: 103–105; Combrink 1990: 247–250; Taalkommissie 2017). Included in the 'Desc' category is (on the left-hand side of Figure 10) complex words ending in affixes that include [NR/·er] and [NR/·skap] that attract the use of the *·s·* linking morpheme when they are used as the left-hand constituent, and the right-hand affixoids like *·gesind* and *·halwe* which behave identically when used as the right-hand constituent. The use of the *·s·* linking morpheme in these cases is governed by rule 19.14 and 19.15 in Taalkommissie (2017).

Neither Kempen (1969) nor Combrink (1990) could be used for the constructional subschema related to the hyphen, since neither of them consider hyphens as linking morphemes. Therefore, the rules postulated by Taalkommissie (2017) form the basis of this subschema. Three of these rules are relevant to compounds and can be summarised as follows:

(12) a. When a number, symbol, initialism, abbreviation or acronym is used as a component of a compound, then it must be separated from the rest of the compound with a hyphen, e.g. *BTW-heffing* 'VAT levy' (rule 7.1, 12.8 and 12.9).

b. In a compound where a cluster of identical or different vowels is present, the word can be interpreted incorrectly, and the use of a hyphen is then compulsory, e.g. *foto-omslag* 'photo cover', and *radio-ingenieur* 'radio engineer' (rule 12.1).

c. If there is a specific semantic relationship between the components in the compound (which includes a coordinative relationship), a hyphen must be placed between the components, e.g. *konkaaf-konveks* 'concavo-convex' (rule 12.7).

These orthographic rules are also represented in Figure 10. The label 'NSIAA' is used to refer collectively to a number, symbol, initialism, abbreviation and acronym (see example in (12a)). The 'V' label stands for a vowel letter that is the last letter of the first component and the first letter of the second component. The vowel letters in the second block under the 'NSIAA' label in Figure 10 are identical (both indicated by V_i), while the vowel letters in the third block are not (represented by V_i and V_j respectively; see examples in (12b)). The rule related to a coordinative relationship between the components is illustrated in the bottom rectangle (see example in (12c)). In Figure 10 the constructional schemas for the hyphen rules are present in combination with the schemas present in the data, even though examples of all the hyphen-related rules are not present in the data. The *·s·* linking morpheme and the hyphen are shown

to be the most productive linking morphemes in Afrikaans. In contrast herewith, it seems that the *·ns·* linking morpheme only appears with (15d), maybe even exclusively.

4.1.2 *Linking morphemes in non-compounds*

It must be reiterated that linking morphemes that appear in affixed forms attach to the affix, as was mentioned in §3.1.2 and discussed in Combrink (1990: 272). The use of linking morphemes in affixation has been studied by making use of the non-compound data set, quantitatively summarised in Table 4.

Table 4: Quantitative analysis of the linking morpheme in the non-compound data set

LK	Example	Count
·d·	*uitvoer·d·er* (export·LK·NMLZ; 'exporter')	23
·e·	*feit·e·lik* (fact·LK·ADJZ; 'actual(ly); factual(ly)')	16
·n·	*rede·n·eer* (reason·LK·VBLZ; 'argue')	12
·en·	*wes·en·lik* (be·LK·ADJZ; 'essential')	13
·-·	*nie·-·ingesetene* (not·LK·resident; 'non-resident')	4
·er·	*kind·er·s* (child·LK·PL; 'children')	4
		72

The *·d·* linking morpheme features most in affixed forms, while the schwa (and variations thereof like *·en·* and *·er·*) occur second most. The distribution of the *·e·* linking morpheme, *·en·* linking morpheme, and *·n·* linking morpheme are quite equal, whereas the hyphen and *·er·* linking morpheme are the least likely to occur in affixed forms. The *·er·* linking morpheme only appears when *kind* 'child' is pluralised, as shown in (13a). Other examples of non-compounds that make use of the *·er·* linking morpheme include (13b) and (13c), as identified by Combrink (1990: 249).

(13) a. [CHILD·LK·PL/kind·er·s]
‘children’
b. [EAST·LK·NR/Oost·er·ling]
‘Easterner’
c. [CALF·LK·ADJR/kalw·er·agtig]
‘calf-like’

From the available data it is apparent that a word that ends on *·eer*, as well as the *·er·* affix (whether as [CMP/·er] or [NR/·er]) requires the use of the *·d·* linking morpheme. The use of the *·e·* linking morpheme is less dependent on the syllable that precedes it and more on the suffix used, specifically [ADJR/·lik] ‘like’. These phonological regions, similar to the approach taken to compounds in § 4.1.1, are represented in the constructional schemas in Figure 11. The differences concerning sketching conventions of the affixed forms from compounds are limited to the labels: the “suf” label indicates suffixes, the ‘pref’ label indicates prefixes, and the ‘aff’ label indicates any affixed form.

As was the case for compounds, the categorisation network for linking morphemes in non-compounds is supplemented using linguistic descriptions. Combrink (1990: 246–247, 249–250) identifies a single prefix that occur with a linking morpheme, viz. [ADJR/a·] in a case like (17a). Similar examples do not occur in the non-compound data set. According to Combrink (1990: 259) the requirements for the use of this specific *·n·* linking morpheme are that: (i) the component structure that it has to combine with has to start with a vowel or an <h> (e.g. (14a) and (14b)); and (ii) it has to have a Greek/Classical origin. In other words, in order to increase the valency of the prefix, another morpheme (*·n·*) is required to create an allomorph that could combine with another constituent.

(14) a. [NEG·LK·ORGANIC/a·n·organies]
‘inorganic’
b. [NEG·LK·HYDRIDE/a·n·hidried]
‘anhydride’

Combrink (1990: 250) also identifies a number of suffixes that regularly combine with linking morphemes. These are [NR/·ling] as in (15a), [ADJR/·loos] as in (15b), [NR/·nis] as in (15c), [ADJR/·rig] as in (15d), and [NR/·ry] as in (15e).

(15) a. [DROWN·LK·NR/drenk·e·ling]
‘drowning person’
b. [SENSE·LK·ADJR/sinn·e·loos]
‘senseless’
c. [CONFESS·LK·NR/belyd·e·nis]
‘confession’
d. [BLOOD·LK·ADJ/bloed·e·rig]
‘bloody’
e. [WALK·LK·NR/stapp·e·ry]
‘walking/pedestrianism’

From the data it is clear that the hyphen is used as a linking morpheme that combines only with the prefix [NIE/nie] ‘not’. However, despite the fact that similar examples are not present in the non-compound data set, provision is made in the categorisation network for hyphens in cases like *sebra-agtig* ‘zebra-like’, as well as for numerous prefixes (e.g. *bi-elektries* ‘bi-electric’) and combining forms (e.g. *gastro-enteritis* ‘gastro-enteritis’) (Taalkommissie, 2017: 27–28).

These linking morphemes are featured in Figure 11 together with the other more prototypical linking morpheme constructional schema, and are connected with correspondence lines to indicate that they are equal constructional schemas.
Seeing that numerous phonological aspects play a significant role when affixation is concerned, the realisation pole of the linking morpheme is more complex than in the case of compounds. A lot more detail was added to the realised structure of non-compounds in Figure 11, improving on previous linguistic descriptions of linking morphemes in non-compounds.

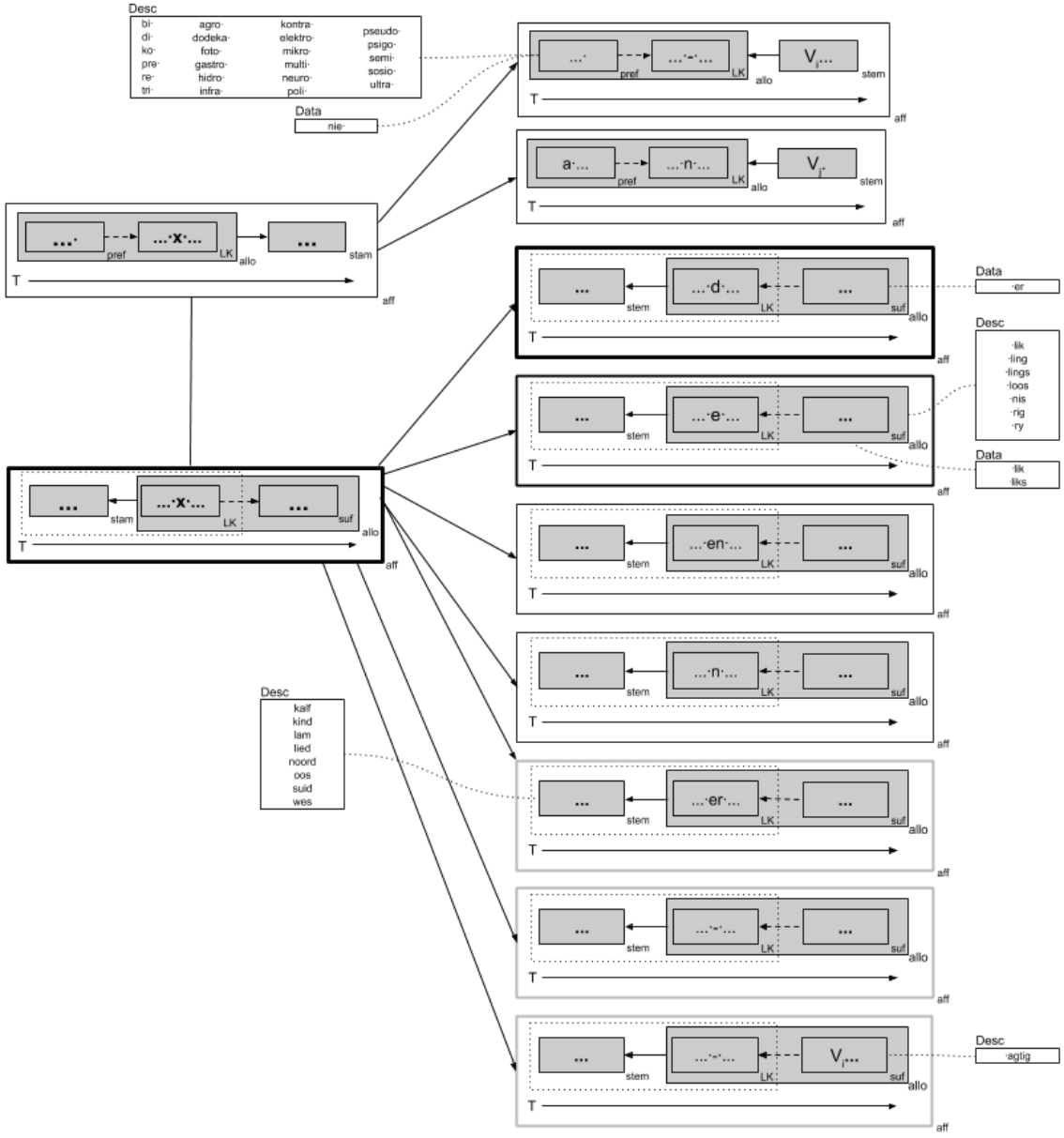


Figure 11: Categorisation network for the realisation pole for the linking morpheme in Afrikaans non-compounds

4.2 *Description: Conceptual pole*

A categorisation network consists of not only a realisation pole, but also a conceptual pole. A discussion of the semantic contexts where linking morphemes tend to appear, and the possibility of the linking morpheme adding any semantic value to the complex morphological construction, is relevant. Since the linking morpheme is prototypically semantically empty, Figure 12 represents the conceptual constructional schema of a prototypical linking morpheme in any Afrikaans complex word – compounds and non-compounds. The shading and ellipse represent the highly schematic semantic content of the linking morpheme. The possibility of elaborating this basic schema with more complex conceptual schemas (especially for compounds), will be discussed in the next sub-section.

Figure 12: Primary conceptual constructional schema of the linking morpheme in Afrikaans

4.2.1 *Linking morphemes in compounds*

The compound data set has been semantically annotated in accordance with the protocol in Verhoeven et al. (2014: 28–50). Six main semantic categories in which noun-noun compounds could be categorised are listed in the protocol. In Table 5 the distribution of the 288 noun-noun compounds in the compound data set is summarised and categorised into these six main categories. Due to the sparseness of the data, conclusions could only be made about the hyphen and the *·s·* linking morpheme.

Table 5: Frequency matrix of the compound data set's semantic categories

	BE	**HAVE**	**IN**	**ACTOR**	**INSTR**	**ABOUT**	**Count**
·s·	4	45	13	9	47	46	164
·-·	49	18	5	4	15	13	104
·e·	3	7	0	3	0	1	14
·ns·	0	1	0	0	0	3	4
·er·	0	1	0	0	0	0	1
·n·	0	0	0	0	0	1	1
Count	56	72	18	16	62	64	288

The first observation from the table is the nearly even distribution of the *·s·* linking morpheme in three of the six categories (HAVE, INSTR, and ABOUT), as illustrated by examples (16a) to (16c). The *·s·* linking morpheme is also represented in the three other categories (ABOUT, BE and IN), which suggests that the *·s·* linking morpheme's conceptual import is varied – if not vague – due to semantic bleaching and over-use. Accompanying (16a)–(16c) are schemas (Figures 13, 14, and 15) representing the conceptual meaning of the categories, with 'tr' labelling the head and "lm" labelling the non-head of the compound.

(16) a. HAVE: [EAGLE·LK·EYE/arend·s·oog] ≈ [EAGLE HAVE EYE]

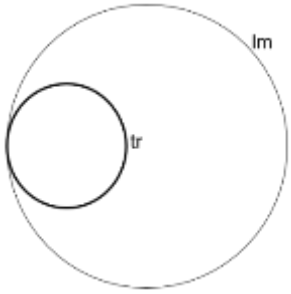


Figure 13: Construal schema of HAVE relationship

In Figure 13 it is shown how the trajectory [EYE/oog] is part of a bigger whole [EAGLE/arend] when a compound is categorised as representing a HAVE relationship.

b. INSTR: [COMBUSTION·LK·PROCESS/ontbranding·s·proses] ≈ [PROCESS IN WHICH COMBUSTION, AS A NON-LIVING ENTITY, IS TAKING PART]

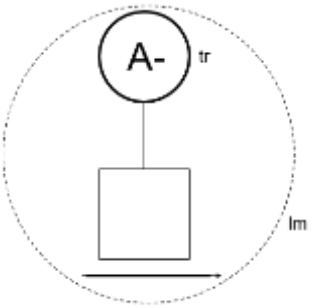


Figure 14: Construal schema of INSTR relationship

The 'A-' label in Figure 14 serves to represent a non-living entity [COMBUSTION/ontbranding] taking part in a process [PROCESS/proses] (the process being represented by the timeline inside the broken-line circle and the square connected to the 'A'- circle).

c. ABOUT: [STABILITY·LK·PROGRAMME/stabiliteit·s·program] ≈ [PROGRAMME ABOUTSTABILITY]

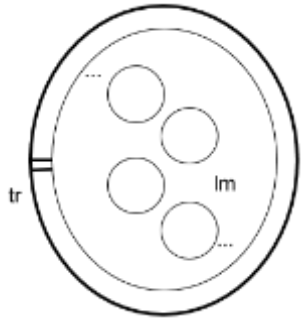


Figure 15: Construal schema of ABOUT relationship

The multiple smaller circles in Figure 15 serve to symbolise the different aspects of the landmark [STABILITY/stabiliteit] that characterises the trajectory [PROGRAMME/program]. In this example 'stability', and that which it entails, forms the base of the relevant 'programme'. However, it is not the trajectory, since the focus is still on the 'programme'.

The hyphen is most often used when there is an IS relationship between the components. For example, [OFFICER·LK·VETERINARIAN/offisier·-·veearts] 'veterinary officer' could be paraphrased as 'a veterinary officer is a veterinarian that IS also an officer'. In this case a conceptual association is apparent – the hyphen serves as an indication of a coordinative relationship between the components (see (12c)), where the two components are both equally central to the meaning of the compound. One could say two trajectors are present, represented by the 'tr' labels in Figure 16. The double line between the circles indicates an equal semantic relationship between the component structures.

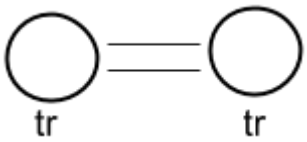


Figure 16: Constructional schema of the coordinative semantic relationship

Of the remaining 55 compounds with hyphens, 52 of the cases use a hyphen to avoid vowel clusters (see (12b)). In as such, these hyphens have a pure orthographical function to prevent misinterpretation of the construction by the reader, without any conceptual import.

4.2.2 *Linking morphemes in affixed forms*

No evidence could be found (in the literature or data) that would indicate that the linking morpheme adds any semantic dimension (regardless the schematicity thereof) to the meaning of complex words like [LAW·LK·ADJR·NR/wett·e·loos·heid] 'lawlessness', or [MANAGE·LK·NR/bestuur·d·er] 'manager'. This aspect coincides strongly with the presence of the *·s·* linking morpheme in compounds after specific derivational affixes and before certain stems (Taalkommissie 2017: 168–169). The implication of this lack of semantic input entails that linking morphemes that combine with affixes, is schematic to the extent that it is semantically totally void. However and importantly, it fulfils an important function in creating allomorphs, in order for such allomorphs to combine with other stems/words (in compounds), or affixes (in non-compounds). The highly schematic constructional schema in Figure 15, which represents the core of the linking morpheme in Afrikaans, is therefore also applicable to affixation.

5. Summary

The primary aim of this article was to construct a comprehensive, theoretically unified description of the linking morpheme in Afrikaans – something that has not been done before. The first step was to collect data for the construal of constructional schemas in the categorisation network. Two annotated data sets, namely a compound set (n=288), and a non-compound set (n=72), were used to postulate these constructional schemas. The data sets were supplemented by existing linguistic descriptions, notably those of Kempen (1969), Combrink (1990), and the Taalkommissie (2017).

In conclusion we can now combine the constructional schemas on the realisation and conceptual poles to form complete categorisation networks; see Figure 17 and Figure 18. Note the added correspondence lines between the schemas on the conceptual and realisation pole. The linking morphemes that appear in compounds are connected to their corresponding conceptual input in Figure 17, as has also been done in Figure 18 for non-compounds. From both categorisation networks one could conclude that the linking morpheme in Afrikaans has extremely schematic conceptual input. Nonetheless, in some cases (like that of the hyphen) we can still postulate conceptual import, as is exemplified by the diagrams in the top left-hand part of Figure 17.

With regard to our secondary aim, i.e. to demonstrate how CG (in combination with CM) can be used as a descriptive framework for morphological constructions, we hope to have shown that various tools from CG could aid enormously in our understanding of morphological constructions. Constructs like composition and entrenchment afford us the opportunity to focus on some of the finer-grained details of the realisation, conceptualisation and distribution of morphological constructions. Categorisation networks not only provide (visually) informative overviews of constructions, but also provides us with the opportunity to adequately account for inheritance relations, deviations from prototypical schemas, etc.

Regarding our last aim, namely to provide a CG perspective on the morpheme status of the linking morpheme, we have argued that it is unproblematic to consider it as a morpheme, albeit one with a highly schematic conceptual import. We therefore concur with Kardela (2014: 25) when he concludes that

> [...] linguistic structure is expected to be cross-cut so that each, even the smallest meaningful linguistic element, becomes a well-structured linguistic unit which is held to involve all "levels" of conceptual organization, including the unit's phonological structure, its morphology, syntax, semantics and pragmatics.

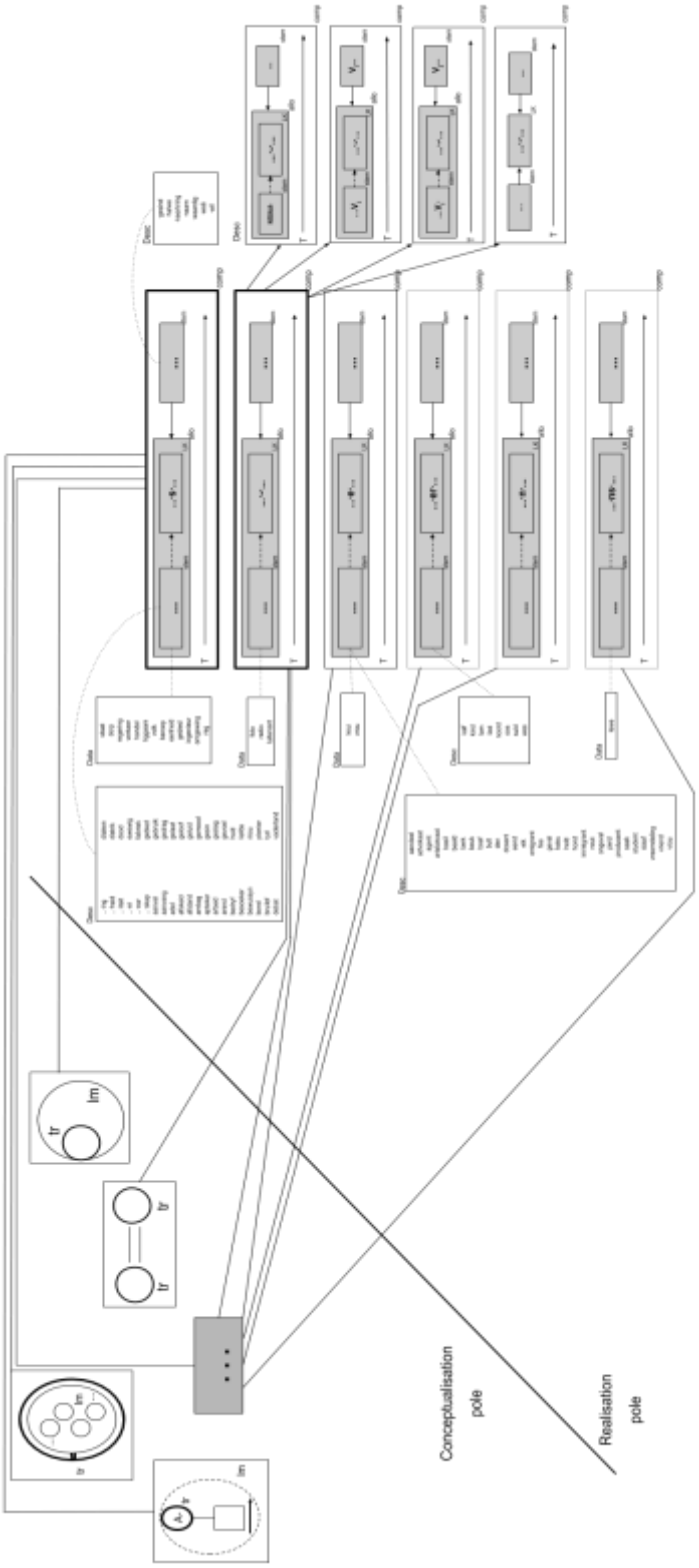


Figure 17: Categorisation network of the linking morpheme in Afrikaans compounds (Full-scale version available at http://gerhard.pro/publications.)

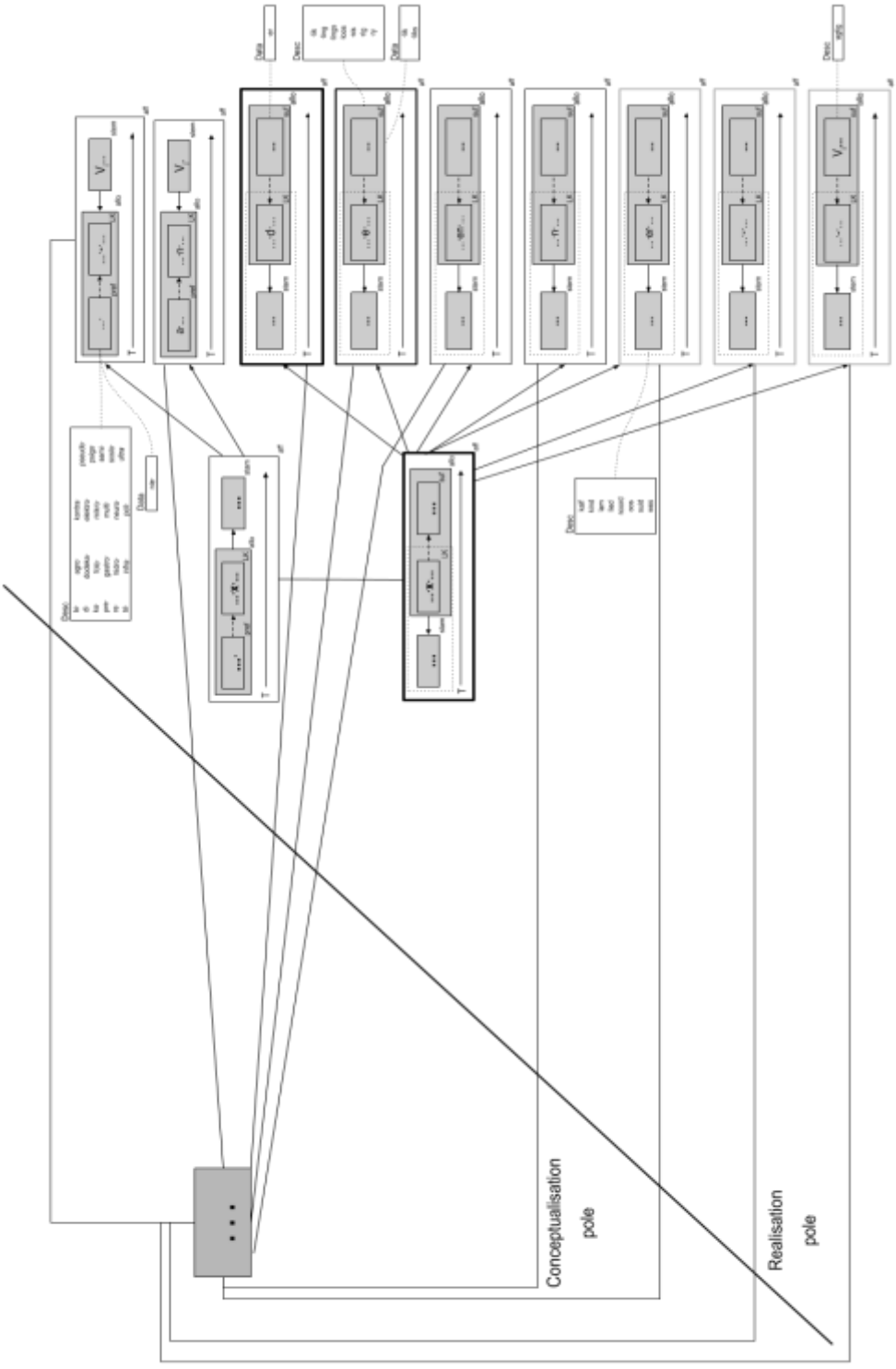


Figure 18: Categorisation network of the linking morpheme in Afrikaans non-compounds (Full-scale version available at http://gerhard.pro/publications.)

Abbreviations

aff	affix(es); affixation
allo	allomorph
AuCoPro	Automatic Compound Processing
BBH	binary branching hypothesis
CG	cognitive grammar
comp	compound(ing)
Desc	Linguistic descriptions, specifically Kempen (1960), Combrink (1990) & Taalkommissie (2017)
LK	linking morpheme
lm	landmark
MLG	morpheme linking Germanic
NCHLT	National Centre for Human Language Technologies
NSIAA	number, symbol, initialism, abbreviation and acronym
POS	part of speech
pref	prefix
SSH	simpler syntax hypothesis
suf	suffix
tr	trajectory

References

Banga, Arina, Hanssen, Esther, Schreuder, Robert & Neijt, Anneke. 2012. How subtle differences in ortography influence conceptual interpretation. Written Language & Literacy 15(2). 185–208.

Banga, A., Hanssen, E., Schreuder, R. & Neijt, A. 2013. Two languages, two sets of interpretations: Language-specific influences of morphological form on Dutch and English speakers' interpretation of compounds. Cognitive Linguistics 24(2).195–220.

Bauer, Laurie. 2003. Introducing linguistic morphology. Washington, D.C.: Georgetown University Press.

Booij, Geert. 2010. Construction Morphology. Oxford: Oxford University Press.

Booij, Geert. 2018. Cohering and non-cohering affixes. Taalportaal. (http://taalportaal.org/taalportaal/topic/pid/topic-13998813296131504) (Accessed 19 April 2018).

Butler, Anna P. 2014. Die deelwoord in Afrikaans: perspektiewe vanuit 'n kognitiewe gebruiksgebaseerde beskrywingsraamwerk. [The participle in Afrikaans: perspectives from a cognitive usage-based descriptive framework]. Potchefstroom: North-West University. (Masters Dissertation).

Combrink, Johan G. H. 1990. Afrikaanse morfologie: Capita Exemplaria [Afrikaans morphology: Capita Exemplaria]. Pretoria: Academica.

Croft, William. 2001. Radical Construction Grammar: Syntactic Theory in Typological Perspective. Oxford: Oxford University Press.

Culicover, Peter W. & Jackendoff, Ray. 2005. Simpler syntax. New York: Oxford University Press.

Evans, Vyvyan & Bergen, Benjamin K. & Zinken, Jörg. 2007. The cognitive linguistic enterprise: An overview. In Evans, V., Bergen, B. K. & Zinken, J. (eds.), The cognitive linguistics reader, 1-36. London: Equinox.

Evans, Vyvyan & Green, Melanie. 2006. Cognitive Linguistics: An Introduction. Edinburgh: Edinburgh University Press.

Fuhrhop, Nanna & Kürschner, Sebastian. 2015. Linking elements in Germanic. In Müller, P. O. & Ohnheiser, I. & Olsen, S. & Rainer, F. (eds.), Word-Formation: An international handbook of the languages of Europe, 568–582. Berlin: De Gruyter.

Goldberg, Adele E. 2006. Constructions at work: The nature of generalization in language. Oxford: Oxford University Press.

Guevara, Emiliano. 2007. Binary branching and linguistic theory: Morphological arguments. In Picchi, M. P. & Pona, A. (eds.), Proceedings of the XXXII Incontro di Grammatica Generativa, 1–18. Allesandria: Edizioni dell'Orso.

Hamawand, Zeki. 2011. Morphology in English: Word Formation in Cognitive Grammar. London: Bloomsbury.

Hanssen, Esther. 2011. Linking elements in compounds: Regional variation in speech production and perception. Nijmegen: Radboud Universiteit Nijmegen. (Ph.D.-thesis).

Janda, Laura. 2011. Metonymy in word-formation. Cognitive Linguistics 22(2). 359–392.

Kardela, Henryk. 2014. The non-modular nature of cognitive grammar. SKASE Journal of Theoretical Linguistics 11. 2–31.

Kempen, Willem. 1969. Samestelling, afleiding en woordsoortelike meerfunksionaliteit in Afrikaans. Cape Town: Nasou Beperk.

Krott, Andrea & Schreuder, Robert & Baayen, R. Harald & Dressler, Wolfgang, U. 2007. Analogical effects on linking elements in German compound words. Language and Cognitive Processes 22(1). 25–57.

Labrune, Laurence. 2014. Featural linking elements in Basque compounds. Morphology 24(4). 377–405.

Langacker, Ronald W. 1987. Foundations of Cognitive Grammar: Theoretical Prerequisites. Vol. 1. Stanford: Stanford University Press.

Langacker, Ronald W. 1990. Concept, Image, and Symbol: The Cognitive Basis of Grammar. Berlin/New York: Mouton de Gruyter.

Langacker, Ronald W. 1991. Foundations of Cognitive Grammar: Descriptive Application. Vol. 2. Stanford: Stanford University Press.

Langacker, Ronald W. 2003. Constructions in cognitive grammar. English Linguistics 20(1). 41–83.

Langacker, Ronald W. 2008. Cognitive Grammar: A Basic Introduction. Oxford: Oxford University Press.

Langacker, Ronald W. 2013. Essentials of cognitive grammar. New York: Oxford University Press.

Manova, Stela. 2011. Understanding Morphological Rules - With Special Emphasis on Conversion and Subtraction in Bulgarian, Russian and Serbo-Croatian. Dordrecht: Springer.

Mattens, Willem, H.J. 1970. De Indifferentialis: een ondersoek naar het anumerieke gebruik van het substantief in het algemeen bruikbaar Nederlands [The Indifferentialis: a study of the non-numerical use of the substantive in generally-used Dutch]. Assen: Van Gorcum.

Neef, Martin. 2015. The status of so-called linking elements in German: Arguments in favor of a non-functional analysis. Word Structure 8(1). 29–52.

Nübling, Damaris. & Szczepaniak, Renata. 2013. Linking elements in German origin, change, functionalization. Morphology 23(1). 67-89.

Schäfer, Roland & Pankratz, Elizabeth. 2018. The plural interpretability of German linking elements. Morphology 28(4). 325-358.

Taalkommissie van die Suid-Afrikaanse Akademie vir Wetenskap en Kuns. 2017. Afrikaanse woordelys en spelreëls [Afrikaans wordlist and spelling rules]. 11th ed. Cape Town: Pharos Woordeboeke.

Taylor, John R. 2002. Cognitive Grammar. Oxford: Oxford University Press.

Taylor, John R. 2015. Word-formation in cognitive grammar. In Müller, Peter O., Ohnheiser, Ingeborg, Olsen, Susan & Rainer, Franz. Volume 1 Word-Formation. An International Handbook of the Languages of Europe. Berlin: De Gruyter Mouton. pp. 145–158.

Tuggy, David. 2003. Abrelatas and scarecrow nouns: Exocentric verb-noun compounds as illustrations of basic principles of cognitive grammar. International Journal of English Studies 3(2). 25–62.

Tuggy, David. 2005. Cognitive approach to word-formation. In Štekauer, P. & Lieber, R. (eds.), Handbook of Word-Formation, 233–265. Dordrecht: Springer.

Van Huyssteen, Gerhard B. 2005. 'n Kognitiewe gebruiksgebaseerde beskrywingsmodel vir die Afrikaanse grammatika [A cognitive-based description model for Afrikaans grammar]. Southern African Linguistics and Applied Language Studies 23(2). 125–137.

Van Huyssteen, Gerhard B. 2010. (Re)defining component structures in morphological constructions: A cognitive grammar perspective. In Michel, S. & Onysko, A. (eds.), Cognitive Approaches to Word-Formation, 97–126. Berlin: Mouton de Gruyter.

Van Huyssteen, Gerhard B. 2017. Morfologie [Morphology]. In Carstens, W. A. M. & Bosman, N., (eds.) Kontomporêre Afrikaanse Taalkunde [Contemporary Afrikaans Linguistics], 177–214. Pretoria: Van Schaik Uitgewers.

Van Huyssteen, Gerhard B. 2018. The *Hulle* and *Goed* constructions in Afrikaans. In Booij, G., (red.) The Construction of Words: Advances in Construction Morphology, 399–437. Cham: Springer.

Van Tiel, Bob & Rem, Margit & Neijt, Anneke. 2011. De historische ontwikkeling van de tussenklank in Nederlandse nominale samenstellingen [The historical development of the linking morpheme in Dutch nominal compounds]. Nederlandse Taalkunde 16(2). 120–140.

Verhoeven, Ben & Van Huyssteen, Gerhard B. & Van Zaanen, Menno & Daelemans, Walter. 2014. Annotation guidelines for compound analysis. Computational Linguistics & Psycholinguistics Technical Report Series. CTRS-005.

Wegener, Heide. 2008. The regrammaticalization of linking elements in German. In Seoane, E. & López-Couso, M. J. (eds.), Theoretical and Empirical Issues in Grammaticalization, 333–355. Amsterdam: John Benjamins Publishing Company.

Eddie Benito Trollip
North-West University, Potchefstroom Campus
School for Languages
Potchefstroom, North-West 2531
South Africa
benmapstieks@gmail.com

Gerhard B van Huyssteen
North-West University, Potchefstroom Campus
CText: Centre for Text Technology
Potchefstroom, North-West 2531
South Africa
gerhard.vanhuyssteen@nwu.ac.za

In SKASE Journal of Theoretical Linguistics [online]. 2018, vol. 15, no. 3 [cit. 2018-13-12]. Available on web page http://www.skase.sk/Volumes/JTL38/pdf_doc/03.pdf. ISSN 1336-782X.

Derivational paradigms and competition in English: a diachronic study on competing causative verbs and their derivatives

Cristina Fernández-Alcaina, University of Granada, Spain
Jan Čermák, Charles University in Prague, Czech Republic

Although there is no clear definition of competition in morphology, it is often described as a situation where two or more forms express the same semantic category (if no restrictions apply). Viewing word-formation as a complex network where elements are interrelated, this paper attempts to describe to which extent a description of derivation in terms of paradigms can help understand morphological competition. This paradigm-competition interaction is expected to be bidirectional, such that the paradigms of two competitors serve as extra evidence for defining competition. At the same time, paradigm theory can benefit from the identification of competing patterns as it may help to specify whether two forms compete for the same semantic niche or not. Based on a sample of 45 Present-Day English verbal clusters where forms in -ize *and zero-derivation compete (or did compete) for the expression of the semantic category* CAUSATIVE*, this paper elaborates on previous research on diachronic competition in two ways. Methodologically, this paper complements the method used in previous research by constructing the subparadigms for the competing verbs while considering both available and unavailable derivatives and using lexicographic and corpus data. At the same time, this approach allows us to offer a more complete description of competition by exploring to what extent the subparadigms of the forms in competition may be used to refine our understanding of the competition and how this can be exploited methodologically.*

Keywords: *derivational (sub)paradigm, availability, competition, -ize, zero-derivation*

1. Introduction

Competition (or *overabundance* in inflection, cf. Thornton 2011) has attracted considerable attention over the past decades as it has proved to be "an inherent and universal feature of natural languages" (Štekauer 2017: 15). With the relatively recent growth of word-based morphology (either hybrid or pure models of word-and-paradigm morphology, see Blevins 2006), there is no doubt that competition should be made part of any morphological account in terms of paradigms (Bonami & Strnadová 2018: 9), both in inflection and derivation. Bonami & Strnadová (2018: 9) suggest viewing "doublets as parallel citizens in a paradigmatic system" and conclude that the problem with doublets is not their representation in a paradigm but the identification of the features that make two or more forms synonymous, and therefore, fillers of the same slot. The definition of synonymy in competing forms remains a challenge that calls for substantial synchronic and diachronic research. This paper elaborates on previous diachronic research on competing clusters[1] by exploring the interaction between derivational paradigms and morphological competition.

Competition is defined as "the coexistence of two or more affixes for the same base and for the expression of the same semantic category, if restrictions (e.g. phonological,

[1] A cluster is defined as "a set of synonymous derivatives morphologically related by their bases but formed with a different affix that can be grouped into doublets, triplets, etc." (Fernández-Alcaina 2017: 168).

morphological) do not apply and no semantic or distributional differences are observed" (Fernández-Alcaina 2017: 166, see also Bauer 2009; Aronoff 2016; Chiba 2016; Fradin 2016).[2] Specifically, this paper relies on a sample of 45 verbal clusters in Present-Day English where forms in *-ize* and zero-derivation compete (or did compete) for the expression of the semantic category CAUSATIVE.

This paper is organized as follows: §2 deals with the interaction between derivational paradigms and competition and with the importance of this interrelation. §3 describes the method used in this paper for data collection and analysis. Results are described in §4, followed by a discussion in §5. Final conclusions are drawn in §6.

2. Derivational paradigms and competition

The *paradigm* has been traditionally viewed as a distinctive feature of canonical inflection that contrasts with the apparently arbitrary organization of derivation. However, increasing evidence against a clear-cut inflection/derivation dichotomy has proved that such distinction is not as straightforward as it was thought to be – or at least, not always (Don 2014: 66–72; Bauer et al. 2015: 533–544). Instead, an account in terms of prototypical categories (Dressler 1989; Plank 1994) and/or subcategories within inflection (Booij 1996) or derivation (Bauer 1997a on evaluative morphology) may offer a more suitable explanation for intermediate cases where the boundaries between inflection and derivation are fuzzy.

Viewing inflection and derivation as the extremes of a continuum implies that the traditional criteria may apply to prototypical instances of inflection and derivation but possibly not to in-between cases. This means that there exists an overlap between inflection and derivation where some of the criteria defined for the former may also apply to the latter – for example, paradigmatic organization. Defective paradigms illustrate the lack of applicability of this criterion in inflection, which together with more or less regular and predictable sequences of derivatives (e.g. *nation–national–nationalize–nationalization* in Bauer 1997b) support a description of derivation in terms of paradigms (van Marle 1985; Bochner 1993; Bauer 1997b; Pounder 2000; Stump 2001: 252–260; Beecher 2004; Booij & Lieber 2004; Booij 2008; Štekauer 2014). However, the definition of the term *paradigm* is still ambiguous and it has been addressed in the literature under various labels depending on the approach: *word family* (Bauer & Nation 1993), *derivational nest* (Horecký et al. 1989 in Štekauer 2014: 364), *derivational family* (Roché 2011; Bonami & Strnadová 2018), *morphological family* (Bauer et al. 2015: 519) or *derivational network* (*Projekt Monika*). Subdivisions to refer to various levels of the derivational paradigm have also been proposed by Beecher (2004) and Bonami & Strnadová (2018). In this paper, derivational paradigms are based on Beecher's (2004: 17) model for a derivational paradigm where the whole set of forms related to a bound root is called *macroparadigm*. The macroparadigm usually consists of several *subparadigms*, where the members contained are related by a common stem. At the same time, these forms can present recursive derivation. Figure 1 illustrates Beecher's model of the macroparadigm for the forms related to the root *popul-*:

[2] This definition refers to competition in complex-word formation. The ambiguity of the term allows different interpretations depending on the approach (see, e.g. Štekauer 2017 for a detailed discussion on competition in complex-word formation and complex-word interpretation).

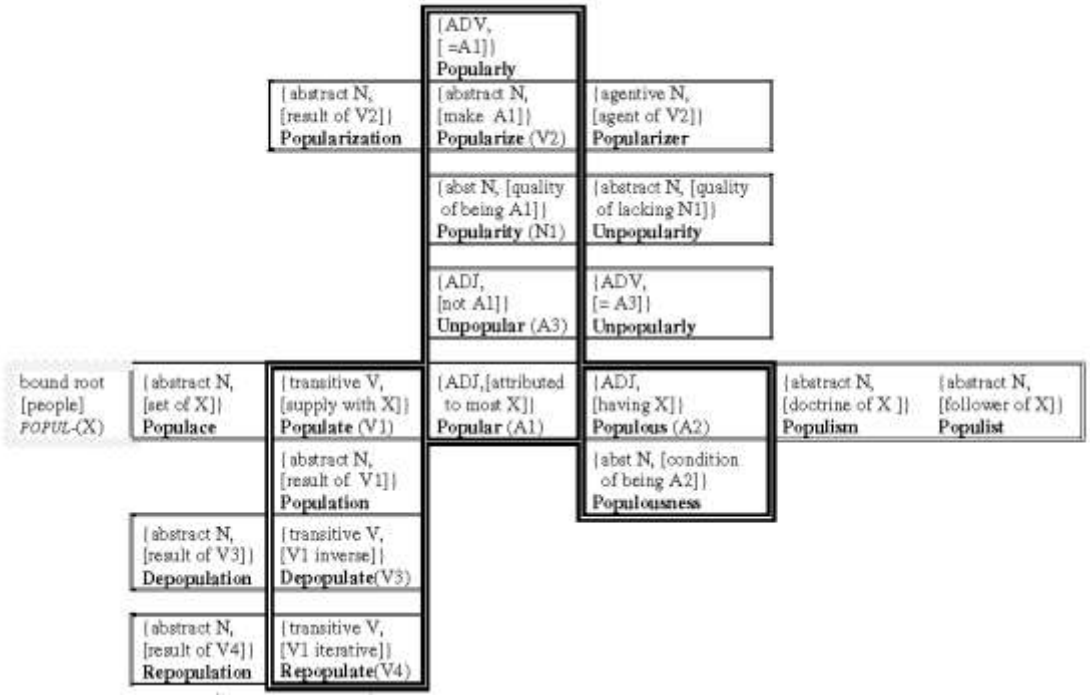


Figure 1: Model for a derivational paradigm proposed by Beecher (2004: 17). Double solid lines represent the subparadigms whose members have a common base (e.g. *populate* (V1) is the base for *population*, *depopulate* and *repopulate*) and which can make room for further derivation (e.g. *depopulation* < *depopulate* < *populate*)

As a way of organizing and systematizing inflectional and derivational data, a description in terms of paradigms must deal with special cases, such as suppletion, gaps in derivational paradigms, defective inflectional paradigms and doublets or overabundance (Bonami & Strnadová 2018), among others. At the same time, the inclusion of competing forms in a paradigm needs an unequivocal definition of competition – desirable but unfortunately still lacking as it has been already mentioned – that would establish the criteria to decide on the synonymy of two (or more) competing forms. However, it seems that, when attempting to delineate the relation between morphological competition and derivational paradigms, the inverse direction is also required.

In that respect, Pounder (2000: 83) addresses the interaction between competition and paradigms in forms with two competing senses (e.g. *kingly* may refer both to 'belonging to a king' and 'like a king') and argues that competition is "at least partly dependent on relations holding between the complete set of lexemes related to the same base". The cases of competition dealt with in this paper are of a slightly different nature to those described in Pounder (2000). For our purpose, two or more forms are considered competitors when they share the same base but differ in the suffix (e.g. in the cluster *mongrel*/*mongrelize* both forms mean 'make mongrel'). In other words, it seems that this base/derivative dependency can be extended to competitors with different suffixes by comparing the relations between the bases and the members of their respective subparadigms.

To the best of our knowledge, previous research usually analyzed clusters of two competing affixes (Kaunisto 2007, 2009; Bauer et al. 2010; Fernández-Alcaina 2017; Lara-

Clares 2017) or groups of three or more competing affixes (Plag 1999; Kjellmer 2001; Díaz-Negrillo 2017), focusing on the forms themselves but disregarding their subparadigms. In this sense, this paper aims at elaborating on the competition between the verbal suffix *-ize* and zero-derivation[3] (Fernández-Alcaina 2017) by exploring to what extent derivational subparadigms provide information on, and possibly influence the result of, the competition of their base forms.

3. Method

3.1 *Data selection*

Previous research (Fernández-Alcaina 2017) relies on an initial sample of 816 verbs in *-ize* expressing the semantic category CAUSATIVE extracted from the *Oxford English Dictionary* (henceforth, OED).[4][5] The sample was filtered by ending (**ize*), keywords contained in the definition ('make' and 'render') and language of origin (English). All the *-ize* verbs were screened for potential competitors in zero-derivation when they express the semantic category CAUSATIVE, yielding a final sample of 45 clusters of verbs in *-ize* and zero-derivation.[6] The main results from this piece of research show that competition displays various profiles of resolution and that, in general, the suffix *-ize* is apparently preferred over zero-derivation for the expression of CAUSATIVE (see Fernández-Alcaina 2017 for further details).

Given the diachronic nature of the present study, data collection was based on two types of evidence and two kinds of sources, lexicographic and corpus. Firstly, derivatives were extracted from the OED, where both obsolete forms and forms in use are recorded. Secondly, available forms with lower frequencies not recorded in the OED were extracted from the *Corpus of Contemporary American English* (henceforth, COCA; Davies 2008–).

As the main aim of this paper is to explore the relation between derivational paradigms and competition, it was necessary to obtain the most complete picture of derivation attainable by filling as many cells in the paradigms as possible. For that reason, this paper follows an inclusive approach in that:

---

[3] Or conversion. For easier reading, the term *zero-derivation* will be used in the rest of the paper.

[4] The cluster *-ize*/zero-derivation expressing CAUSATIVE was selected from a previous sample extracted from the entire frequency list of the *British National Corpus* (henceforth, BNC) (Davies 2004–).

[5] For methodological clarity, it is important to highlight that most of the data analyzed in this paper belong to entries that have been updated in the third edition of the OED (OED3). Forms whose diachronic data are still based on the OED2 will be duly specified in each case.

[6] The cluster *scheme*/*schematize* classified as CAUSATIVE in Fernández-Alcaina (2017) has been disregarded here after being re-interpreted as another semantic category such as MANNER ('represent as a scheme'). In contrast, the cluster *Latin*/*Latinize* ('make Latin'), not included in Fernández-Alcaina (2017), has been included here, again after re-interpretation.

For this reason, the number of clusters with forms in an ongoing or resolved competition as well as the profiles of resolution have been updated in order to be compared with those obtained in this paper. Specifically, the number of patterns classified as instances of ongoing competition in Fernández-Alcaina (2017) has changed from 16 to 15 and those where competition was resolved, from 29 to 30. However, the results after these changes have been made do not differ substantially as the same pattern of competition remains. In fact, the results obtained after the above-mentioned revisions point even more clearly in the direction of the resolution of competition in favor of the *-ize* form (*Latinize*) (see §4).

i) it considers available and unavailable forms in the creation of the subparadigms with bases in *-ize* and zero-derivation, and

ii) apart from affixation, the following analysis also encompasses zero-derivation, neoclassical compounding by combining forms and formations with affixoids. As we are aware of the difficulties of delimiting combining forms and affixoids from compounding (which has been excluded from the following analysis), this paper includes only the combining forms and affixoids in the derivatives extracted from the OED and the COCA. Table 1 represents the list of combining forms and affixoids classified by their position:

Table 1: Combining forms and affixoids used for data selection

Initial position						**Final position**
anti-	*mega-*	*nano-*	*pro-*	*re-*	*supra-*	*-like*
demi-	*micro-*	*non-*	*proto-*	*semi-*	*ultra-*	*-some*
half-	*mid-*	*post-*	*pseudo-*	*sub-*	*under-*	*-wise*
hyper-	*multi-*	*pre-*	*quasi-*	*super-*		

For the identification of derivatives in the OED, forms have been searched for by using the expression **lemma** (e.g. **tender*). [7] This has allowed us to obtain a list containing a high number of derivatives from a particular base. The lists were then analyzed to exclude irrelevant cases of accidental formal identity (e.g. *pretender* < *pretend* 'a person who makes a profession or assertion, esp. falsely or hypocritically') and compounds (e.g. *tender-foreheaded* 'modest, meek').

As we have already mentioned, the lexicographic data were complemented with data from the COCA. Despite the wide range of corpora available, this paper narrowed down the choice of the corpus to two of the principal corpora of English, namely, the BNC and the COCA. Table 2 offers a comparison of these two corpora that justifies the choice of the latter:

Table 2: Comparison between the BNC and the COCA

	BNC	**COCA**
Size	100,000,000	560,000,000+
Data source	1960s–1993	1990s–present day
Sample balance	10%S vs. 90%W	20%S vs. 80%W
Sample classification	Fine-grained	Less fine-grained
Hapaxes	Lower number	Higher number

As Table 2 shows, the COCA is larger in size (560,000,000 vs. 100,000,000) and contains updated information (the latest form attested in the BNC dates back to the 1990s, while the latest form in the COCA is recorded in 2017). In fact, the COCA gets expanded with 20 million words each year since 1990, evenly distributed into five genres with texts from various

[7] In some bases, such as *discipline*, the last grapheme is dropped as it is one of the requirements for some suffixes to attach (e.g. *disciplinable*, *disciplinize*).

sources.[8] Therefore, it is not surprising to find a higher number of forms (including hapaxes) contained in the COCA than in the BNC. This feature is relevant for the type of research carried out in this paper as it is conducive to obtaining as high a number of members of a paradigm as possible. For this purpose, the COCA is, to the best of our knowledge, the only corpus that "provides data for ongoing changes in English that are not available from any other source" (Davies 2011: 462).

Despite the importance of the corpus size and the number of hapaxes it contains, these features are not necessarily related proportionally to a high(er) number of forms that may be part of a specific paradigm. One of the limitations of low-frequency forms in corpora is that some of them belong, in fact, to unwanted items that may disrupt data collection. For that reason, the concordances of forms with a frequency lower than 20 occurrences were screened in order to discard names (e.g. *Bacon*, *Beghetto*) or forms from other languages (e.g. *entender* which does not correspond to the English verb *entender* 'make tender', but is recorded as part of an extract in Spanish meaning 'understand').

3.2 *Data analysis*

The data thus obtained were analyzed following the template in Table 3, which is partly based on the template designed for the international *Projekt Monika* (Pavol Jozef Šafárik University, Košice, the University St. Kliment Ohridsky, Sofia, and the University of Granada) on cross-linguistic derivational networks. An example of the partial paradigm of the base *mongrel* ('the offspring or result of cross-breeding, miscegenation, mixed marriage') is given in Table 3:

Table 3: A sample of the data file where the word-class of the base, the timeline and the hyperonymic definition is based on OED data. Forms are semantically classified according to Bagasheva (2017)

Base	**1st Der**	**W-class**	**Attested**		**Meaning**	**Sem.cat.**	**2nd Der**	**W-class**	**Attested**		**Meaning**	**Sem.cat.**
			*	†					*	†		
mongrel	*mongrel*	V	1602	1662	make (mongrel)	CAUSATIVE						
	mongrelize	V	1629	-	make (mongrel)	CAUSATIVE	*mongrelization*	N	1868	-	action of making (mongrel)	ACTION
							mongrelizing	N	1922	-	action of making (mongrel)	ACTION
							mongrelized	Adj	1857	-	made (mongrel)	QUALITY

According to lexicographic data, the verbs *mongrel* and *mongrelize* began to compete around 1630 (when the form in *-ize* is first attested), but in the second half of the 17th century, the zero-derived form was lost and only the *-ize* verb remained. The preference for the *-ize* verb is

[8] For a detailed description of the sources and the number of words per genre and year see https://corpus.byu.edu/coca/.

supported by further derivation in *-ation* (*mongrelization*), *-ing* (*mongrelizing*[N]) and *-ed* (*mongrelized*).[9]

4. Results

4.1 *Overview*

A total of 510 forms extracted from the OED and the COCA as described in §3.1 were analyzed following the template in Table 3. Although our previous research focused only on the competition between *-ize* and zero-derivation, this paper also includes other verbal bases in *-ate*, *-ify*, *-en* and the prefix *en-*, for a more comprehensive picture of verbal competition.

Figure 2 shows the number of forms attested only in the OED and those attested in the OED and the COCA classified by the affix in their base (*-ize*, zero-derivation or other affixes such as *-ate, -ify*, *-en* and the prefix *en-*):

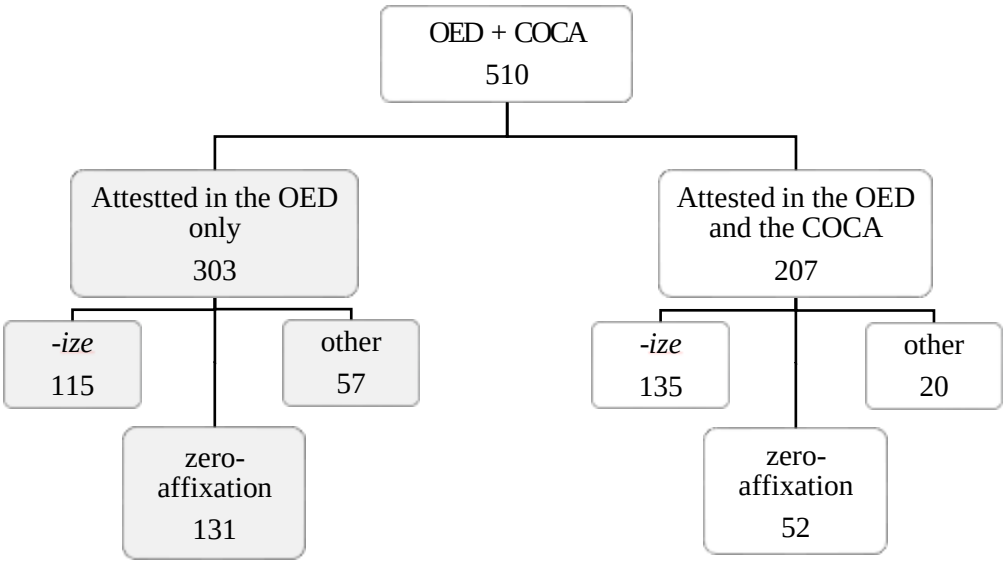


Figure 2: (Non-)attested forms with bases in *-ize*, zero-derivation and other affixes as extracted from the OED and attested in the COCA

Of the 510 forms extracted, more than a half (303) are attested only in the OED. Of those 303, 131 are forms with a base in zero-derivation, followed by 115 forms with their bases in *-ize* and 57 forms with other affixes. In contrast, of the 207 forms that are attested both in the OED and the COCA, most (135) are forms with their bases in *-ize*, while 52 are forms in zero-derivation. Again, the group containing the lowest number of forms (20) is that where bases end in affixes other than *-ize* and zero-derivation, i.e. *-ify*, *-ate*, *-en* and the prefix *en-*.

[9] In order to follow the most inclusive approach possible, the suffixes *-ed* and *-ing* are included in this paper when they are recorded in the OED as separate entries, either as adjectives (in *-ed* or *-ing*) or as nouns (in *-ing*), despite their controversial nature as in-between cases on the inflection/derivation cline.

A comparison of the total number of forms in *-ize* and zero-derivation (whether attested only in the OED or both in the OED and the COCA) shows that the two affixes follow opposite patterns. The number of forms with their bases in *-ize* that are attested in the Dictionary and the Corpus (135) is higher than those only attested lexicographically (115). In contrast, the number of attested forms in zero-derivation in the OED and the COCA is markedly lower (52) than those only attested in the OED (131). Forms in or with their bases in *-ify*, *-ate*, *-en* and the prefix *en-* represent the smallest group both as attested lemmas in the OED (57) and in both sources (20). This apparent preference for the suffix *-ize* in the COCA is in line with previous diachronic research (Fernández-Alcaina 2017). Specifically, the main results for the competition between *-ize* and zero-derived verbs showed that:

i) apparently, the suffix *-ize* has gradually replaced zero-derivation for the expression of CAUSATIVE in the clusters analyzed while zero-derivation seems to be more frequent in the expression of INSTRUMENT and MANNER, and
ii) the resolution of competition takes various shapes: the loss[10] of one of the forms, semantic specialization or the loss of the two forms, sometimes in favor of a third form. Cases of an ongoing competition are also observed where evidence does not clearly support a preference for one or the other verb. Specifically, 15[11] out of the 45 clusters analyzed were described as displaying unresolved competition.

The rest of this section is divided as follows: §4.2 describes the increasing preference observed for the suffix *-ize* over zero-derivation. §4.3 compares the outcome of competition in previous research and in this paper, with a focus on the profiles found in §4.3.1 and with a focus on the special cases observed in §4.3.2. The final section (§4.4) is a summary of the results obtained.

4.2 *Increasing replacement*

Research into the competition between the suffix *-ize* and zero-derivation over time showed that the introduction of the former led to losses in the use of zero-derivation for the expression of CAUSATIVE (Fernández-Alcaina 2017), which is especially marked from the 17th century onwards. The pattern is confirmed by the results presented in this paper where the availability of a form also depends on the availability of its derivatives in the Corpus. Figures 3 and 4 show a comparison of the diachronic development of verbs in *-ize* and zero-derivation in previous research and in this paper. The timelines for the verbs in *-ize* (Figure 3) and zero-derivation (Figure 4) are according to the number of available forms created from the 14th century onwards until the end of the 20th century (axis *y*). The green line represents forms in *-ize* and zero-derivation that were attested in the OED as in use and recorded in corpora, specifically, in the BNC and the COCA, (axis *y*). Data represented by the blue line also take into account whether *-ize* or zero-derived forms have served as bases for further derivation by considering as available those verbs that serve as bases for derivatives attested in the OED as in use and recorded in the COCA by the attachment of affixes other than *-ed* or *-ing*.[12] In other words,

[10] The term *loss* is used in this paper following Tichý's (2018) terminology.
[11] See footnote 6.
[12] Data from the BNC are not used in this paper as previous research showed that the forms recorded in the BNC were usually also recorded in the COCA. This can be exemplified using the clusters *soberize*/*sober*, and *fossilize*/*fossil*: *soberize* is recorded neither in the BNC nor in the COCA, in contrast to *sober*, which has a normalized

although *-ed* and *-ing* forms were included in data collection, if they are the only derivatives attested, they are not counted as evidence for the availability of a verb in Figures 3 and 4.

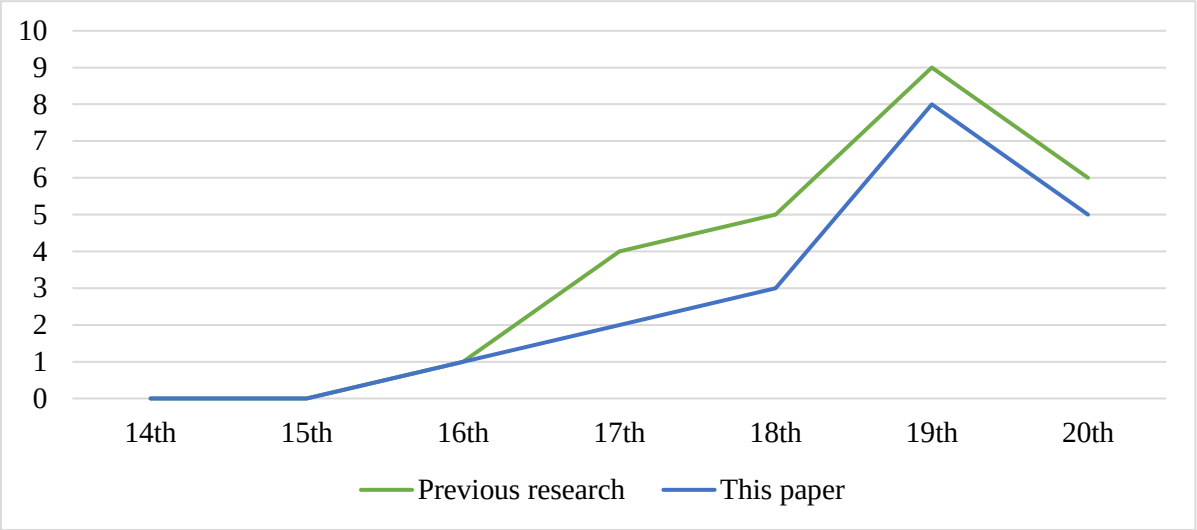


Figure 3: Comparison of the availability of forms in *-ize* from previous research (in green) in contrast to this paper (in blue), where the derivatives of the competing verbs are considered

Since the introduction of the suffix *-ize* in the 16th century, its use has gradually increased until the 19th century. Figure 3 shows that the same pattern is found occurring in this paper in which further derivation of the bases in competition is considered. Let's exemplify this difference with the cluster *ghetto*/*ghettoize*:

(1) *Jews, who are **ghettoed** under the racial legislation.*
(*Times*. 15 Feb. 11 vs. 3, 1936, OED2)

(2) *Arcand's attempt … to **ghettoize** a minority.*
(*Canadian Jewish Chron*. 4 Aug. 3, 1939, OED2)

According to the OED, both forms date back to the 1930s and are classified as in use by the OED, even if only the latter is recorded in the COCA. However, is this enough evidence to conclude that the *-ize* suffix wins out over zero-derivation? The claim that competition has been resolved in a cluster such as *ghetto*/*ghettoize* is risky because both forms are relatively new and, moreover, their entries in the OED have not been updated since 1989. However, the fact that only the *-ize* form presents further derivation suggests that it is the preferred option for the expression of CAUSATIVE because it allows further derivation related to the semantic category expressed by the base. In the same example, only *ghettoize* has derivatives attested in the COCA (e.g. *ghettoization* is recorded in the COCA with a normalized frequency of 0.08).

frequency of 1.52 in the BNC and 1.32 in the COCA. Similarly, *fossilize* is recorded in the BNC with a normalized frequency of 0.20 and of 0.22 in the COCA. Its competitor, *fossil*, is recorded in neither of the two corpora.

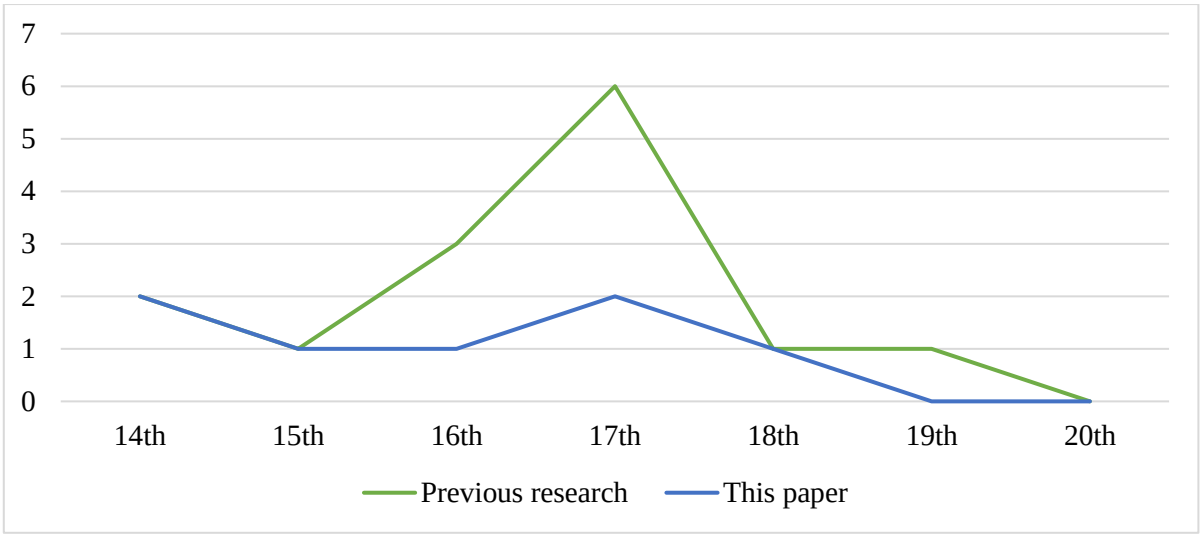


Figure 4: Comparison of the availability of forms in zero-derivation from previous research (in green) in contrast to this paper (in blue), where the derivatives of the competing verbs are considered

Likewise, the development of zero-derivation in this paper shows similarity to that observed in previous research. Consider the cluster *English*/*Englishize*:

(3) *A New York tailor is advertising: Let us take your Stateside suit and* ***English*** *it up.* (*Evening Standard* 10 Dec. 6/6, 1965, OED3)

(4) *Why then do they tend to* ***'Englishize'*** *the pronunciation of Italian words, but not the French or Hispanic?*
(*Post-Standard* (Syracuse, N.Y.) (Nexis) 19 Feb. d 3, 2006, OED3)

The two forms apparently co-exist according to the OED, but only the zero-derived form is recorded in the COCA as a verb. A look at their subparadigms shows that only *English* has derivatives recorded in the COCA (e.g. *Englishable*) or other derivatives such as *re-English*[V], which are not recorded in the Corpus but attested in the OED as neologisms.

A comparison of the patterns followed by verbs in *-ize* (Figure 3) and zero-derivation (Figure 4) in previous research and in this paper shows again that the 17th century is a turning point in the competition between the two forms. While the suffix *-ize* has been increasingly used since its first record in the 14th century (in the clusters analyzed), derivatives in zero-derivation have started to decrease from the 17th century onwards.

4.3 *Outcomes of the competition between* -ize *and zero-derivation*

The members of the subparadigms where the competing forms serve as bases also provide information regarding the outcomes of the diachronic competition between *-ize* and zero-derivation. Specifically, the existence of derivatives mapping on the competing sense of one of the verbs can provide clues about the preference for one of the forms to express a certain semantic category. Figure 5 compares the number of clusters with forms in an ongoing and resolved competition in Fernández-Alcaina (2017) and in this paper:

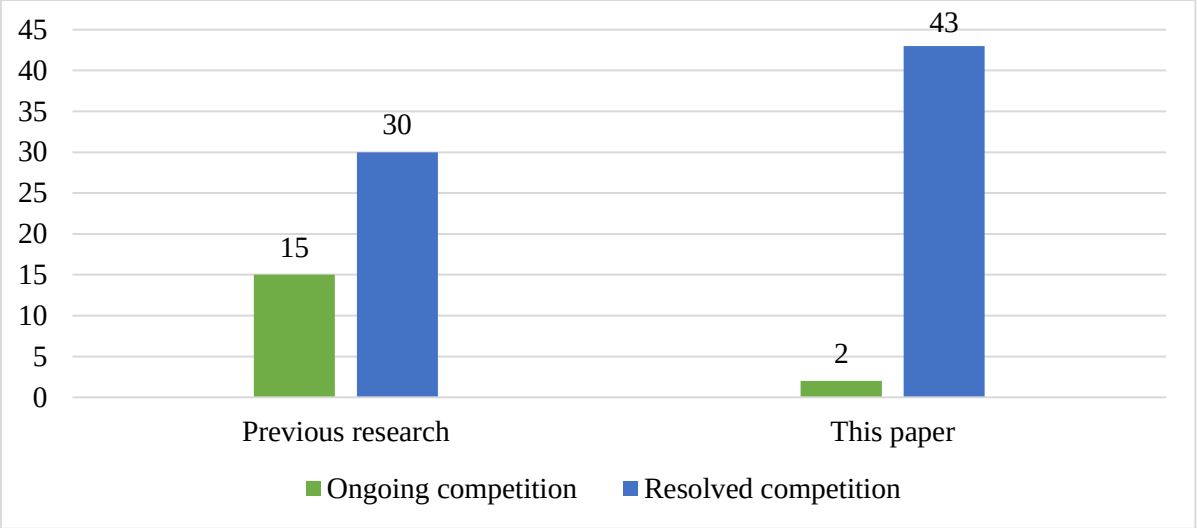


Figure 5: Total number of clusters in ongoing competition (in green) and resolved competition (in blue) both in Fernández-Alcaina (2017) and this paper

As Figure 5 shows, the number of clusters where forms co-exist decreases when the subparadigms created by the bases are considered as additional evidence. Thus, of the 15 clusters (out of 45) that were found to exhibit an ongoing competition in previous research, only two clusters remain as cases of ongoing competition in this paper. Table 4 expands upon the new classification of the 13 clusters, now most of them classified as instances of resolved competition:

Table 4: Data comparison between previous and the present paper

	Previous research	**This paper**
PDE competition?	15	2
-ize = CAUSATIVE	12	19
Zero-derivation = CAUSATIVE	4	7
Semantic specialization	8	8
Both forms are obsolete	6	9

Of the 13 clusters that have changed their status from an ongoing to resolved competition, in seven of them the suffix *-ize* is preferred over zero-derivation for the expression of CAUSATIVE, while only in three clusters zero-derivation wins out over the suffix *-ize*. For example, based

on lexicographic data, forms in the cluster *ghetto*/*ghettoize* were initially considered to co-exist in use, but only *ghettoize* was found attested in the COCA. Besides, only the form in *-ize* presents further derivation by derivatives recorded in the Corpus (e.g. *ghettoization*). Therefore, the resolution of competition points at a preference for the suffix *-ize* over zero-derivation for the expression of CAUSATIVE.

The opposite pattern is found in clusters such as *discipline*/*disciplinize*. The latter was marked as rare by the OED, and the latest attestation in the Dictionary dates back to 2003. Furthermore, only *discipline*^{V} has attested derivatives in the COCA (e.g. *disciplining*^{N}, *undisciplinedness*). The only derivative found for *disciplinize* (*disciplization*) is marked as obsolete by the OED (1706–1706) and is not attested in the COCA.

The verbs in the remaining nine clusters are obsolete either because a third form may be preferred (e.g. *grand*/*grandize*/*aggrandize*) or because the forms simply disappear (e.g. *pemmican*/*pemmicanize*). According to the OED, *grand*^{V} was last attested in the 17th century while *grandize* was last attested in 2014. A closer look at this specific cluster reveals that:

i) none of the forms presents further derivation, and
ii) other forms have competed for the expression of the same semantic category CAUSATIVE at some point in history: *aggrandize* (1634–), *engrandize* (1652–1883), *engrand* (1655–1655) and *grandify* (1665–).[13]

Apart from the OED dates of attestation, none of the verbs in (ii) is recorded in the COCA, except for *aggrandize* and its derivatives. However, it is worth mentioning that, in forms such as *grandify* and *grandize* (both last attested in the 20th century in the OED), there is an interval of around a century between the last two records, as (5) and (6) show:[14]

(5) *Repudiating, as I do, all idea of grandifying London at a coup, or to any great extent formalising it.*
(*Brit. Architect* 6 Aug. 93/2, 1897, OED3)

(6) *It would have been two cottages that were joined in the 18th century, with the pediment added in an attempt to 'grandify' it.*
(*Sunday Tel.* (Nexis) 21 Aug. Stella 42, 2011, OED3)

A similar case is observed in the OED records for *grandize*, where the time gap is even bigger because the record leaps from the 17th century (7) to the 19th century (8) and to the 21th century (9):

(7) *Both [love and fear] together, are to the sanctified Soul, as Ballast to a Ship, to keep it steady, and doth grandize, elevate, and enlarge each affection.*
(J. Harrington, *Horæ Consecratæ* (1682) 154, a1680, OED3)

[13] Even though the form is marked as rare by the OED, there is an attestation dating back to 2011.
[14] The observed leaps in lexicographic records may be a consequence of the well-known limitations of historical dictionaries (see §5). However, we cannot exclude the possibility of dealing, in cases such as these, with instances of Bauer's (2014) renewed availability (see §5 for further details).

(8) *I have been so grandized, so dazzled, so overawed, that I have scarcely been able to breathe.*
(F. J. Hall, *Next of Kin II*. i. 32, 1854, OED3)

(9) *The more beautiful I made her, the more she hated it. She accused me of grandising her journey*.
(*Newcastle* (Austral.) *Herald* (Nexis) 8 Mar. 28, 2014, OED3)

4.3.1 *Profiles of resolved competition*

Table 4 has presented the possible paths of resolution of competition in the 45 verbal clusters under study. This section elaborates on clusters where competition is resolved (or is on its way to be resolved) by describing the profiles observed in the clusters when subparadigms are considered. Profiles have been labeled using a representative cluster within each group: *revolutionize*-like clusters, *ghettoize*-like clusters and *sober*-like clusters. In the first two profiles, the *-ize* form is the preferred option to express the CAUSATIVE meaning. The difference lies in the fact that in *ghettoize*-like clusters, both forms in zero-derivation and in *-ize* present further derivation whereas only the *-ize* verb has derivatives in *revolutionize*-like clusters. In the case of *sober*-like clusters, the resolution of competition occurs in favor of the zero-derived verb.

***Revolutionize*-like clusters**

In these clusters only the *-ize* form has derivatives that convey the sense 'make X'. These derivatives may be recorded in the Corpus and some of them show recursive derivation through the attachment of combining forms or affixoids.

A first set of clusters displaying this profile was already identified as exhibiting cases of resolved competition in Fernández-Alcaina (2017), where the information provided by their derivatives supported a scenario of competition resolved in favor of the suffix *-ize*:

i) *public*/*publicize*
ii) *coward*/*cowardize*
iii) *mongrel*/*mongrelize*
iv) *idol*/*idolize*
v) *parallel*/*parallelize*
vi) *romantic*/*romanticize*
vii) *oxide*/*oxidize*

Other clusters were described as displaying an ongoing competition in previous research, usually because lexicographic and corpus data did not point at any clear preference between *-ize* and zero-derivation. However, when their derivatives have been taken into account, they indicate a preference for the *-ize* form as its derivatives map on the sense 'make X':

i) *revolution*/*revolutionize*
ii) *glamour*/*glamourize*
iii) *pauper*/*pauperize*
iv) *slender*/*slenderize*
v) *aerosol*/*aerosolize*

vi) *legend*/*legendize*
vii) *canal*/*canalize*

Table 5 exemplifies this profile using the cluster *revolution*/*revolutionize*. Although *revolutionize* and *revolution*^{V} are first attested within a short span of time (1795 and 1805, respectively) and both forms are recorded as in use by the OED, only the former allows further derivation:

Table 5: Subparadigms for the cluster *revolution*/*revolutionize* with specification of their base (*revolution*^{N}), the level of derivation, the dates of the earliest and latest attestation based on the OED and the semantic category following Bagasheva (2017)

Base	1st. Der.	W-class	Attested *	†	Sem.cat.	2nd Der.	W-class	Attested *	†	Sem.cat.	3rd Der.	W-class	Attested *	†	Sem.cat
revolution	*revolutionize*	V	1795	-	CAUSATIVE	*re-revolutionize*	V	1803	-	ITERATIVE					
						revolutionization	N	1871	-	ACTION					
						revolutionized	Adj	1798	-	QUALITY	*unrevolutionized*	Adj	1797	-	PRIVATIVE
						revolutionizing	N	1797	-	ACTION					
						revolutionizing	Adj	1797	-	QUALITY					
						revolutionizement	N	1820	1820	ACTION					
						revolutionizer	N	1798	-	AGENT					
						counter-revolutionize	V	1827	1827	OPPOSITIVE					
	revolution	V	1805	-	CAUSATIVE										

Other clusters, such as *uniform*/*uniformize* and *tender*/*tenderize*, were initially classified as exhibiting semantic specialization where the zero-derived verb expressed the semantic category CAUSATIVE, whereas the *-ize* verb was restricted to a specific domain. Specifically, according to the OED, *uniformize* is used in mathematics, whereas *tenderize* is used about food with the meaning 'make meat tender'. This has been partly confirmed in this paper: in both cases the *-ize* verb is specialized in meaning, but in none of the clusters does the zero-derived form keep conveying the general sense of 'make X'. For example, *uniformization* is not shown to be restricted to the field of mathematics in the COCA:

(10) *The Nazarite matrons' IsiZulu represents the first-ever true **"uniformization"** of one regional variant of twentieth century Zulu speakers' folk attire.*
(COCA: 2004 ACAD African Arts)

In the cluster *tender*/*tenderize*, the corpus data support the latter as a term specifically used about food:

(11) *It called for more research but concluded the risk was only slightly higher: about seven additional illnesses due to **tenderization** for every billion steak servings.*
(COCA: 2003 NEWS Atlanta)

Moreover, the position of *tender* in the competition as reflected by the COCA appears to be further weakened by the fact that the concordances for *tender* in the COCA mostly refer to its

homonym *tender*^{V} (from French *tendre*^{V} 'extend, give') meaning 'to offer or present formally for acceptance', and not to 'make tender':

(12) *Therefore, I have no choice but to* ***tender*** *my resignation.*
(COCA: 2017 FIC Analog Science Fiction & Fact)

Figure 6 shows the timelines for *revolutionize*-like clusters based on the dates of the earliest and latest attestation provided by the OED:

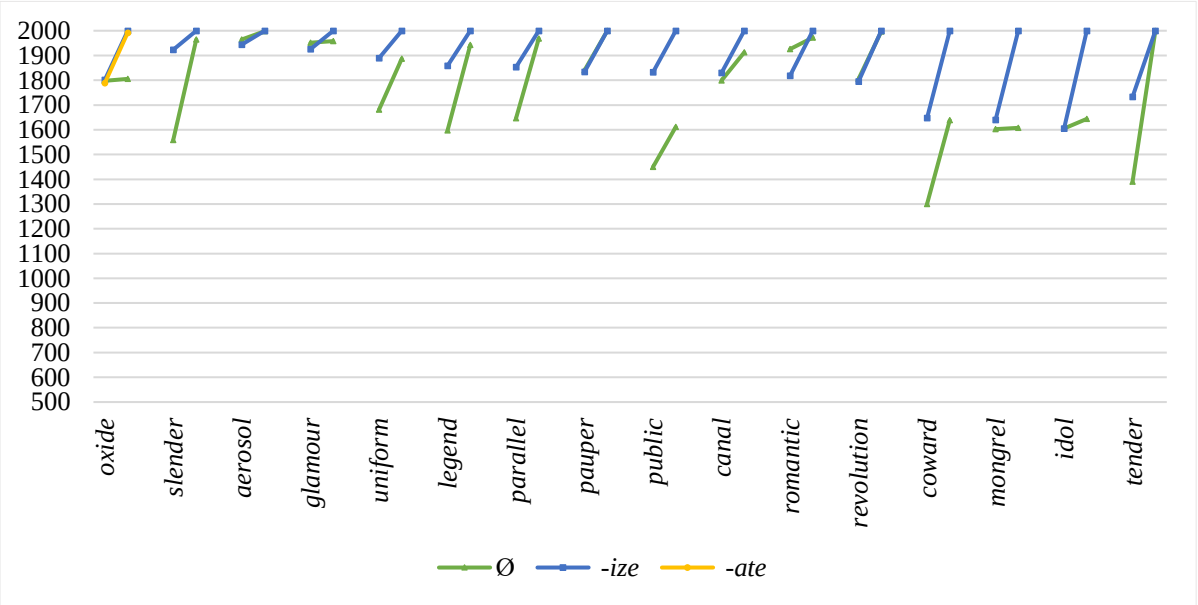


Figure 6: Timeline for the *revolutionize*-like clusters in Ø (in green), *-ize* (in blue) and *-ate* (in yellow) when forms compete for the expression of CAUSATIVE

All the zero-derived forms in the clusters shown in Figure 6 are first attested well before their *-ize* competitors (*slender*/*slenderize*, *uniform*/*uniformize*, *legend*/*legendize*, *parallel*/*parallelize*, *public*/*publicize*, *coward*/*cowardize* and *tender*/*tenderize*), or within a short interval of time with respect to the *-ize* verb (*oxide*/*oxidize*, *aerosol*/*aerosolize*, *glamour*/*glamourize*, *pauper*/*pauperize*, *canal*/*canalize*, *romantic*/*romanticize*, *revolution*/*revolutionize*, *mongrel*/*mongrelize* and *idol*/*idolize*). In general, what all the clusters have in common is that zero-derived forms fail to be further derived, unlike their competitors in *-ize*, including cases when *-ize* is first attested much later.

Ghettoize-like clusters

The second profile groups clusters where both forms in *-ize* and in zero-derivation have derivatives attested in the OED, but where the former seems to be preferred.

Some of the clusters had already been identified as cases of resolved competition in previous research, either due to the loss of the zero-derived verb, as with the zero-derived form in the cluster *immune*/*immunize*, or via semantic specialization, as in *union*/*unionize* (Fernández-Alcaina 2017: 196). The remaining clusters were originally classified as instances

of ongoing competition, but the existence of derivatives with *-ize* bases may indicate a preference for this suffix over zero-derivation:

i) *ghetto*/*ghettoize*
ii) *oval*/*ovalize*
iii) *fossil*/*fossilize*
iv) *proselyte*/*proselytize*
v) *Latin*/*Latinize*

Table 6 illustrates this profile using the cluster *ghetto*/*ghettoize*:

Table 6: Subparadigms for the cluster *ghetto*/*ghettoize* with specification of their base (*ghetto*^{N}), the level of derivation, the dates of the earliest and latest attestation based on the OED and the semantic category following Bagasheva (2017)

Base	1st. Der.	W-class	Attested		Sem.cat.	2nd Der.	W-class	Attested		Sem.cat.
			*	†				*	†	
ghetto	*ghetto*	V	1936	-	CAUSATIVE	*ghettoed*	Adj	1970	-	QUALITY
	ghettoize	V	1939	-	CAUSATIVE	*ghettoization*	N	1939	-	ACTION
						ghettoized	Adj	1990	-	QUALITY
						ghettoizing	N	1990	-	ACTION
						ghettoizer	N	1997	-	AGENT

In all the cases, *-ize* forms are recorded in the COCA, as well as most of their derivatives: *ghettoized* is recorded in the COCA with a frequency of 0.05, and *ghettoization*, 0.08. *Ghettoizing*^{N} is recorded only once.[15] In contrast, apart from *-ed* adjectives, zero-derived verbs do not show further derivation through the attachment of other affixes, unlike *-ize*. None of the zero-derived verbs has prefixed derivatives, except for *proselyte*^{V} and its negative form *unproselyte*, which is marked as obsolete by the OED and is not recorded in the COCA.

Figure 7 shows the timelines for the forms in *ghettoize*-like clusters with the dates of the earliest and latest attestation provided by the OED:

[15] In some clusters, only *-ize* derivatives serve as bases for compounds (e.g. *self-ghettoization*). As compounding has fallen out of the scope of this paper such cases have not been considered, even if they may prove relevant in the interaction between paradigms and competition within clusters.

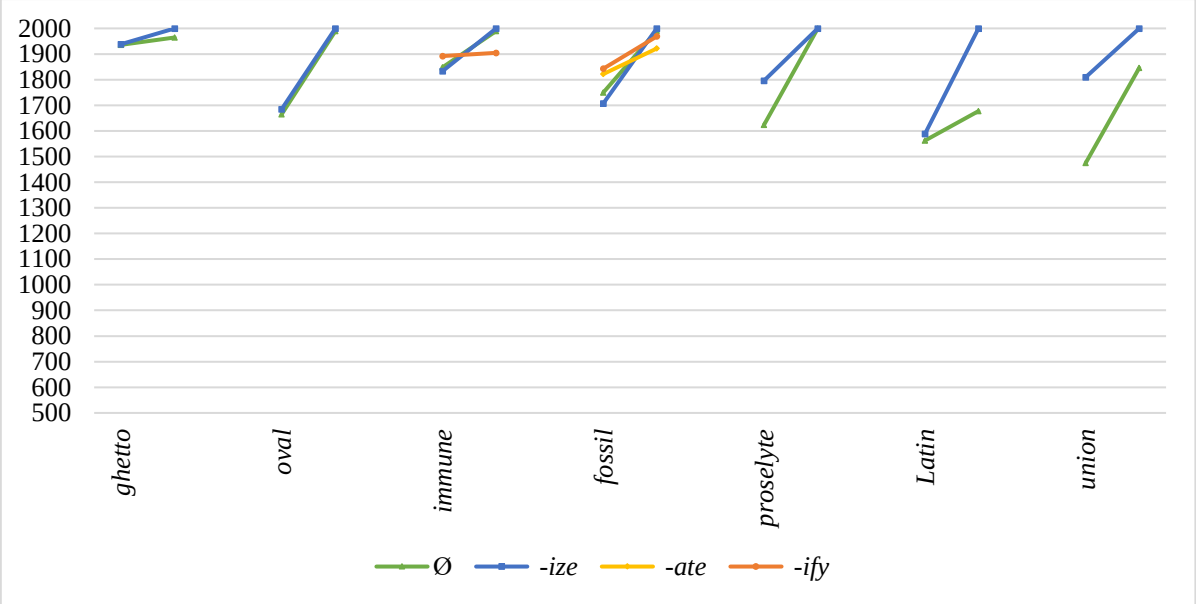


Figure 7: Timelines for the *ghettoize*-like clusters in Ø (in green), *-ize* (in blue), *-ate* (in yellow) and *-ify* (in orange) when forms compete for the expression of CAUSATIVE

Figure 7 shows that there are clusters (*immune*/*immunize*, *fossil*/*fossilize*) where more than two forms have competed for the expression of CAUSATIVE, either in *-ify* (*immunify*, *fossilify*) or in *-ate* (*fossilate*). In both clusters, *-ify* and *-ate* forms are first attested in the 19th century, later than their competitors in *-ize* and zero-derivation. However, they are marked as obsolete by the OED in the first half of the 20th century, thereby reducing competition, in theory, to two forms. As for dates of attestation, the pattern found is similar to that in Figure 6. Zero-derived forms are usually attested before their *-ize* competitors in the clusters *proselyte*/*proselytize* and *union*/*unionize*, or attested within a short timespan as regards their competitors in the clusters *ghetto*/*ghettoize*, *oval*/*ovalize*, *immune*/*immunize*, *fossil*/*fossilize* and *Latin*/*Latinize*.

***Sober*-like clusters**

In all the clusters where zero-derivation is preferred over the suffix *-ize*, most forms have derivatives but only zero-derivation shows further derivation apart form *-ed* and *-ing*. Clusters such as *cuckold*/*cuckoldize* and *gentle*/*gentilize* were already described as cases of resolved competition: *cuckoldize* is marked as obsolete in the OED and *gentilize* refers to 'live like a Gentleman', rather than to 'make gentle'. In contrast, the clusters *sober*/*soberize*, *English*/*Englishize*, *discipline*/*disciplinize* and *quiet*/*quietize* were initially identified as instances of ongoing competition. Table 7 exemplifies this profile with the cluster *sober*/*soberize*:

Table 7: Subparadigms for the cluster *sober*/*soberize* with specification of their base (*sober*[ADJ]), the level of derivation, the dates of the earliest and latest attestation based on the OED and the semantic category following Bagasheva (2017)[16]

Base	1st Der.	W-class	Attested *	Attested †	Sem.cat.	2nd Der.	W-class	Attested *	Attested †	Sem.cat.	3rd Der.	W-class	Attested *	Attested †	Sem.cat.
sober	*ensober*	V	1651	1651	CAUSATIVE										
	sober	V	1797	-	CAUSATIVE	*sobered*	Adj	1797	-	QUALITY					
		V	1820	-	PROCESS	*soberer*	N	1849	-	AGENT					
						sobering	Adj	1510	-	QUALITY	*soberingly*	Adv	1923	-	MANNER
						sobering	N	1510	-	ACTION					
						unsober	V	1856	-	PRIVATIVE					
	soberize	V	1707	-	CAUSATIVE	*soberized*	Adj	1840	-	QUALITY					
						soberizing	Adj	1860	-	QUALITY					
		V	1831	1831	PROCESS										

Figure 8 shows the timelines for the *sober*-like clusters:

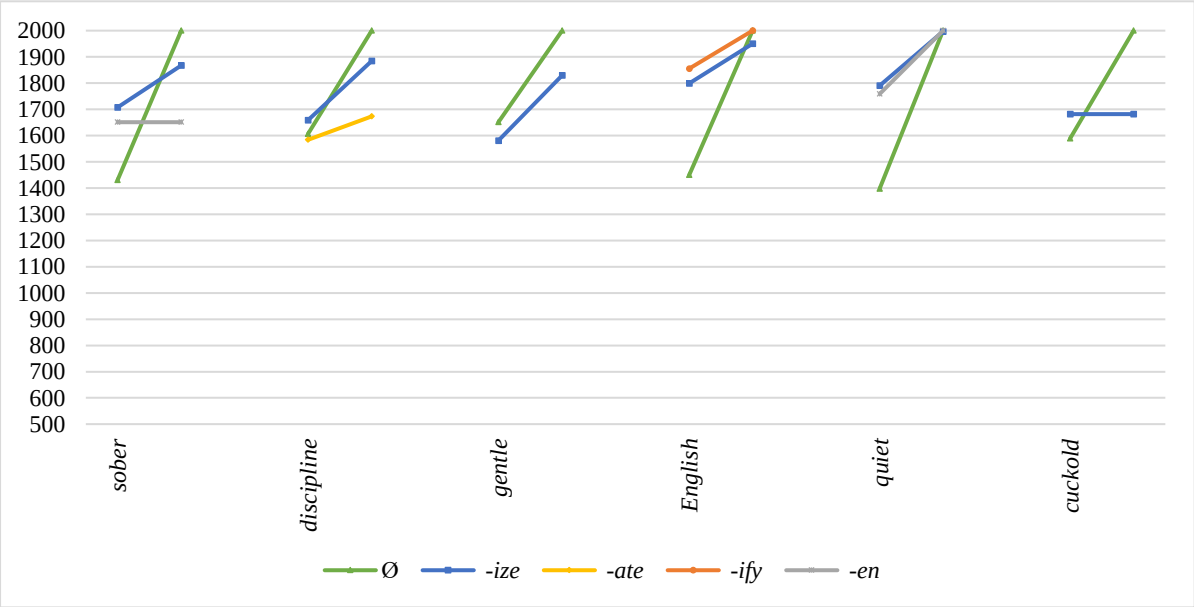


Figure 8: Timelines for the *sober*-like clusters in Ø (in green), *-ize* (in blue), *-ate* (in yellow) and *-ify* (in orange) and *en-* (in grey) when forms compete for the expression of CAUSATIVE

In all the clusters, and in line with Figures 6 and 7, zero-derived verbs appear much earlier than the *-ize* form (*English*/*Englishize*, *quiet*/*quietize*) or within a short timespan (*sober*/*soberize*, *discipline*/*disciplinize*, *gentle*/*gentilize* and *cuckold*/*cuckoldize*). The difference from *revolutionize*-like clusters and *ghettoize*-like clusters is that, in the clusters in Figure 8, the zero-derived form presents further derivation mapping on the sense 'make X'. It is possible

[16] The subparadigm for *soberize* is a case of an impoverished set as it does not contain forms derived by affixes other than *-ed* and *-ing*.

that the *-ize* verbs in this particular cluster appeared as unsuccessful innovations, even if some of them may not be marked as obsolete by the OED.

4.3.2 *Special cases*

Although the clusters in §4.3.1 were described as cases of resolved competition, some of them may still overlap in meaning, especially if those are attested in the 20th century.[17] For instance, the boundary between the semantic categories CAUSATIVE and INSTRUMENT or MANNER sometimes appears fuzzy in the cluster *aerosol*/*aerosolize*. Examples (13) and (14) illustrate this:

(13) *Self-assembling nano-bots can be **aerosoled** (perhaps "have been" would be more appropriate) and breathed…*
(http://forbiddenknowledgetv.net/stockholm-the-mark-of-the-beast-is-here/)

(14) […] *experiment, which is being conducted to determine the maximum temperature output of the coils used to **aerosolize** the e-liquid in the e-cigs under different configurations and conditions.*
(COCA: 2016 MAG Medical Xpress)

Likewise, the study of competition is also hindered by incomplete lexicographic information and/or lack of corpus data. For example, information in the OED (2) on the cluster *finite*/*finitize* is scarce and neither of the forms is recorded in the COCA. Any assumption about the resolution of competition in this cluster can therefore only be tentative:

(15) *The Lord to be in them, there to personate and **finite** himself.*
(H. Bushnell, *Christian Nurture* ii. v. 301, 1861, OED2)

(16) *The Unconditional has been under a necessity to **finitise** Itself.*
(S. S. Laurie, *Synthetica II*. 859, 1906, OED2)

In some other cases, the senses that are marked as in use by the OED do not reflect in the data provided by the Corpus. For example, according to the OED, *woman*^{V} and *womanize* have competed for the sense 'make womanly', but competition has apparently been resolved in favor of *womanize*. Nevertheless, the corpus data usually refer to the intransitive sense 'to engage in casual sexual or romantic encounters with women'. Some derivatives (e.g. *womanizer*, *womanizing*) appear to map on this sense:[18]

(17) *Dara may have tolerated Jon's **womanizing**, but according to the FBI, Jon's latest romance with a Thai woman seemed like more than a fling…*
(COCA: 2010 SPOKEN)

[17] Corpus frequency may be a guiding factor in the resolution of the competition in this type of cases, but as the information on the (un)availability of forms provided by the COCA is sometimes scarce or lacking altogether, the description of the resolution of competition could again benefit from the use of historical corpora.

[18] This assumption must be taken with caution as it is based on mere observation. More information on this issue could be extracted from a thorough analysis of the semantic categories or senses expressed by their corpus concordances. For details on the relevance of this type of analysis for the study of competition see Lara-Clares (2017) and Lara-Clares (2018) on nominal competition.

In some other cases, the OED marks one of the forms as associated with a special use. For instance, the zero-derived form in the cluster *wanton*/*wantonize* is marked as *poetic*, but corpus evidence is again lacking:

(18) *The same breeze that had uncivilized him seemed to have* ***wantoned*** *her.*
(S. Carroll, *Bride Finder* xvi. 232, 1998, OED3)

(19) *If they meet a girl who is not wanton, they* ***wantonize*** *her in their minds.*[19]
(C. Stead, *Letty Fox* xiii. 116, 1946, OED3)

There are also examples where the resolution of competition between zero-derivation and the suffix *-ize* occurs, but the remaining form is in a complementary distribution with a third form. A case in point is the *quiet*/*quietize*/*quieten* cluster. With *quietize* ousted, the use of the verbs in the reduced cluster *quiet*/*quieten* seems to depend on regional varieties of English: the verb *quiet* shows a frequency of 0.39 in the BNC and 2.77 in the COCA. By contrast, *quieten* is recorded in the BNC with a frequency of 1.64, and 0.06 in the COCA. No obvious semantic difference has been observed in their use:

(20) *He tried to find a compromise that would satisfy his artistic urge and* ***quieten*** *his conscience*.
(BNC: K8R W_fic_prose)

(21) *Daniel held up his hand to* ***quiet*** *them, and it took them longer than usual to fall silent…*
(COCA: FIC_Analog Science Fiction & Fact)

4.4 *Recapitulation*

The results presented in this paper confirm previous research regarding the increasing preference for the suffix *-ize* for the expression of CAUSATIVE in clusters where it competes with zero-derivation. In particular, of the 510 forms extracted from the OED and the COCA for the creation of the subparadigms, 131 derivatives with bases in zero-derivation are not recorded in the COCA, in contrast to the 52 forms recorded. On the other hand, 115 *-ize* derivatives do not appear in the Corpus whereas 135 do. Opposite patterns regarding the number of derivatives attested and not attested in the COCA are supported from a diachronic perspective. In fact, the development of both affixes in Figures 3 and 4 above shows that the introduction of the suffix *-ize* led to losses of the zero-derived verbs expressing the semantic category CAUSATIVE from the 17th century onwards, when zero-derivation reached its peak. At the same time, this decrease in the use of zero-derivation was accompanied by an increase in the number of *-ize* verbs, which continued growing until the 19th century.

Regarding the profile displayed by the clusters where competition is resolved (or, apparently, on its way to resolution), CAUSATIVE is mainly expressed by forms in *-ize*, although there is a small group of clusters where zero-derivation wins out over forms in *-ize*. In all these clusters, labeled here as *sober*-like clusters, zero-derived forms are attested before their *-ize*

[19] The choice for the form *-ize* in this particular example may be influenced by the role of immediate language context by contrasting to the adjective *wanton* in the same sentence. Context-based lexical and stylistic choices of this type may be worth considering among influential factors in the future study of competition in word-formation.

competitors (i.e. before the 17th century) and are the only ones of this type whose derivatives are recorded in the COCA. This implies that *-ize* forms are usually preferred but, once a paradigm has been created around a zero-derived verb, a change in favor of the *-ize* counterpart appears to be less likely.

5. Discussion

The results in §4 show that the description of the subparadigms created by two competing forms contributes to the study of morphological competition because subparadigms may provide additional data on the preference for one or the other form.

The fact that the one form rather than the other triggers further derivation may indicate that this form is better established in English, and thus tends to determine the outcome of the competition of forms within a cluster. Even if it were the case that the two competitors have derivatives, it would not necessarily imply that they co-exist. In fact, although the sense in the derivatives usually map on the sense of their bases, "the mapping is never complete, and it is not infrequently narrowed down to the central senses of the base" (Bauer & Valera 2015: 83). Therefore, identifying the sense to which the derivatives refer is relevant for the study of competition, as it may indicate whether the sense for which two or more forms compete is more central in some competitors. In fact, a comparison between the results in this paper and those of a previous study (Fernández-Alcaina 2017) shows that cluster classification may change if members in the subparadigms are considered.

Another aspect of competition that can be better understood if the role of subparadigms is taken into account is the profile displayed by the clusters where competition has been resolved (or is on the way to be resolved). In the set of competitors where the suffix *-ize* wins out over zero-derivation for the expression of CAUSATIVE, the zero-derived form is usually attested earlier than its *-ize* counterpart. This leaves some room for zero-derived verbs to be derived further. In contrast, in the clusters where zero-derivation wins out over the suffix *-ize* in CAUSATIVE senses, the latter is usually short-lived and does not act as the base for any derivative. It seems that, in the clusters analyzed, once a causative zero-derived verb has derivatives mapping on this sense, *-ize* verbs are less likely to replace them, and thus, the subparadigm of the zero-derived competitor seems to support a preference for its base. Figure 9 shows the timelines for the verbs *English* and *Englishize* and their respective derivatives:

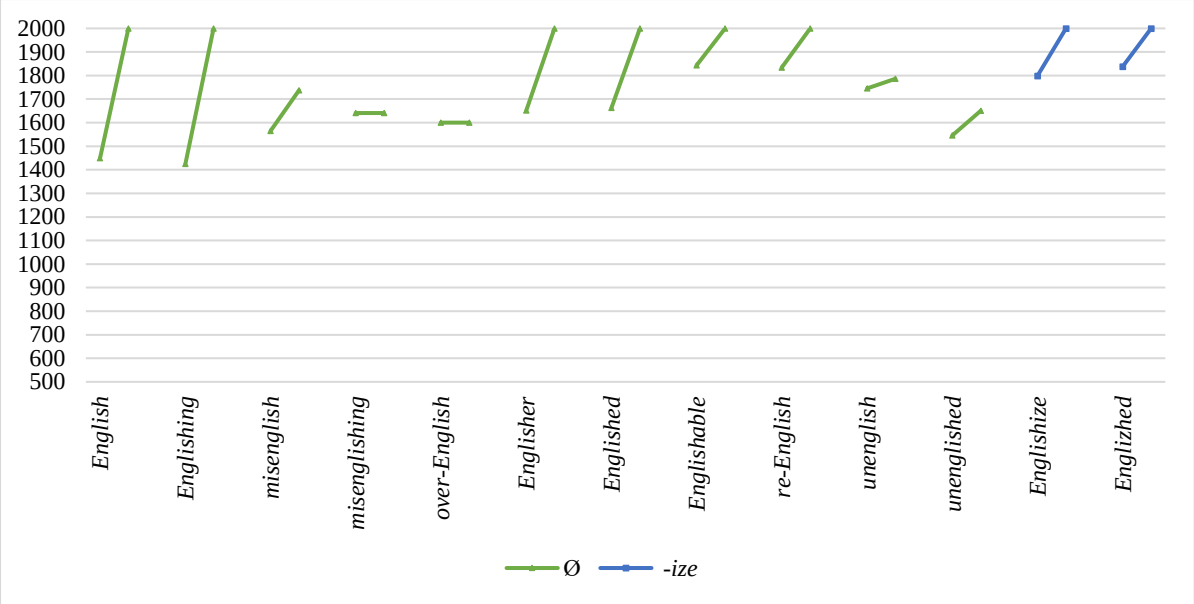


Figure 9: Timelines for the forms in Ø (in green) and in *-ize* (in blue) and their derivatives in the cluster *English*/*Englishize*

The subparadigm of *English* exemplifies cases where the derivatives from the zero-derived form are attested earlier than the *-ize* competitor, including those that remain in use. In contrast, although the latest dates of attestation for *Englishize* and *Englishized* belong to the 20th century, it may be partly a consequence of the relative youth of these forms.

However, there are clusters where it is difficult to prove whether the subparadigm indicates the direction in which the resolution of competition will occur or not. This appears especially so if the two forms are attested within a short timespan, e.g. *ghetto* (1936–, OED2)/*ghettoize* (1939–, OED2). Confirmation of the extent to which this applies requires an analysis of the same pattern in clusters where the two forms have close dates of the earliest attestation and where the competition has been resolved, e.g. *Latin*/*Latinize*. The timelines for the verbs in this cluster and their derivatives are shown in Figure 10:

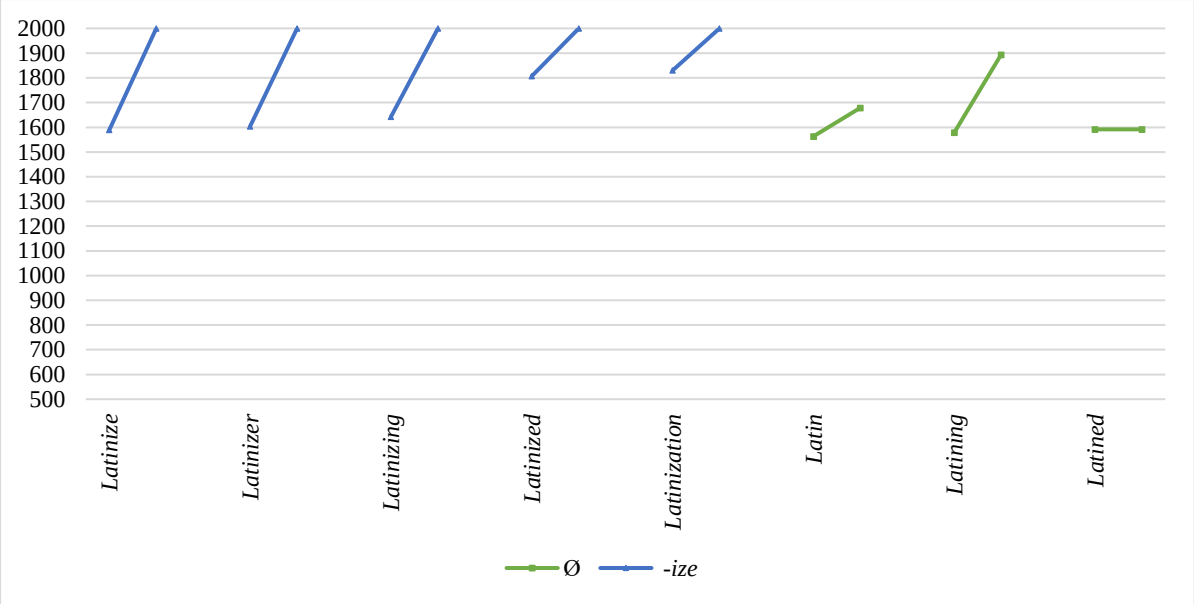


Figure 10: Timelines for the forms in Ø (in green) and in *-ize* (in blue) and their derivatives for the cluster *Latin*/*Latinize*

The cluster *Latin*/*Latinize* shows an opposite view of the paradigm in the resolution of competition to that observed in Figure 9. In this case, both verbs are first attested at the end of the 16th century and both have derivatives that date back to the same period. However, only those derived employing *-ize* remain in Present-Day English.

All in all, the clusters in Figures 9 and 10 reinforce the interaction between competition and paradigms, such as that the latter serve as additional data for supporting the resolution of the competition in favor of the one or the other affix. Furthermore, Figures 9 and 10 also seem to imply that the influence of the (sub)paradigms of two forms in competition may be twofold as further derivation may support the preference for an already existing form over a later attested competitor. Alternatively, when both competitors are first attested within the same period, further derivation may in some way guide the resolution of competition in favor of one of the forms. Again, these assumptions must be considered cautiously as the results may be no more than a consequence of the limitations of the lexicographic resources used.

Moreover, other factors may be at work here as well. It may be the case that, in some instances of competition at least, the *-ize* form did not succeed simply because it was used with a stylistic effect that did not prove permanent. This hypothesis is illustrated by the cluster *cuckold*/*cuckoldize*. In contrast to *cuckold* (1589–, OED2), *cuckoldize* is only attested once in the OED, in the 17th century with a label of "obsolete, rare":

(21) *Can dry Bones Live? or Skeletons produce The Vital Warmth of* ***Cuckoldizing*** *Juice?*
(N. Tate & Dryden, *2nd Pt. Absalom & Achitophel* 11, 1682, OED2)[20]

[20] The forms in the cluster *cukold*/*cuckoldize* were last updated in 1989 (OED2). Cases like this exemplify the limitations in the use of lexicographic data and the need in future research of exploring historical corpora for the centuries under scrutiny.

However, the lack of lexicographic record makes it difficult to decide whether forms such as *cuckoldize* were potential competitors for the zero-derived verb or just unsuccessful stylistic innovations. Similarly, the gaps between dates of attestation observed in forms such as *grandize*[21] also hinder the study of availability and competition. Explanations for the gaps found in attestations tend to be varied, because forms may have:

i) been in use, but failed to be recorded by the dictionary makers due to limitations in the lexicographic practice; or
ii) come to be restricted to certain domains (e.g. medicine, mathematics) or been used with a specific stylistic purpose; or
iii) been lost at some point in the history and later re-activated again (Bauer 2014) as potential competitors (as it may be the case of e.g. *grandify*/*grandize* above).

Although the issue of (un)availability is a complex one and further, methodologically varied research is needed to account for its complexity fully, it is clear that taking into account the information provided by the derivatives in the subparadigms created by the bases in competition may help us gain insights into the diachronic availability of forms such as *grandize*. However, assumptions based on the information provided by subparadigms need to be considered with caution. The lack of derivatives may evidence resolved competition, but in forms such as *revolution*^{V} (1805–, OED3) or *ghetto*^{V} (1936–, OED2), which are attested in the 19th century or in the course of the 20th century, respectively, the unavailability of derivatives may also be a consequence of their as yet short existence in the language and, therefore, of their availability to yield further derivation in the future.

6. Conclusion

This paper elaborates on, and partly confirms, the results obtained from an analysis in our previous research in two ways: from a descriptive, analytical and theoretical perspective, paradigms add additional evidence that supports the resolution of competition in favor of one of the suffixes under study and suggest that such resolution can occur following various profiles. Methodologically, the promising results show that the study of availability of (competing) forms is often hindered by lexicographic and corpus limitations and call for more research situated at the interface of synchrony and diachrony.

Regarding the interaction of derivational (sub)paradigms and competition, the results in this paper support those in the previous research in that the suffix *-ize* seems to be the preferred option for the formation of causative verbs. However, the results obtained also show that in seven of the 26 clusters where competition has been resolved, zero-derivation is preferred over the suffix *-ize*. In all of them, the former is attested earlier and does serve as the base for further derivation.

The inclusion of paradigms in the study of competition also suggests that there are cases where once one of the competing forms has developed derivatives, a synonym may be less likely to be coined (e.g. in the cluster *discipline*/*disciplinize*, the *-ize* verb is first attested after its zero-derived competitor once the latter has already attested derivatives). This is in some way reinforced by the opposite pattern found in clusters where *-ize* wins out over zero-

[21] See §4.3, examples (7)–(9).

derivation. In most cases, *-ize* is last attested later than the zero-derived form, but none or few derivatives with zero-derived bases are attested in the OED or recorded in the COCA. The comparison of both patterns suggests that the *-ize* form is preferred if there is no earlier form with already attested derivatives.

Therefore, the results support the assumption that the description of the patterns of resolved competition is enhanced by including derivational (sub)paradigms, but whether they are the cause or the consequence of such resolution needs further research. The analysis of the results obtained suggests that paradigms may have an effect on the competition between their base and other form(s). In detail, the subparadigms in the cluster *English*/*Englishize* illustrate how the derivatives of the zero-derived form may support its prevalence over a later attested competitor in *-ize*. In contrast, the subparadigms for the verbs *Latin*/*Latinize*, where both forms are first attested in the same period, apparently guide the resolution of the competition in favor of *-ize*.

Whether variation is variety-related is a question that has remained unanswered in this paper, although choices exemplified by *quiet* and *quieten* may be a consequence of this factor. Corpus data reflect that while *quiet* is apparently preferred in AmE, *quieten* is more common in BrE. This case clearly illustrates the necessity of considering inter-variety differences in future research as well as all the possible affixes that may enter in competition in a cluster (e.g. *quiet*/*quietize*/*quieten*).

From a more general perspective, the study of availability may also benefit from an analysis of the members of the derivational (sub)paradigm in question. Specifically, the availability of derivatives may provide information about the availability of their bases. This is especially relevant in those cases where there are gaps in lexicographic records. However, although the availability of the members of the subparadigms may suggest that their bases have remained available even if records are lacking for some centuries, this assumption needs to be taken with caution. The lack of derivatives and gaps in the dates of attestation do not necessarily mean that the forms are unavailable because they may be instances of renewed availability (Bauer 2014). In fact, the lack of attestations may mean just the opposite: verbs first attested during the 19th or 20th centuries may have been created with a restricted sense that may not be obvious from the records provided by the OED or may yet become competitors of available verbs. As both obsolete forms and neologisms are sometimes excluded from corpus record, identifying whether the type of verbs mentioned above belong to the first or second group needs supporting evidence.

In fact, the questions posed about the availability of forms highlight first, the problems of etymological and historical dictionaries whether they get updated (like the OED, which is rare) or not (which is typical), and second, the limitations of synchronic corpora. Specifically, the use of historical corpora could complement the information provided by the OED. By doing so, the patterns of competition observed so far could be further detailed and/or reformulated.

This study has further demonstrated complex interactions between derivational paradigms and competition in word-formation. Our results imply that derivational (sub)paradigms add extra data to the study of diachronic competition. In turn, a fuller description of competition allows constructing more detailed paradigms, and thus, makes a more complete picture of derivational morphology possible.

Acknowledgements

We would like to express our gratitude to the reviewers for their helpful comments and suggestions. This paper has been supported by the Spanish State Research Agency (SRA, Ministry of Economy and Enterprise) and European Regional Development Fund (ERDF) (Ref. FFI2017-89665-P) and by the European Regional Development Fund-Project *Creativity and Adaptability as Conditions of the Success of Europe in an Interrelated World* (No. CZ.02.1.01/0.0/0.0/16_019/0000734).

References

Aronoff, Mark. 2016. Competition and variation. (Paper presented at the *16th International Morphology Meeting*, Vienna, 18–21 February 2016.)

Bagasheva, Alexandra. 2017. Comparative semantic concepts in affixation. In Santana-Lario, Juan & Valera, Salvador (eds.), *Competing patterns in English affixation*, 33–66. Bern: Peter Lang.

Bauer, Laurie. 1997a. Evaluative morphology: A search for universals. *Studies in Language* 21(3). 533–575.

Bauer, Laurie. 1997b. Derivational paradigms. In Booij, Geert & van Marle, Jaap (eds.), *Yearbook of Morphology 1996*, 243–256. Dordrecht: Kluwer.

Bauer, Laurie. 2009. Competition in English word-formation. In van Kemenade, Ans & Los, Bettelou (eds.), *The handbook of the history of English*, 177–198. Malden, MA: Blackwell.

Bauer, Laurie. 2014. 'What is the plural of mouse?' and other unhelpful questions for morphologists. (Plenary lecture delivered at the *47th Annual Meeting of the Societas Linguistica Europaea*, Poznań, 11–14 September 2014.)

Bauer, Laurie & Nation, Paul. 1993. Word families. *International Journal of Lexicography* 6(4). 253–279.

Bauer, Laurie & Díaz-Negrillo, Ana & Valera, Salvador. 2010. Affixation vs. conversion: The resolution of conflicting patterns. In Rainer, Franz & Dressler, Wolfgang U. & Kastovsky, Dieter & Luschützky, Hans Christian (eds.), *Variation and change in morphology: Selected papers from the 13th International Morphology Meeting*, Vienna, February 2008, 15–32. Amsterdam: John Benjamins.

Bauer, Laurie & Lieber, Rochelle & Plag, Ingo. 2015. *The Oxford guide to English morphology*. 2nd edn. Oxford: Oxford University Press.

Bauer, Laurie & Valera, Salvador. 2015. Sense inheritance in English word-formation. In Bauer, Laurie & Körtvelyéssy, Lívia & Štekauer, Pavol (eds.), *Semantics of complex words* (Studies in Morphology 3), 67–84. Dordrecht: Springer.

Beecher, Henry. 2004. *Derivational paradigm in word formation* (http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.94.9071) (Accessed 2018-07-20.)

Blevins, James P. 2006. Word-based morphology. *Journal of Linguistics* 42(3). 531–573.

Bochner, Harry. 1993. *Simplicity in Generative Morphology*. New York, NY: De Gruyter Mouton.

Bonami, Olivier & Strnadová, Jana. 2018. Paradigm structure and predictability in derivational morphology. *Morphology*. (https://doi.org/10.1007/s11525-018-9322-6) (Accessed 2018-07-20.)

Booij, Geert. 1996. Inherent versus contextual inflection and the Split Morphology hypothesis. In Booij, Geert & van Marle, Jaap (eds.), *Yearbook of Morphology 1995*, 1–16. Dordrecht: Kluwer.

Booij, Geert. 2008. Paradigmatic morphology. In Fradin, Bernard (ed.), *La raison morphologique: Hommage à la mémoire de Danielle Corbin*, 29–38. Amsterdam: John Benjamins.

Booij, Geert & Lieber, Rochelle. 2004. On the paradigmatic nature of affixal semantics in English and Dutch. *Linguistics* 42(2). 327–357.

Chiba, Shoju. 2016. The lexical realization patterns and their relationship to productivity: The Finnish symbiotic derivation. (Paper presented at the *16th International Morphology Meeting*, Vienna, 18–21 February 2016.)

Davies, Mark. 2004–. *BYU-BNC* (Based on the *British National Corpus* by Oxford University Press). (http://corpus.byu.edu/bnc/) (Accessed 2018-07-20.)

Davies, Mark. 2008–. *The Corpus of Contemporary American English (COCA): 560 million words, 1990-present*. (http://corpus.byu.edu/coca/) (Accessed 2018-07-20.)

Davies, Mark. 2011. The Corpus of Contemporary American English as the first reliable monitor corpus of English. *Literary and Linguistic Computing* 25(4). 447–465.

Díaz-Negrillo, Ana. 2017. On the identification of competition in English derivational morphemes. The case of *-dom*, *-hood* and *-ship*. In Santana-Lario, Juan & Valera, Salvador (eds.), *Competing patterns in English affixation*, 119–162. Bern: Peter Lang.

Don, Jan. 2014. *Morphological theory and the morphology of English*. Edinburgh: Edinburgh University Press.

Dressler, Wolfgang U. 1989. Prototypical differences between inflection and derivation. *Zeitschrift für Phonetik, Sprachwissenschraft und Kommunikationsforschung* 42(1). 3–10.

Fernández-Alcaina, Cristina. 2017. Availability and unavailability in English word-formation. In Santana-Lario, Juan & Valera, Salvador (eds.), *Competing patterns in English affixation*, 163–206. Bern: Peter Lang.

Fradin, Bernard. 2016. What can we learn from doublets? (Paper presented at the *16th International Morphology Meeting*, Vienna, 18–21 February 2016.)

Horecký, Ján & Buzássyová, Klára & Bosák, Ján et al. 1989. *Dynamika slovnej zásoby súčasnej slovenčiny*. Bratislava: Veda.

Kaunisto, Mark. 2007. *Variation and change in the lexicon: A corpus-based analysis of adjectives in English ending in* -ic *and* -ical. Amsterdam: Rodopi.

Kaunisto, Mark. 2009. The rivalry between English adjectives ending in *-ive* and *-ory*. In McConchie, Roderick W. & Honkapohja, Alpo & Tyrkkö, Jukka (eds.), *Selected Proceedings of the 2008 Symposium on New Approaches in English Historical Lexis (HEL-LEX 2)*, 74–87. Somerville, MA: Cascadilla Press.

Kjellmer, Göran. 2001. Why *weaken* but not **strongen*? On deadjectival verbs. *English Studies*. 82(2). 154–171.

Lara-Clares, Cristina. 2017. Competition in Present Day English nominalization by zero-affixation vs. *-ation*. In Santana-Lario, Juan & Valera, Salvador (eds.), *Competing patterns in English affixation*, 207–244. Bern: Peter Lang.

Lara-Clares, Cristina. 2018. Nominal competition in Present-Day English affixation: zero-affixation vs. *-ness* with the meaning STATIVE. (Paper presented at the conference *Word-Formation Theories III & Typology and Universals in Word-Formation IV*, Košice, Slovakia, 27–30 June 2018.)

Plag, Ingo. 1999. *Morphological productivity: Structural constraints in English derivation*. Berlin: De Gruyter Mouton.

Plank, Frans. 1994. Inflection and derivation. In Asher, Ron E. (ed.), *Encyclopedia of Language and Linguistics* 3. 1671–1678. Oxford: Pergamon.

Pounder, Amanda. 2000. *Process and paradigms in word-formation morphology* (Trends in Linguistics 131). Berlin: De Gruyter Mouton.

Projekt Monika. 2018. *Cross-linguistic research into derivational networks.* Pavol Jozef Šafárik University (Košice, Slovakia), University St. Kliment Ohridsky (Sofia, Bulgaria), University of Granada (Spain). (https://www.ugr.es/~svalera/Monika/index.html) (Accessed 2018-08-07.)

Proffitt, Michael. 2018. *The Oxford English Dictionary*. (http://www.oed.com) (Accessed 2018-07-20.)

Roché, Michel. 2011. Quelle morphologie? In Roché, Michel & Boyé, Gilles & Hathout, Nabil & Lignon, Stéphanie & Plénat, Marc (eds.), *Des unités morphologiques au lexique, langues et syntaxe*, 15–39. Plymouth: Hermes Science Publishing.

Štekauer, Pavol. 2014. Derivational paradigms. In Lieber, Rochelle & Štekauer, Pavol (eds.), *The Oxford handbook of derivational morphology*, 354–369. Oxford: Oxford University Press.

Štekauer, Pavol. 2017. Competition in natural languages. In Santana-Lario, Juan & Valera, Salvador (eds.), *Competing patterns in English affixation*, 15–32. Bern: Peter Lang.

Stump, Gregory. 2001. *Inflectional morphology*. Cambridge: Cambridge University Press.

Thornton, Anne Marie. 2011. Overabundance (multiple forms realizing the same cell): A non-canonical phenomenon in Italian verb morphology. In Maiden, Martin & Smith, John Charles & Goldbach, Maria & Hinzelin, Marc-Olivier (eds.), *Morphological autonomy: Perspectives from Romance inflectional morphology*, 358–381. Oxford: Oxford University Press.

Tichý, Ondřej. 2018. Lexical obsolescence 1700–2000. In Kopaczyk, Joanna & Tyrkkö, Jukka (eds.), *Applications of pattern-driven methods in Corpus Linguistics*, 81–103. Amsterdam: John Benjamins.

van Marle, Jaap. 1985. *On the paradigmatic dimension of morphological creativity*. Dordrecht: Foris.

Cristina Fernández-Alcaina
c/o Salvador Valera Hernández
Departamento de Filologías Inglesa y Alemana
Facultad de Filosofía y Letras
Campus de Cartuja
18071 Granada
España - Spain
cristinaalcaina@correo.ugr.es

Jan Čermák
Department of English Language and ELT Methodology
Charles University, Faculty of Arts
nám. Jana Palacha 1/2
116 38 Prague 1
Czech Republic
jan.cermak@ff.cuni.cz

In SKASE Journal of Theoretical Linguistics [online]. 2018, vol. 15, no. 3 [cit. 2018-13-12]. Available on web page http://www.skase.sk/Volumes/JTL38/pdf_doc/04.pdf. ISSN 1336-782X.

Evolutionary trajectories of word-formation processes in the Old High German language

Larysa Shchyhlo, Sumy State University, Ukraine

The purpose of this article is to model the evolution of the word-formation system in the ancient stage of the diachrony of the High German language in a multidisciplinary focus, especially from the standpoint of a special direction of the synergetic paradigm – evolutionary linguistic synergetics. The expediency of using the ideas of synergetics as a transdisciplinary methodology to the development of the linguistic system in general and its subsystems (in our case, word-formation) in particular is determined and substantiated; the mechanisms and conditions for the self-organization of the German word-formation system are analyzed; the degree of extralinguistic factors influence on the self-development of the German word-formation system is established.

Keywords: *synergetics, linguistic synergetics, evolutionary linguistic synergetics, word-formation, the Old High German language.*

1. Introduction

Synthesis of different disciplines takes place in modern science within synergetics – interdisciplinary (and broader – transdisciplinary) scientific direction of the study of complex, open, dynamic, self-organizing, non-linearly evolving systems of different ontologies. Interdisciplinarity of synergetics consists in cooperation of methods and principles of research of complex systems developed in the framework of various scientific disciplines, which are united by the object of research – complex systems, studying stages of their self-organization, revealing general patterns of their functioning, establishing general principles of their evolution.

Synergetics is not a fundamentally new paradigm of scientific knowledge. We are in agreement with scientists who are considering synergetics as

> a special area of system research, as the next stage in the theory of systems development, aiming at disclosure of changability's nature, revealing the mechanisms of system transition to a qualitatively new state,'unpacking' the concept of 'dialectical leap' in the evolutionary development of a complex system (Dombrovan 2013: 68).

Nowdays we can say about the successful cooperation of synergetics and linguistics, having been embodied in a new integrative interdisciplinary (and broadly – transdisciplinary) direction in the science of language – linguistic synergetics or linguosynergetics.

In modern linguistics, the following directions are clearly defined: (1) psycholinguistic approach or psycholinguistics (Haken 2000; Moiseeva 2007) integrates synergetics and the theory of speech activity; (2) functional linguosynergetics, proposed by Ponomarenko, combines system-functional and synergetic perspectives of language / speech / discourse (Ponomarenko 2010); (3) linguistic synergetics (Keller 1994; Piotrovskij 2006) is oriented, first of all, on quantitative data of synergetic mechanisms of language / speech and their statistical processing; (4) diachronic linguosynergytics (Dombrovan 2013; 2018) is focused on modeling of a particular language system evolution with prediction of variants of its subsequent

changes, depending on multi-vector bifurcations and the diversity of potential attractors; (5) contradiction-synergistic approach (contradiction-synergetic linguistics) (Myshkina 1999), which introduces the energy dimension into the linguosynergic perspective of the language; (6) the linguosynergic approach in the genre studies is aimed to reveal the mechanisms of various language genres development (Pikhtovnikova 1999); (7) using the gestalt-synergetic approach Ljudmila Kushnina develops the applied aspect of linguistic synergetics in order to create a synergetic semantic translation model – the translation space, which is considered as an open, dynamic, developing system (Kushnina 2004); (8) the semiotic-synergetic approach integrates synergetic and semiotic principles, which in particular allowed to investigate the postmodern fiction discourse as a self-organizing semiotic space (Oliz'ko 2009).

Substantiating theoretical and conceptual foundations of sunergetic approach to the study of language system, Ukrainian researcher Tetiana Dombrovan schematically depicts the conceptual basis of linguistic synergetics in the form of a radial diagram containing three basic components that represent the main methodological principles of the three components of linguosynergetics: philosophy, linguistics, synergetics (Figure 1).

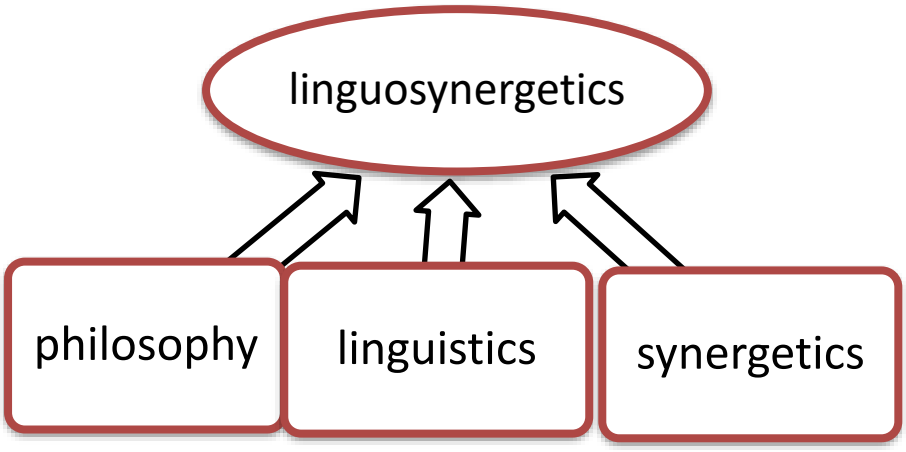


Figure 1: Conceptual and methodological platform of linguistic synergetics
(Dombrovan 2013: 70)

The diagram reflects the integrative approach to cognitive activity within the framework of linguistic synergetics, the combination of methodologies for the study of humanitarian and natural sciences (Dombrovan 2013: 70).

So, as we see from the scheme, the basic foundation of linguistic synergetics is philosophy. The relevance of the synergistic approach to the philosophical understanding of language is caused by its nature, its multifunctionality and the properties as an open, evolving system.

One of the constituent structures of the conceptual and methodological platform of linguistic synergetics is linguistics itself, because there is a need to develop the language concept as a complex hierarchically organized megasystem, the components of which are coherently linked.

The excursus in the history of linguistics shows that ideas similar to synergetic were proclaimed already in the first quarter of the 19th century in the works of the German scientist Vilhelm von Humboldt. This is the idea of self-regulation of the language: “Languages did not arise out of arbitrariness or arrangement, but they came out of the hiding plases of human nature

and are self-regulating and developing sound elements" (von Humboldt 1984: 324); as well as the idea of a spontaneous self-organization of language: "Whatever the natural assumption of gradual formation of languages, they could only arise immediately. Man is a man only through language; but in order to create a language, it must already be a man" (von Humboldt 1964: 80). And further:

> It is impossible to imagine a language as something previously given, because in this case it is completely incomprehensible how a man could understand this reality and force it to serve itself. Language, of course, arises from a man and, of course, little by little, but so that the body of language does not lie dead in the darkness of the soul, but is a law determing the mental function of a person, therefore the first word already defines and assumes the existence of the whole language. If this unique ability of a man to try to compare with something else, then you have to remember about the animal instinct and call the language the intellectual instinct of the mind [ibid.].

It should be emphasized that Vilhelm von Humboldt gives an opinion about the evolutionary-dynamic approach to language, considering language not so much as a product of activity (ergon), but as an activity itself (energeia) (von Humboldt 1984: 7).

An appeal to synergetics in the process of studying the language system is fully justified, because it is caused by the language properties as a dynamic, complex organizing synergetic system.

So, language as a complex self-organizing system is the object of transdisciplinary study – linguistic synergetics or linguosynergetics. Linguosynergetics is a transdisciplinary network structure of science, where linguistic knowledge varies depending on the chosen scientific perspective, and other relevant knowledge is united on the basis of evolutionary-synergetic principles (Bronnik 2012: 9). Evolutionary linguistic synergetics or evolutionary linguosynergetics, basing on the principle of universal evolutionism is one of the methodological approaches within linguosynergetics for studing the phenomenon of spontaneous occurrence – the self-organization of structures in various diachronic stages of the linguistic system development. The main task of the proposed theory is to reveal the internal and external laws of the language system evolution (Shchyhlo 2017: 87).

The development of the linguistic theory in general and the theory of word-formation in particular, the experience gained in the field of word-formation analysis of separate stages of the German language development allowed us to recreate the development of the German word-formation system on the basis of the evolutionary linguosynergetics. Such interdisciplinary orientation of the research contributes to a comprehensive, thorough description of the factors and mechanisms of changes in the word-formation macrosystem at a qualitatively new level.

2. Research methods

To ensure the reliability of the results and conclusions the following research methods are used: observation of investigated phenomena and processes, analysis and descriptions of their changes and functioning, induction (for summarizing the results of the observation), deduction (for checking of general statements on a particular language material). The method of component analysis is used in this paper for identification and definition of the semantic structure of derivatives and complicated words.

The modeling method is applied for reproducing the history dynamics of the German word-formation system.

3. Presentation of the main research material

The diachrony of the Germanic languages covers more than a thousand years. They passed this long way in various linguistic and extra-linguistic conditions. There were some general tendencies and a unified direction of the language system formation, but each language was characterized by its own specific features and was unique. The development and establishment of the Germanic languages continues in the pre-literary period. The Great Migration of peoples in the 4^{th} – 5^{th} centuries had a significant influence on this process, which resulted in the formation of the German nation from the West Germans with its own language – German. In the VI century begins the second movement of consonants, which was manifested primarily in the South Germanic dialects. All changes in the German language (including the word-formation system) in the pre-literary period are prone to the influence of Vulgar Latin and Gallo-Romance. The specificity of the German language evolution in this period was its development in conditions of dialect variety. As generally known, the Old High German language is the language of Germanic tribes, fixed in written sources, dating from about 750 year to 1050. The mentioned language was represented by a group of East Germanic dialects of the tribes of Franks, Alemans and Bavarians. They became the base for the future development of the High German language.

3.*1 Substantive derivation in the Old High German language*

Derivatives with an agent suffix -âri. The suffix *-âri* comes from the Latin *-arius* and was borrowed from the Latin language together with the words into which it was included (for example, *molinarius* 'miller'). Obviously, this suffix served, for the most part, to create new words from the stems of nouns, which is confirmed by the materials of one of the oldest Germanic languages – Gothic. Compare, for example, *bôkareis* (der Schriftgelehrte), 'a scribe, one who is understood in the scripture' from *bôka* 'letter'; *wullareis* is 'a specialist in making wool' from *wulla.* In the Old High German language there were also words derived from the stems of nouns with the help of this suffix, for example: *ambahtâri* 'servant' from the *ambaht* 'service'; *fâtâri* 'tempter' from *fâra* 'temptation, danger'; *gartâri*, *gartinâri* 'gardener' from *garto* 'garden'; *lêrâri* 'teacher' from *lêra* 'teaching'; *mâdâri* 'harvester' from *mâd* 'harvests'; *scâhâri* 'robber' from *scâh* 'robbery'. Among the nouns there were nomina agentis) derived from the corresponding verbs, for example: *suonâri* 'judge' – *suona* 'court' – *suonen* 'judge'; *lêrâri* 'teacher' – *lêrà* 'teaching' – *lêren* 'teach', etc. This circumstance contributed to the formation of new nouns directly from the verbal bases. Thus, the meaning of the suffix itself was redefined; it began to express the creator of action, and not just an attitude to action. In the Old High German language a large number of nouns with this suffix are derived from the verb: *fiskâri* 'fishing' from *fiskôn* 'fish', *jagâri* 'hunter' from *jagôn* 'hunt', *heilâri* 'healer' from *heilen* 'heal', *helfâri* 'helper' from *helfan* 'help'. The new word-formation type, which was formed as a result of borrowing of the word-formation component was so productive that it began to supplant the old suffix *-n*, having been served then for the same purposes. Compare, for example, the Old High German – *helfo* 'helper' – *helfâri* 'one who helps', *nemo* 'taker' – *nemâri* 'one who takes', as well as the Middle High German – *geber* is 'one who gives'

widersacher ‘an opponent’ with the Old High German – *gebo* ‘one who gives’, *widersacho* ‘an opponent’. From the old formations with *-n* in the Modern language only a few words have remained, for example, *der Bote* ‘missioner’ (the Old High German *boto* ‘missioner’). This word-formation type retained its productivity nowadays, having for this the preconditions in objective reality. In the Modern German there are nouns in the meaning of nomina agentis almost derived from verbal bases (rarely from nominal ones, for example, *Künstler* ‘artist’, *Gärnter* ‘gardener’).

Often, the suffix *-âri* prior was preceded by any more ancient word-formation component; for example, such derivatives as *lugina* ‘lie’ – *luginâri* ‘liar’, *satal* ‘saddle’ – *satalâri* ‘saddler’, which contributed to the time of the process of redefinition of the bases and the isolation of modern suffixes *-ler* and *-ner*. Already in the Old High German period as a result of the stem redistribution, there is a suffix *-nari*, compare *gartinâri* ‘gardener’ from *garto* ‘garden’ (stem on *-n*) and *sculdinâri* ‘debtor’ from *sculda* ‘debt’. As the analysis shows, there are no examples with the suffix *-lâri* in the Old High German language.

In addition to the suffix *-âri* in the Old High German language for the formation of nomina agentis were used the German suffix *-(i)l*, for example, *biril* ‘carrier’ from *beran* ‘worn’, *wibil* ‘beetle’ from *weban* ‘weave’, *wachtil* ‘guard’ from *wachten* ‘guard’. However, as the analysis shows, the number of such nouns was small and gradually diminished in the process of historical development, ousting the formations on *-er*.

Derivatives with the instrumental suffix -il. The instrumental suffix *-il* was the most widely used in the Old High German language formations in the meaning of tools: *sluzzil* ‘key’ from *sliezen* ‘to close’, *zugil* ‘leads’ from *ziehen* ‘to pull’, *slegil* ‘felled’ from *slahan* ‘beat, hit’, *leffil* ‘spoon’ from *laggan* ‘draw’, *fezzil* ‘bundles’ from *fazzon* ‘grab, skid’. In some cases, the instrumental suffix *-il* served for forming of nomina agentis: compare the Old High German *biril* ‘bearer’ with *berero* from *beran* ‘bear’, *tregil* ‘carrier’ from *tragan* ‘carry’.

Derivatives with the suffix -unga. This suffix was widespread in all Germanic languages in the patronymic meaning: compare *Amalungi* (in Goths), *Carolingi* (in Francs). In the Old High German language this suffix was used precisely for the formation of abstract nouns of the female genus from verbal bases (renowned formations are rarely found). So, in poetry in Otfrid (9^{th} century) it occurs in several cases as verbal nouns with a subject meaning: *manunga* ‘reminder’, *samanunga* ‘collection’. The development of abstract nouns on *-unga* are most closely observed in the clerical prose of the 7^{th} – 10^{th} centuries in translations from Latin: compare *wirkunga* ‘action’ (Latin *operatio*), *zeigunga* ‘definition’ (latin *determination*), *sceidunga* ‘division’ (latin *division*), *heilagunga* ‘sanctification’ (latin *sanctification*). However, it should be noted that the most “nutritious” source for abstract units of this kind was a scientific prose: *betunga* ‘request’ from *betôn* ‘to ask’, *beitunga* ‘expectation’ from *beitôn* ‘expect’, *korunga* ‘test’ from *korôn* ‘to test’; *scouwunga* consideration from *scouwôn* ‘to consider’; *bezzirunga* ‘improvement’ from *bezzirôn* ‘to improve’; wehsilunga ‘change’ from *wehsilôn* ‘to change’; *zwîfalunga* ‘doubt’ *zwîfalôn* ‘to doubt’.

This type of word-formation is still productive, having been used for many centuries to form a large number of nouns.

Derivatives with the suffix -ida. We find the suffix *-ida* in other Germanic languages, for example: Gothic – *iþa* (*háuhiþa* ‘height’), Modern English *length, breadth, warmth, strength,* etc. Interestingly, that this suffix, which has become unproductive in Modern German and Modern English, is widespread in Modern Dutch, compare *diepte, stilte, warmte* and others. Already in the Old High German period derivatives with this suffix formed from the basics of adjectives and verbal bases were few, for example: *breitida* ‘width’, *ewida* ‘eternity’,

frewida ‘joy’, *reinida* ‘purity’, *salida* ‘happiness’. In manuscripts that can be considered more consistent with live use (Otfrid, Tatian, Notker), their frequency of use is low. As it can be assumed, these derivatives were characteristic of certain language styles (religious and clerical prose). In Modern German there are a few remnants of ancient derivative nouns with suffix *-ida*: *Begierde* ‘desire’, *Freude* ‘joy’, *Zierde* ‘adornment’. This suffix was superseded by the relatively ‘young’ suffixes *-heit* and *-ung*.

Derivatives with the suffix -heit. The suffix *-heit* belongs to ‘young’ suffixes that have evolved from independent words. In all Germanic languages we find the corresponding independent word meaning ‘person’, ‘position’, ‘mode’. Compare, for example, Gothic *haidus* ‘kind, way’; Old English *hád* ‘state, genus, property, species’, Old Scandinavian *heiðr* ‘honour’; Old High German *heit* ‘person, gender, rank, condition’. As a means of generating abstract nouns, this suffix evolves from an independent word only in West Germanic languages. Among the words formed from the basis of nouns are those whose first part denotes any person, for example: *deganheit* ‘valor’, *diubheit* ‘theft’, *kindheit* ‘childhood’, *narraheit* ‘folly’, and so on.

However, the most prevalent in the Old High German language were derivative nouns with the suffix *-heit*, formed from the basics of adjectives. In combination with the basics of adjectives *-heit* becomes the most common means of the formation of abstract nouns and competes with older formations with the suffix *-î*, for example: *armheit* ‘poverty’, *bitterheit* ‘bitterness’, *bôsheit* ‘anger’, *frîheit* ‘freedom’, *kuonheit* ‘courage’, *tumbheit* ‘foolishness’, *wîsheit* ‘wisdom’. The valency specificity of the suffix *-heit* is to combine its potency with the derivative bases of adjectives, in particular, with adjectives that have the suffixes *-ag, -îg* due to this in the Middle High German period a new morpheme – the suffix *-keit* is singled out.

Derivatives with the suffixes -scaf, -scaft. As the analysis testifies, the ‘young’suffix *-scaf* is used in all German languages not only as an independent word, but also as derived nouns from this basis, as well Gothic *gaskafts* ‘creation’; Old High German *giscaft* ‘creature, form, state, fate’; Old English *gesceaft* ‘creation, talent’, occurs several times in texts in the meaning of an image, property, as well as verbs *scephen* and *scafan* ‘create’, adjective *unscaf* ‘unusual, abnormal’, *scafalos* ‘shapeless, ugly’. The original meaning of the suffix *-scaf* was, obviously, ‘an image, a property’ (compare with Modern German *Beschaffenheit* ‘feature’, ‘characteristic’). As to the form of this suffix, it should be noted that up to the IX century it acts in the form of *-scaf*. From the IX century the form *-scaft* begins to compete with it, which is genetically derived from the ancient abstract formation with the suffix *-ti* – Old High German (*gi*) *scaft* ‘creation, a creature’ that eventually replaced the first one.

Like *-heit, -scaft* was joined the basics of nouns and adjectives, but the number of derivatives of the last kind is negligible, for example: *fîantscaf* ‘enmity’, *holdscaf* ‘affection, friendliness’. The differential sign from *-heit* in this case is that derivatives with *-heit* denote usually a specificity or property, derivatives of *-scaft* point to activity, state, behavior, or attitude. A significant number of nouns created using the suffix *-scaf* is characterized by the value of collectivity, for example: *bruoderscaf* ‘brotherhood’, *geselliscaf* ‘society’, *kunniscaf* ‘generation’.

Derivatives with the suffix -tuom. The ‘young’ suffix *-tuom* as a corresponding word was available in all German languages and was encountered in self-use, for example: Gothic *dôms* ‘judgment’; Old Scandinavian *dómr* ‘court, decision’; Old English *dōm*; Old High

German *tuom* ‘custom, power, domination’, compare also with Old Indian dhāman ‘place, motherland, power, honor, court’, etc.

A set of derivative meaning with the suffix *-tuom* is different in its diversity. Derivative nouns with the suffix *-tuom* can denote the state, custom, often they are close to the derivatives of *-heit* or *-scaft*. Usually the suffix *-tuom* is valent to the basics of nouns, very rarely from the adjectives, for example: *alttuom* ‘age’, *arzetuom* ‘art of medicine’, *diornutuôm* ‘innocence’, *heidantuom* ‘paganism’, *kaisartuom* ‘kingdom’, *rîhtuom* ‘wealth’.

3.2 *Adjective derivation in the Old High German language*

Derivatives with suffix -isc. This is a very old suffix, which occurs in other Indo-European languages, compare Greek ιχος *(-ikós*), Latin *-ikus.* The suffix indicates origin, as well as belonging to something. With this suffix from the names of living und unliving objects formed many adjectives, compare *diutisc* ‘folk’ from *diot* ‘people’, *himilisc* ‘heaven’ from *himil* ‘sky’, *irdisc* ‘earth’ from *erda* ‘earth’, *kindisc* ‘young, gentle’ from *kind* ‘child’, and so on. This suffix is usually used to create adjectives from the names of areas or peoples: *Rôm* ‘Rome’ – rômisc, *rumisc* ‘Roman’; *Franco* ‘Franc’ – *francisc* ‘Franc’; *Walah* Celtic – *walahisc* ‘Celtic’, etc. It is very stable and therefore has survived to this day, expanding the sphere of its application, compare, for example, the formations from own names: *bayerisch* ‘Bavarian’, *goethisch* ‘belonging to Goethe’s pen’.

Derivatives with the suffixes -ag, -îg (modern *-ig*). In the Old High German language this suffix was one of the most used. Adjectives derived from the noun stems with the help of this suffix meant either any connection with the phenomenon, which called the noun or pointed to a property characteristic of the subject or concept, for example: *frost – frostag* ‘frosty’; *lust – lustag* ‘joyful, cheerful’; *muot – muotag* ‘courageous’. A small number of adjectives is also formed from the basics of other adjectives, for example, *reht – rihtîg* ‘right’.

Among the adjectives formed from the nouns of action there were those, in which the link to the verb was clearly traceable. This feature served as the basis for the formation of adjectives with this suffix directly from the verbs, for example: *birîg* ‘fruitive’ and *beran* ‘to bear, to give birth’; *firsûmîg* ‘careless’ and *firsûmen* ‘to miss’; *slâfag* ‘sleepy’ – *slâf* ‘dream’ – *slâfan* ‘to sleep’. Some adjectives could be formed from particles and adverbs, for example: *hiutu* ‘today’ – *hiutìg* ‘today’; *ofto* ‘often’ – *oftîg* ‘frequent’.

The suffix *-îg* is especially often encountered in compoud derivative words (Zusammenbildungen), that is, in the method of creating new words, which combines compounding of the stems and derivation, for example: *filo-spräch-îg* ‘boastful’*; -jâr-îg in zwî-, drî-, fimfjârîg* ‘two-three-five year old’, *vridomachîg* ‘peaceable’. This suffix is characterized by persistence and eventually displaces the suffixes *-oht, -aht* earlier equivalent in use.

Derivatives with the suffix -lîch. Like nouns, adjectives in the Old High German language form a new system of suffixes with abstract meaning through the formation of words with primordially independent words. So, the suffix *-lîch* comes from Gothic *leiks*, Old High German *lîh* with the meaning ‘body’ (Leib, Körper). Along with the noun *-lîh* were: a homonymous adjective with a meaning ‘similar, equal, suitable’, verbs *lîchên* ‘like’, *lîchîson* ‘to compare’ and a noun with an abstract meaning *analîchî* ‘similarity’. Compare also Gothic *galeiks*, Old High German *gilîch* ‘equal, identical’.

The question of which part of the language was the second component of compound words with *-lîch* a noun or adjective is controversial. Without a doubt, we can only say that these were compound words such as Bahuvrihi, for example, Gothic *waira-leiks* courageous is

interpreted as ‘the one that has a man's behavior’. From this meaning thereafter follows a more abstract meaning of similarity in general; compare derivatives such as *friuntlîch* ‘friendly’, *gastlîch* ‘hospitable’, *wiblîch* ‘feminine’, or relevance to what is called the basis from which the adjective is formed, for example: *fridulîch* ‘peaceful’; *scantelîch* ‘disgraceful’. There are also derivatives from the basics of adjectives, for example: *langlîh* ‘long’, *suozlîh* ‘sweet’. Simple adjectives and derivatives from them with the suffix *-lîh* at that time correlated with each other, quite possibly as a Modern German *gut* ‘good’, on the one hand, and *von guter Art* ‘of good kind’ or *gutartig* ‘good-natured’on the other, that is, they differed stylistically. Essentially their meaning was identical, but in the adjective with the suffix *-lîh* there is a tendency to use with abstract nouns. Subsequently, language carries a clear differentiation in meaning and in use (for example, modern *arm* ‘poor’– *ärmlich* ‘poorly*’, alt* ‘old’ – *ältlich* ‘elderly’, etc.). The correlation between some derivative adjectives with both a noun and a verb (*klagalîh* ‘pity’, *klaga* ‘pity’, *klagôn* ‘to pity’) contributes to the emergence of subsequent periods of diachrony of German derivatives directly from the basics of verbs, for example: *begreiflich* ‘understandable’ – *begreifen* ‘understand’**,** *erfreulich* ‘joyful’ – *erfreuen* ‘delight’, etc. (Filicheva 2003: 121).

Derivatives with suffix -haft. The suffix *-haft* (Gothic adjective *hafts*, Old High German *haft* extended form *haftig*) comes from the verb Gothic *hafjan*, Old High German *heffen* and is used in the meaning of ‘the one who has’, ‘the one who owns’. Compare the Old High German language *minnihaft* ‘minimal’; the New High German language *dauerhaft* ‘lasting’ expanded form *teilhaftig* ‘partial’, *wahrhaftig* ‘truthful’, etc. Along with this, the suffix in the Old High German language and in the Middle High German language was also a homonymous adjective with the meaning ‘bound, connected’ and the noun *haft* ‘captured, arrested’; compare, for example, Latin *captus* from *capio*. Perhaps the initial meaning of the suffix was ‘encumbered, equipped with something’, for example *namahaft* ‘famous, known’, *scamahaft* ‘shy’, *sigihaft* ‘victorious’. Some derivative adjectives correlate both with nouns and with verbs, for example, *lôbhaft* ‘praiseworthy’ – *lôb* ‘praise’ – *lôbôn* ‘praise’, *mezhaft* ‘shy, moderate’ – *mez* ‘measure’ – *mezzan* ‘measure’ that later serve to form adjectives directly from verbal bases, compare, for example, modern *schmeichelhaft* ‘flattering’ – schmeicheln ‘flatter’, *zaghaft* ‘timid’ – zagen ‘hesitate’, etc. The investigated suffix not only survived to this day, but is also quite productive, for example: *fehlerhaft* ‘faulty’, *krankhaft* diseased’, *naschhaft* ‘greedy’.

Derivatives with the suffix -bâri (expanded form *-bârig*). The suffix *-bâri* comes from the verb *beran* with the meaning of ‘bearer’ and, as you can suppose, functioned as an independent word, compare Old High German *unbâri* ‘fruitless’ (literally, ‘one that does not bear’), *bârig* ‘fruitful’. In compound words, this suffix was used in the meaning of ‘bearer’, ‘causing’, for example, Old High German *fluohbâri* is ‘hated’ (the one bearing the curse), *trostbâri* ‘comforting’ (the one that carries the fun), *dankbâri* ‘thankful’ (literally, the one that is thankful). Some derivatives correlate with the verb and noun, for example, *dankbâri* ‘thankful’ – *dank* ‘thank’*, dankôn* ‘to thank’. The group of verbs is most productive in the New High German period, compare modern: *ausführbar* ‘executable’ – *ausführen* ‘executre’, *entzündbar* ‘inflammable’ – *entzünden* ‘inflame’, *trinkbar* ‘drinkable’ – *trinken* ‘drink’, etc.

Derivatives with suffix -sam. The suffix *-sam* corresponds to the same pronoun as the adverb, *sam*, *samo*, as well as, for example: Gothic *same*, Old Scandinavian *samr*, Modern English *same*. Initially, the derivatives indicated conformity, similar to that expressed on the basis, and subsequently becomes a means of indicating the characteristic properties, abilities, inclinations, for example, *arbeitsam*. Some derivative adjectives correlate both with the noun

and with the verb, for example: *heilsam* ‘healing’ – *heil* ‘healing’ – *heilen* ‘heal’; *sorgsam* ‘caring’ – *sorga* ‘care’ – *sorgen* ‘to take care (care)’, etc. In Modern German there is a large number of verbs that denote the ability to act, called the basis (compare modern: *folgsam* ‘obedient’, *schweigsam* ‘taciturn’).

3.3 *Verbal derivation in the Old High German language*

Prefixation is a word-formation method most characteristic for the system of verb, while suffixation due to its classification function – a noun. According to the origin ancient verbal prefixes *bi-*, *gi-*; *int-*, *ant-*; *ar-*, *ur-*; *far-*, *fir-*; *zi (r)-*, *za (r)-*, *zur-* come from prepositive verbal adverbs, for the most part, locative semantics. The prefixes *bi-*, *ir-*, *ar-*, *ur-* and also *far-*, *fir-* in the Old High German language still correspond to prepositions that have evolved also from the corresponding adverb, for example: *bî-* by, with, around; *ur- from*; *furi* and *vora* in front of. The rest of the prefixes in the Old High German language had no correspondence in the form of independent words.

In the Old High German language these prefixes were used, mostly, to modify the meaning of verbs already available in the language, for example: *bintan* (*pintan*) ‘bind, skid’, *firbintan* ‘to join’, *far-bioton* (*firbiotan*) ‘prohibit, disallow’, and so on. In the process of further development of the German language, they acquire the function of ‘verbalization of nouns’ and become one of the most important means of enriching the vocabulary of the language with verbal innovations, stems for which not only verbs but also nouns served, for example, *Anspruch* ‘claim’ – *beanspruchen* ‘to claim’, *Gift* ‘poison’– *entgiften* ‘detoxify’, etc. (Filicheva 2003:127).

3.4 *Composite derivation as the most ancient productive word-formation process*

The ability to combine with each other the stems already existing in the language for the expression of any new notion was characteristic for the most ancient Indo-European languages from ancient times. However, one should say that in the Indo-European languages this ability developed in different ways. Thus, in Latin the number of complex words was insignificant, while in the Germanic languages, in particular in the German language, the vocabulary is enriched mainly by composition.

The oldest way of composition is considered to be the stem composition, in which the first part uses a pure basis without indexes of a case and a number. A similar type of composition is rooted in the distant past, when the flexion was not yet developed and when the simple arrangement of the stems served as a means of their combination in the sentence. These archaic combinations finally turned into compound words with the appearance of flexion, because their feature was the lack of flexion in the first component. This type of word-formation, called by Jacob Grimm *eigentliche Composita* (complete composites), could exist in pure form only when there were clearly different types of stems, which made it possible, without much effort to distinguish the necessary basis of each noun.

Complex combinations are widely represented in the Old High German language, for example: *erdrîhhi* ‘kingdom on earth’, *rouchvaz* ‘thurible’. In addition, in the Old High German language another kind of composition, which J. Grimm gave the name *uneigentliche Composita* (incomplete composites) was also presented. Complex words in this case arose originally from the combination of such words, between which the syntactic links found their expression in the case endings. Compare, for example: *donaræstag* ‘Thursday’, *Frankonofurt*

'Frankfurt'. The development of this type of composition took place along with the restructuring of the nominal system declination. In addition, the former flexion was reinterpreted as a connecting element (*Sonnenstrahl* 'sunray'). This contributed to the emergence of many formations by analogy. The last type of composition prevails in the Modern German language, which was in the Old High German language at the stage of formation.

So, it should be noted that the degree of the word-formation system development in the Old High German period is evidenced by the fact that the word-formation system of the VIII–XI centuries practically did not have its own word-formation models and word-formation means and as a result of the fact that the Old High German language was greatly influenced by the Greek and Latin languages, borrowed word-formative elements. Thus, the most word-formation models and means of this period are tracing calks from Latin or Greek, which did not allow to develop properly the German word-formation elements.

4. Conclusions and perspectives

Thus, the description of word-formation processes in the Old High German language from the standpoint of evolutionary linguistic synergetics shows that the word-formation system manifests the property of fractality. However, it should be emphasized that the fractality in the word-formation we see

> not in the use of certain suffixes or prefixes for the formation of new words, nor even in the simple coexistence of stems. These are all – mechanisms, models, schemes, 'trajectories' that contribute to the formation of new units. Fractality in word-formation is revealed in 'unpacking' of the word, deepening in the history of its appearance becomes clear that it is complicated and includes as components sometime independent words transformed as a result of interaction with other language components. It turns out that the inner limits of the word (the so-called morphemic seams) are mobile, dynamic and that any increase in the scale of consideration opens up new secrets: auxiliary inventory in the form of prefixes, suffixes and connecting elements, one that today can not be used independently, but only is a combinative part of the basic unit, which in the distant (or perhaps not so distant) past was full-valued, full-fledged, independent word (Dombrovan 2013: 302).

Results of the analysis of word-formation changes in the system of the Old High German language in the aspect of evolutionary linguistic synergetics contributes to the construction of a synergetic model of the German word-formation development as a self-organizing, complex, open, nonlinear macrosystem and outlines new perspectives for the study of historical word-formation dynamics.

References

Bronnik, Larisa. 2012. *Language dynamics in discourse: cognitive-synergistic approach*. Moscow: Moscow Military University. (Doctoral Thesis.)

Dombrovan, Tetiana. 2013. *A synergetic model of the English language development.* Odessa: Odessa National University named after Ivan Mechnikov. (Doctoral dissertation.)

Dombrovan, Tetiana. 2018. *An introduction to linguistic synergetics*. Cambridge: Cambridge Scholars Publishing.

Filicheva, Ninel'. 2003. *History of the German language*. Moscow: Publishing Center 'Academy'.

Haken, Herman. 2000. *Information and Self-organization: A Macroscopic Approach to Complex Systems*. Berlin: Springer-Verlag.

Keller, Rudi. 1994. *On Language Change: The Invisible Hand in Language*. London: Routledge.

Kushnina, Lyudmila. 2004. *The interaction of languages and cultures in the translation space: a gestalt-synergistic approach*. Per'm State University. (Doctoral dissertation.)

Moiseeva, Irina. 2007. *Synergetic model of the text formation*. Chelyabinsk: Chelyabinsk State University. (Doctoral Thesis.)

Myshkina, Nelli. 1999. Text linguistic dynamics: a contradiction-synergistic approach. Perm: Perm State University.

Oliz'ko, Nataliya. 2009. *Semiotics-synergetic interpretation of the features implementation of the categories of intertextuality and interdiscursivity in postmodern artistic discourse*. Chelyabinsk: Chelyabinsk State University. (Doctoral Thesis.)

Pikhtovnikova, Lidiia S. 1999. *Synergy of the fable style: a German rhymed fable in the XIII – XX centuries.* Kharkiv: Bisnes Inform.

Piotrovskiy, Raymond. 2006. *Linguistic synergetics: initial positions, first results, prospects*. St. Petersburg: St. Petersburg State University (Faculty of Philology).

Ponomarenko, Evgeniya. 2010. *Linguosynergetics of business communication from the standpoint of the competence-based approach (based on the English language*. Moscow: MGIMO-University.

Shchyhlo, Larysa. 2017. Language as an object of synergetic studies. In Betil'mirzaeva, Maret, Dreval', Irina, Mishenina, Tetiana & et. al., *Science and innovation in the modern world: philosophy, literature and linguistics, culture and art, architecture and construction, history*, 81 – 89. Odessa: Kuprienko SV. doi:10.21893/978-617-7414-05-5.0.

von Humbolt, Vilhelm. 1964. *On comparative study of languages in relation to different stages of their development. Part I.* Moscow: Education.

von Humbolt, Vilhelm. 1984. *About the influence of different character of languages on literature and spiritual development.* Moscow: Progress.

Larysa Shchyhlo
Department of German Philology
Sumy State University
Ukraine
l.schiglo@gmail.com

In SKASE Journal of Theoretical Linguistics [online]. 2018, vol. 15, no. 3 [cit. 2018-13-12]. Available on web page http://www.skase.sk/Volumes/JTL38/pdf_doc/05.pdf. ISSN 1336-782X.

Stress Assignment in Words with *-i* Suffix in Hebrew*

Ora (Rodrigue) Schwarzwald, Bar-Ilan University, Israel

The purpose of this paper is to describe, analyze and explain the stress patterns of words formed with the homonymic suffix -i which conveys a variety of derivational and inflectional morphological functions in Hebrew, e.g. yaldut-i *'childish' (derivation) and 'my childhood' (inflection). The suffix* -i *functions in two inflectional and three derivational categories: (a) second person singular feminine in verbs (e.g.* šví *'sit down! [F.SG]',* takúmi *'you[F.SG] will get up'); (b) first person singular in nouns, prepositions, and several other parts of speech (e.g.* 'aví *'my father',* kamóni *'like me'); (c) adjectival formation (e.g.* 'olamí *'worldwide',* cíni *'cynical'); (d) gentilic affiliation (e.g.* germaní *'German',* síni *'Chinese'); (e) affectionate expression (e.g.* xamúdi *'sweetie'). Polysemy seems to occur in adjectives (c) and gentilic words (d), however, as each of the functions creates different word classes – only adjectives in (c) and both nouns and adjectives in (d), these categories will be differentiated here. The stress in words with the suffix -i is not fixed, as can be seen in the examples above. In most cases the suffix is stressed (except for (e) which is always unstressed). Penultimate stress is determined in each category by various phonological and morphological rules as well as by other non-linguistic factors which will be described and explained in this paper.*

Keywords: *derivation, homonymy, inflection, stress, suffixes, word class*

1. Introduction

Affixes are morphological devices in word formation and in inflection. They are affixed to certain stems in order to indicate features such as person, number, gender, or tense in inflection, and to create substantives from verbal stems, adjectives from nouns and other words in derivation. In general, specific affixes carry certain morphological and semantic loads; some of which prove to be polysemic (Lieber 2005: 403; Efthymiou 2016: 94 ff.).

Stem is defined in this paper as the form to which an affix is attached, and can be one of the allomorphs of a lexical item. Thus, for instance, the past tense stem of a verb like *lišrok* 'to whistle' is *šarak-*, the future and imperative is *-šrok-*, and the present is *šorek-* (e.g. *šarákti* 'I whistled', *šarka* 'she whistled', *tišrok* 'you.M.SG will whistle', *yišreku* 'they will whistle'; *šoréket* 'whistle.PRS.F.SG', *šorkim* 'whistle.PRS.M.PL').[1] All these words are derivatives of the root *√šrk* which take the CaCaC verbal template *pa'al*.[2] *Šarak, šrok* and *šorek* are the inflectional stems; *šorek* can also be the derivational stem (e.g. *šorkani* 'whistling.ADV').

* This is an expanded version of the paper that I presented at the conference "Word Formation Theories II", Kosice, 27-30 June 2018. I would like to thank the anonymous readers for their valuable comments.

[1] Stress in not marked when ultimate, excluding the cases where the distinctions between ultimate and penultimate are relevant. The sign √ marks a consonantal root.

[2] Many Hebrew words are derived by a combination of consonantal root and vocalic template, sometimes with additional consonants. For example, the following words can be derived from the root *√kšr* when combined with the templates CaCaC, CéCeC, tiCCóCet, CiCuC, CiCeC: *kašar* 'tie.V; signal operator.N', *kéšer* 'knot', *tikšóret* 'communication', *kišur* 'connection', *kašur* 'tied, connected.ADJ', *kišer* 'connect.V'.

Most affixation in Hebrew is suffixed. Suffixes are added to different word classes for inflection and to certain stems for derivation. The analysis of suffixes raises two issues to be discussed in this paper: homonymy in suffixes and conditions for suffix stress patterns.

The suffix *-i* demonstrates a clear case of homonymy: it encompasses various syntactic and semantic functions, and can serve to denote inflection as well as derivation. It has five different functions. The stress of these homonymic suffixes is not fixed, thus different words containing the different instances of *-i* suffixes can be either stressed or unstressed as determined by various phonological, morphological, and other linguistic and non-linguistic factors. Whatever the word class, the stem is quite clear when *-i* is suffixed.

2. The suffix -i

The suffix *-i* is unique because it can be attached to stems belonging to various word classes and possess different, unrelated meanings. Apparently, there are other Hebrew suffixes that can be attached to various word classes; however, unlike the suffix *-i*, these carry the same meaning. For instance, the suffixes *-nu* and *-a* can be attached to various word classes, but their meanings are the same in all instances: suffix *-nu* always indicates first person plural, e.g. *lamád-nu* 'we studied.V', *pí-nu* 'our mouth.N', *lá-nu* 'to.PREP us'; suffix *-a* always marks the third person singular feminine, e.g. *lamd-á* 'she studied.V', *pkid-á* 'clerk.N.F', *l-a* 'to.PREP her'.[3]

The suffix *-i* functions in the following cases:

(a) Second person singular feminine when affixed to a verb in the future tense and the imperative, e.g. *tipl-í* 'you.F.SG will fall', *kúm-i* 'get up.F.SG!' (2.1)

(b) First person singular marker: possessive when suffixed to singular nouns; pronominal when suffixed to prepositions and other word classes, e.g. *yad-í* 'my hand', *šel-í* 'mine, of me', *kamó-n-i* 'like me' (2.2)

(c) Adjective, when suffixed to nouns, adjectives and other word classes, e.g. *yaldut-í* 'childish', *rišon-í* 'initial', *ba'albát-i* 'landlord-like' (2.3)

(d) Gentilic affiliation, when suffixed to names of locations, religions and sources, e.g. *rus-í* 'Russian', *súr-i* 'Syrian', *nocr-í* 'Christian' (2.4)

(e) Affection and diminution when suffixed to proper names and some adjectives, e.g. *rút-i* 'Ruthie', *xamúd-i* 'cutie' (2.5)

Functions (a) and (b) are inflectional whereas (c)–(e) are derivational.[4] Function (a) is unique in that there is no other way in which to express second person feminine in the future or the imperative. Function (d) is also special because it is the standard way of forming gentilic substantives. In function (b) the suffix *-i* occurs in complementary distribution with *-ay* which is attached to plural nouns and some prepositions (cf. *yad-ay* 'my hands', *'el-ay* 'to me, towards me').[5] Functions (c) and (e) can be formed using other morphological devices. Adjectives can be formed with a consonantal root and template combination as in (1) or by other suffixes as in (2).

---

[3] In a limited number of cases, an unstressed *-a* serves as a directionality marker, e.g. *'axór-a* 'backwards', *cafón-a* 'to the north' (Hitin-Mashiah 2005).

[4] Ornan (1997) considers (d) to be inflectional, but Schwarzwald (1998a) refutes this.

[5] Phonetically, the suffix *+áy* can also be attached to nominal stems to derive agent nouns, e.g. *bank+áy* 'banker', *xašmal+áy* 'electrician'. However, this suffix is spelled using the letters *aleph* and *yod*, unlike the first person singular suffix which is spelled only with a *yod*.

(1) √*šmn* + CaCeC > *šamen* ‘fat’, √*lmd* + meCuCaC > *melumad* ‘educated’, √*šlm* + muCCaC > *mušlam* ‘perfect’, √*kdm* + CaCuC > *kadum* ‘ancient’, √*rgz* + CaCCan > *ragzan* ‘grumpy, irritable’

(2) *zá'af* ‘anger, rage’ + *-an* > *za'afán* ‘quick-tempered’, *klum* ‘nothing’ + *-nik* > *klúm-nik* ‘good-for-nothing person’

Affection and diminution (e) can also be expressed by single stems, as in (3), by consonantal root duplication as in (4), or by other diminutive suffixes as in (5) (Bat-El 2005; Bolozky 2013).

(3) *xamúda* ‘cutie.F’ (cf. *xamudá* ‘cute.F’), *mótek* ‘honey’

(4) *klavlav* ‘puppy’ (*kélev* ‘dog’), *xataltul* ‘kitten’ (*xatul* ‘cat’), *sfamfam* ‘small mustache’ (*safam* ‘mustache’)

(5) *sus-on* ‘little horse’ (*sus* ‘horse’), *kar-it* ‘cushion, pad’ (*car* ‘pillow’), *yáfa-le*, *yáf-inka*, *yáf-ki*, *yáf-čuk* ‘Yafa.AFFC’, *mótek-le* ‘sweetie’, *áb-uš* ‘Dad-dear’ (*'ába* ‘Dad’), *sarít-uš* ‘Sarít-dear’ (cf. Dressler 2005: 274)

Stress is only steady on the stem for function (e) of suffix *-i*. For all other functions (a)–(d) the stress varies. Ultimate stress is mostly unmarked, namely, the suffix is stressed. However, sometimes the stress is penultimate and is placed on the final syllable of the stem. Variations are determined in each case by a range of phonological, morphological, and other factors which will be explained in the following sections.

2.1 *Second person feminine suffix in the future and imperative forms*

The stress of the second person feminine suffix *-i* should be examined together with other verbal system vocalic suffixes in the, past, future and imperative verbal forms: *-a* (PST.3F.SG), *-u* (PST and FUT.3PL; FUT and IMPR2PL).

The distribution of the affixes in the Hebrew verbal system is listed in Table 1 and exemplified below using the verb *šilem* ‘pay’:

As can be seen in Table 1 and in the following example, a number of person suffixes begin with the consonants *t* or *n*, while others start with a vowel. The suffixes that start with a consonant are always unstressed; the stress stays on the stem vowel, as in (6). The stem form of various verbal templates in the past tense is given before each group of examples shown here.[6]

[6] There are five active verbal templates in MH into which consonantal roots are interwoven. These are formed as: CaCaC, niCCaC, CiCeC, hiCCiC, and hitCaCeC. Each verbal template is called *binyan*, and they are named *pa'al*, *nif'al*, *pi'el*, *hif'il*, and *hitpa'el*. The passive templates are *pu'al* (CuCaC) and *huf'al* (huCCaC), and are seldom used in speech. It is possible to use the passive templates for the future tense but not for the imperative, e.g. *tešulví* ‘you.F.SG will be integrated’, *tuxšerí* ‘you.F.SG will be trained’.

Table 1: Past, future and imperative verb affixes

Number	Person	Past Tense	Future Tense	Imperative
SG	1M 1F	— *-ti*	*ʔ-~y-* — *-Ø*	
	2M	— *-ta*	*t-* — *-Ø*	— *-Ø*
	2F	— *-t*	*t-* — *-i*	— *-i*
	3M	— *-Ø*	*y-* — *-Ø*	
	3F	— *-a*	*t-* — *-Ø*	
PL	1M 1F	— *-nu*	*n-* — *-Ø*	
	2M	— *-tem*	*t-* — *-u*	— *-u*
	2F	— *(-ten)*[7]	*(t-* — *-na)*	— *(-na)*
	3M	— *-u*	*y-* — *-u*	
	3F	— *-u*	*(t-* — *-na)*	

Past tense: *šilám-ti, šilám-ta, šilám-t, šilém, šilm-á, šilám-nu, šilám-tem, šilm-ú*
Future tense: *'a-šalém~ye-šalém, te-šalém, te-šalm-í, ye-šalém, te-šalém, ne-šalém, te-šalm-ú, ye-šalm-ú*
Imperative: *šalém, šalm-í, šalm-ú*

(6) *katav* ‘write’ (*√ktv, pa'al*): *katáv-ti* ‘I wrote’, *katáv-tem* ‘you.PL wrote’
šilev ‘integrate’ (*√šlv, pi'el*): *šiláv-ta* ‘you.M.SG integrated’, *šiláv-t* ‘you.F.SG integrated’
niršam ‘register’ (*√nršm, nif'al*): *niršám-nu* ‘we registered’, *niršám-ti* ‘I registered’
hitxil ‘begin’ (*√txl, hif'il*): *hitxál-tem* ‘you.PL started’, *hitxál-ta* ‘you.M.SG started’
hitlabeš ‘get dressed’ (*√lbš, hitpa'el*): *hitlabáš-ti* ‘I got dressed’, *hitlabáš-nu* ‘we got dressed’
huksam ‘be fascinated’ (*√ksm, huf'al*): *huksám-ta* ‘you.M.SG were fascinated’, *huksám-nu* ‘we were fascinated’

The unmarked stress of vowel suffixes is ultimate in most verb templates, and the stem final vowels *e, a, o* are deleted (Graf & Ussishkin 2003; Bat-El 1993, 2018). A few examples are listed in (7).

(7) *katav* ‘write’ (*√ktv, pa'al*): *katv-á* ‘she wrote’, *yixtev-ú* ’they will write’
šilem ‘pay’ (*√šlm, pi'el*): *šilm-ú* ‘they paid’, *tešalm-í* ‘you.F.SG will pay; pay!’
nixnas 'enter' (*√kns, nif'al*): *tikans-í* ‘you.F.SG will enter’, *nixnes-ú*[8] ‘they entered’
yašav ‘sit' (*√yšv, pa'al*): *šv-í* ‘sit down.F.SG!’, *yašv-ú* ‘they sat’
natan ‘give’ (*√ntn, pa'al*): *yitn-ú* ‘they will give’, *tn-í* ‘give.F.SG!’
hitbayeš ‘be embarrassed’ (*√byš, hitpa'el*): *titbayš-ú* ‘you.PL will be embarrassed; shame on you!’

[7] The affixes in parentheses are only used in very high registers, not in standard speech, and will therefore not be exemplified here.
[8] The vowel *e* is inserted here between *n* and *s* and between *t* and *v* in *yixtev-ë* above in order to avoid a three consonantal word cluster (Schwarzwald 2004).

kubac ‘be gathered’ (√*kbc*, *pu'al*): *yekubc-ú* ‘they will be gathered’, *kubc-á* ‘she/it was gathered’

The vowel suffixes *-i*, *-a*, and *-u* are not stressed in the verbal system in two circumstances: (a) the *hif'il* template, as in (8) and (b) bi-consonantal verbs, as in (9).[9] The examples in (10) feature both circumstances. In all cases the stem vowels are never deleted.

(8) *hitxil* ‘begin’ (√*txl*): *tatxíl-i* ‘you.F.SG will begin’, *hitxíl-a* ‘she began’, *yatxíl-u* ‘they will begin’, *tatxíl-u* ‘you.PL will begin; begin.PL!’
hirkid ‘lead a dance’ (√*rkd*): *tarkíd-i* ‘you.F.SG will lead a dance’, *hirkíd-u* ‘they led a dance’
hipil ‘drop’ (√*npl*): *hipíl-a* ‘she dropped’, *tapíl-i* ‘you.F.SG will drop’[10]
hošiv ‘seat’ (√*yšv*): *hošív-u* ‘they seated’, *yošív-u* ‘they will seat’
himci ‘invent; discover’ (√*mc'*): *himcí'-a* ‘she invented’, *tamcí'-i* ‘you.F.SG will invent; discover!’

(9) *šar* ‘sing’ : *tašír-i* ‘you.F.SG will sing; sing!’, *yašír-u* ‘they will sing’ (√*šr, pa'al*)
kam ‘get up’ : *kám-a* ‘she got up’, *kúm-u* ‘get up.PL!’, *kúm-i* ‘get up!’ (√*km, pa'al*)
ba ‘come’ : *bá'-a* ‘she came’, *bó'-u* ‘come.PL!’, *tavó'-i* ‘you.F.SG will come; come.PL!’ (√*b', pa'al*)

(10) *hekim* ‘establish’ (√*km, hif'il*): *takím-i* ‘you.F,SG will establish’, *hekím-a* ‘she established’, *hekím-u* ‘they established’
he'et ‘slow down’ (√*'t, hif'il*): *ta'ít-i* ‘you.F.SG will slow down; slow down!’, *he'ét-a* ‘she slowed down’, *he'ét-u* ‘they slowed down’
hegen ‘protect’ (√*gn, hif'il*): *tagén-i* ‘you.F.SG will protect’, *hegén-u* ‘they protected’

The common feature for both circumstances is that the stem final vowel CVC is not deleted, unlike the cases presented in (7). If we adopt Podolsky's (1991) and Melčuk & Podolsky's (1996) classification regarding mobile (shifting) versus stable (fixed) stress, the explanation is clear. The stem vowels in (7) are classified as mobile and can therefore potentially be deleted so that the suffixed vowel takes the stress. The stem vowels in (8–10) are stable and thus cannot be reduced with the addition of a vowel suffix;[11] in these cases then, the stem vowel is stressed and the suffix is unstressed.

The examples in (11) seem to contradict the circumstances given above. Although they appear to be bi-consonantal and take the *hif'il* template, they nonetheless behave like the examples given in (7), namely, their vowel suffixes are stressed.

(11) *hikna* ‘instill, acquire.TRNS’ (√*kny*): *takn-í* ‘you.2F.SG will instill’, *hikn-ú* ‘they instilled’
her'a ‘show’ (√*r²y*): *tar'-í* ‘you.2F.SG will show, show!’, *her'-ú* ‘they showed’, *yar'-ú* ‘they will show’

[9] Graf & Ussishkin (2003: 252–252) use the name monosyllabic rather than bi-consonantal verbs. These verbs are traditionally identified as CwC/CyC (hollow) or $C_1C_2C_2$ (geminate) roots.

[10] The consonantal root of these apparently two consonantal verbs is clear from other occurrences of the root, e.g. *nafal* 'fall', *yašav* 'sit'.

[11] It should be noted that the stem vowel in the *hif'il* template is high, and that some of the forms in (9) also include the stem high vowels *i* or *u*.

hiška ‘water.TRNS’ (√*šky*): *tašk-í* ‘you.2F.SG will water; IMPR.F.SG!’, *tašk-ú* ‘you.PL will water; water.PL!’,*hišk-ú* ‘they watered’
hifna ‘turn.TRNS’ (√*fny*): *tafn-í* ‘you.2F.SG will turn’ *yafn-ú* ‘they will turn’

The difference between the verbs given in (11) and (8–10) is that the stems in (11) end in a vowel, rather than in a consonant. The stems of all the verbs shown in (8) and (10) end with a CVC syllabic structure rather than a CV one (cf. *hikna, her'a, hiška, hifna* in 11 versus *hitxil, hošiv, hipil, hegen* in 8 and 10).

Thus we can formulate the following generalizations:
G1. Suffixes that start with a consonant in the past, future or imperative forms of a verb are always unstressed.
G2. Suffixes that start with a vowel in the past, future or imperative forms of a verb, are stressed unless the verb ends in CVC and takes the *hif'il* template or is a bi-consonantal verb, in which case the stress remains on the final stem vowel.

The second generalization G2 applies to the suffix *-i* which is the focus of our discussion. In general it is stressed, unless it appears within the *hif'il* template and in the case of a few bi-consonantal verbs which in spite of their paucity are frequently used.

2.2 *First person marker*

The suffix *-i* is added to singular nouns in order to indicate the possessive pronoun 'mine'. It is also appended to prepositions and to a few other closed class categories as the pronominal first person singular. The suffix is always stressed when attached to nouns (2.2.1) and to most prepositions. It is not stressed when an *-n-* is inserted before the suffix in prepositions and other closed class words (2.2.2). Although the addition of *-i* to nouns and to other closed class categories is partly optional, it is obligatory when added to prepositions.

2.2.1 *Possessive*
There are two ways to express possession in nouns; the synthetic way by adding the pronominal suffixes, and the analytic way by using of inflected *šel* ‘of’. The suffix *-i* is part of a set of suffixes presented in Table 2, and this discussion applies to them as well, although the examples given only refer to the suffix *-i*.[12]

[12] In section 2 I mentioned the suffix *-ay* which is part of another set of bound pronominal suffixes which are added to plural nouns and to some prepositions. I have not listed this here because it is not relevant to the discussion regarding the suffix *-i*. Unlike the person suffixes in the verb discussed above, there is no stress difference between suffixes starting with a consonant or a vowel in this set of suffixes in Table 2. They are always stressed when attached to nouns.

Table 2: Bound pronominal suffixes in Hebrew

	Singular		Plural	
Person	Masculine	Feminine	Masculine	Feminine
1	*-i* (*-ni*)		*-ánu, -énu*	
2	*-xa*	*-ex, -ax, -x*	*-xem*	(*-xen*)
3	*-o* (*-no, -nu, -hu*)	*-a(h) -ha* (*-na*)	*-am, -hem*	(*-an, -en*)

Even though synthetic possession is not very productive in Modern Hebrew due to analytic possession with the inflected preposition *šel* 'of', there are certain categories where it is still actively used. The examples in (12) show parallels between the synthetic (12a) and analytic (12b) possession of nouns which have the same meaning, although the synthetic expression belongs to a higher register.

(12) a. Synthetic possession: *dod-í* 'my uncle', *roš-í* 'my head', *ba-xalom-i* 'in my dream'

b. Analytic possession: *dod šel-í* 'my uncle (lit: uncle of-mine)', *ha-roš šel-í* 'my head (lit: the-head of-mine)', *b-a-xalóm šel-í* 'in my dream (lit: in-the-dream of-mine)'

The inflected nouns in the following categories are quite regularly used in everyday speech and writing (Dubnov 2000; Avioz 2004: 179–198; Ravid & Chahana-Amitay 2005):

a. Family members and close relations, e.g. *'axot-í* 'my sister', *'im-í* 'my mother', *dodat-í* 'my aunt', *'išt-í* 'my wife', *ba'al-í* 'my husband'
b. Body parts, e.g. *p-í* 'my mouth', *roš-í* 'my head', *lešon-í* 'my tongue', *yad-í* 'my hand', *lib-í* 'my heart'
c. Judgment and consideration, e.g. *da'at-í* 'my opinion', *haxlatat-í* 'my decision', *'acat-í* 'my advice', *ta'anat-í* 'my argument', *tor-í* 'my turn'
d. Abstract nouns indicating verbal or adjectival contents, e.g. *bakašat-í* 'my request', *zxut-í* 'my right', *'ezrat-í* 'my help', *'ašmat-í* 'my fault'
e. Lifetime periods, e.g. *yaldut-í* 'my childhood', *ziknat-í* 'my old age'

Inflection also occurs in some fossilized expressions, e.g. *le-da'at-í* 'in my opinion' (lit: to-opinion-mine), *le-metav zixron-í* 'as far as I remember' (lit: to-best memory-mine), and the fossilized addressing expressions *'ax-í* 'my brother, friend' and *yakirat-í* 'my dear (feminine friend)'.

The following generalization summarizes the fact that the possessive suffix *-í* in singular nouns is always stressed:
G3: The possessive first person suffix *-i* is stressed when added to singular nouns.

2.2.2 *Prepositions and other closed lexical categories*
All Hebrew prepositions are inflected to signify person. The stressed suffix *-i* indicates the first person singular,[13] as in (13).

[13] As mentioned earlier, the suffix *-ay* indicates the same person in several other prepositions, e.g. *'al-ay* 'on me', *'el-ay* 'towards me'.

(13) *l-í* ‘to me’, *'ecl-í* ‘by me’, *b-í* ‘in me’, *betox-í* ‘inside me’, *šel-í* ‘of me’, *biglal-í* ‘because of me’, *bišvil-í* ‘for me’, *'ot-í* ‘ACC me’, *'it-í* ‘with me’, *negd-í* ‘against me’, *mul-í* ‘in front of me’

There are two prepositions where the consonant *-n-* is inserted before the suffix *-í* and the stress stays on the last syllable of the stem rather than on the suffix, as in (14).

(14) *kmo* ‘like, as’ + *-i* > *kamó-n-i* ‘like me’; *min* ‘from’ + *-i* > *mimé-n-i* ‘from me’

In classical Hebrew, the *-n-* was regularly inserted into verbs and gentilic substantives ending in a vowel if the suffix *-i* was to be added to them, in order to prevent hiatus.[14] This fact explains the occurrence of *n* in *kmo*, but not in *min* where the inflection is irregular in all regards.

Very few other words that belong to closed class categories are inflected. The examples in (15) show two quantifying words, *kol* ‘all’ and *rov* ‘most’ and the adverbials *levad* ‘alone, by oneself’ and *le'at* ‘slowly’, that behave in the same way as regular prepositions.

(15) *kul-í* ‘all of me’, *rub-í* ‘most of me’, *levad-í* ‘by myself’, *le'it-í* ‘I slowly’[15]

In the examples given in (16), inflected synthetic words alternate with analytic ones using the independent personal pronoun. In the examples where *n* is inserted before the suffix *-i*, the stress remains on the stem, except for *hinení~hinéni* (16c). The examples in (17) demonstrate actual use. These inflected forms belong to a high literary register and are never used in day-to-day, ordinary speech.

(16) a. *'eyn* ‘no, there is no’ – *'eyn-í~'eyné-n-i ~ 'eyn 'ani* ‘I am not’ (*'ani* ‘I’).

b. *'od* ‘still’ – *'od-í ~ 'odé-n-i ~ 'od 'ani* ‘I am still’

c. *hine* ‘here, behold’ – *hine-n-í~hiné-n-i ~ hiné~híne 'ani* ‘here I am’

d. *harey* ‘indeed, behold, is it not that’ – *harév-n-i ~ harey 'ani* ‘I behold’

(17) a. *'eyn-i xošév-et ~ 'eyné-n-i xošév-et ~ 'eyn 'ani xošév-et* ‘I.F do not think-F’

b. *be-'od-í mehases ~ be'odé-n-i mehases ~ be'od 'ani mehases* ‘while I.M am still hesitating’

c. *hine-n-í mešalem ~ hine 'ani mešalem* ‘I.M hereby pay’

---

[14] a. In fact, the suffix *-i* can theoretically be added to verbs to indicate the accusative first pronoun, as in classical Hebrew, e.g. *ra'á-n-i* ‘he saw me’, *hizkartá-n-i* ‘you.M.SG reminded me’, but it is hardly used in Modern Hebrew. The analytic way is preferred whereby the inflected accusative is used, namely, *ra'a 'oti*, *hizkarta 'oti~li*. In these obsolete forms an *n* is inserted before the suffix as in the examples (14) and (16–17).

b. The *n* is added only before the suffix *-i*, and not before any of the other pronominal suffixes.

[15] Vowel alternations (*u-o, a-i*) occur in many bi-consonantal (geminate) roots to which the words belong (√kl, √rb/v, √'t from historical √kll, √rbb, and √'tt). Alternations of *b–v*, *k–x* and *p–f* are residues of the historical Spirantization Rule in Hebrew (Schwarzwald 2001: 14–16).

d. *haréy-n-i modía* ‘I.M hereby announce’ ~ *hare(y) 'ani modía* ‘eventually I.M announce’

All the forms in (14) and (16) which include an additional *-n-* before the suffix are inherited from classical Hebrew. This formation would originally have been phonologically motivated, but today it is simply a remnant of the old processes. While the prepositions in (14) are frequently utilized, the use of the examples in (16-17) is quite rare and needs to be memorized by rote. The following generalization can be stated:

G4: The first person singular pronominal suffix *-i* is stressed when it appears in prepositions and in the words *kol* ‘all’, *rov* most’, *levad* ‘alone’, and *le'at* ‘slowly’.

The prepositions *kmo* ‘as’ and *min* ‘from’, as well as the words *'eyn* ‘no, there is no’, *'od* ‘still’, *hine* ‘here, behold’, and *harey* ‘indeed, behold’ are exceptional because the stress stays on the stem for these words when an *n* is inserted before the suffix *-i*. They are residues of forms inherited from the classical Hebrew periods and need to be memorized by rote.

2.3 *Adjectives*

As stated above, the suffixation of *-i* is only one of the ways in which adjectives are formed (Werner 1983; see examples 1–2 above).[16] The formation of adjectives with the suffix *-i* is very productive, and words from different classes (except for verbs) serve as the base for adjectival stems, although most of these are derived from nouns, as in (18).

(18) a. From nouns: *'iš-í* ‘personal’ (*'iš* ‘man, person’), *tarbut-í* ‘cultural’ (*tarbut* ‘culture’), *sod-í* ‘secretive’ (*sod* ‘secret’), *pnim-í* ‘internal, interior’ (*pnim* ‘inside, interior’), *keyc-í* ‘summery’ (*káyic* ‘summer’)

b. From adjectives: *tipš-í* ‘silly [-ANIMATE]’ (*tipeš* ‘stupid [+HUMAN]’), *rišon-í* ‘basic, elementary’ (*rišon* ‘first’), *'evil-í* ‘silly’ (*'evil* ‘stupid.N’)

c. From adverbs: *'adkan-í* ‘updated’ (*'ad kan* ‘until here’), *yomyom-í* ‘daily’ (*yom-yom* ‘every day’), *'axšav-í* ‘contemporary’ (*'axšav* ‘now’)

d. From prepositions: *le'umat-í* ‘oppositional’ (*le'umat* ‘against’), *bil'ad-í* ‘exclusive, unique’ (*bil'adey* ‘without’)

Although most adjectives with the suffix *-i* are derived from singular nouns, the stem is occasionally taken from a noun’s plural form, as in (19).

(19) *naš-í* ‘feminine’ (*naš-im* ‘women’; *'iša* ‘woman’), *cdad-í* ‘sideways.ADJ, lateral’ (*cdad-im* ‘sides’; *cad* ‘side.N’), *'imah-í* ‘maternal’ (*'imah-ot* ‘mother-s’; *'em* ‘mother’, *'ima* ‘Mom’), *raš-í* ‘main, head.ADJ’ (*raš-im* ‘heads’; *roš* ‘head’)

[16] The only template where the ending *-i* occurs is the adjectival CCiCi which is used to form the ordinal numbers from 3–10 and a few other words, e.g. *šliši* ‘3rd’, *revi'i* ‘4th’, *xamiši* ‘5th’, *švi'i* ‘7th’, *šmini* ‘8th’, *tši'i* ‘9th’, *'asiri* ‘10th’, *recini* ‘serious’, *'ariri* ‘alone, childless’. The vowel *e* is added after the first consonant in *revi'i, 'asiri, recini,* and *'ariri* because of a restriction on consonant clusters in word initial position (Schwarzwald 2004). This template is not listed anywhere in any list of noun templates (Bolozky 1999).

In a few very rare cases the stem is not based on a Hebrew lexical item, but is rather modified from a foreign word which does not exist in the Hebrew lexicon, as in (20).

(20) *xad-sitr-í* 'one way (street) (*xad* 'one', Aramaic *sitra* 'side' *-i*), *xakla'-í* 'agricultural' (Aramaic *ḥaqla* 'field'), *siton-í* '(commerce) wholesale' (Greek *sitones* 'grains trader')

Homonyms are sometimes created when the suffix *-i* indicating a first person possessive noun (see 2.2.1 above) and the adjectival suffix *-i* are added to the same nominal stem, e.g. *'arc-í* 'my country; national' (*'érec* 'country, land'). The context always enables speakers to distinguish between these, as the examples in (21) show.

(21) *zot yaldut-í*
this.F childhood.F.-1SG
'This is my childhood'

be-yaldut-í gár-nu be-giv'atáyim
in-childhood.F-POSS1SG live.PST-1PL in-Givatayim
'In my childhood we lived in Givatayim'

ze yaldut-í
this.M childhood.F-ADJ
'It is childish'

yeš l-o parcuf yaldut-i
there is to-him face childhood.F-ADJ
'He has a childish face'

The examples given in (18–21) include adjectives where the suffix *-i* is stressed. The suffix *-i* is not stressed in non-integrated words, especially loan words and acronyms (Schwarzwald 2013), as in (22).

(22) a. Loan stems: *'idyót-i* 'idiotic', *pedagóg-i* 'pedagogical', *didákt-i* 'didactic', *téxn-i* 'technical', *histór-i* 'historical', *'ortodóks-i* 'orthodox', *generatív-i* 'generative', *kolegyál-i* 'collegial', *banal-i* 'banal', *telefón-i* 'telephonic', *ba'albát-i* 'landlord-like' (Yiddish-like formation: *bá'al* 'owner' [*báyit* 'house'], plural *bat-im* 'houses' *-i*)

b. Acronyms: *mankál-i* 'of chief director' (*MaNKaL* 'general manager' = ***me****n**a**hel* 'director', ***kl****al-i* 'general'); *xagám-i* 'lacking person, number and gender' (*XaGáM* = ***x****asar* 'lack', ***g****uf* 'person', ***m****in* 'gender', ***m****ispar* 'number'), *mafmár-i* 'of supervisor of the center of the subject' (*MaFMaR* 'director of high-school studies in a certain subject' = ***me****f**akéax* ***me****r**kaz mikcóa* 'supervisor of center subject')

There is a clear distinction here between originally Hebrew words and non-integrated words (mainly loan words). In Hebrew stems the suffix *-i* is stressed whereas in non-integrated words the stress remains on the stem. Non-integrated stems need to be marked in the lexicon by the lexical rule L1 to denote penultimate stress when the suffix *-i* is attached. The final syllable of

the stem is stressed, although the basic word may not retain an ultimate stress, e.g. *télefon, koléga* ‘colleague’.

L1: Non-integrated words are penultimately stressed when the adjectival suffix *-i* is attached to them.

Unmarked stems are Hebrew in origin where the stress is assigned to the suffix *-i*. This is stated by the grammatical rule G5.

G5: Adjectival suffix *-i* is stressed.

2.4 *Gentilic marker*

The category of gentilic substantives indicating ethnic, geographic or religious origin is relatively small considering the number of lexical items which it includes.[17] It is different from adjectival formation with the suffix *-i* described in 2.3 because all the words in this category are both adjectives and nouns whereas words which are adjectives can only occasionally be used syntactically as nouns. Adjectives can be formed as a result of other word formation devices as demonstrated in (1), whereas gentilic words are only formed by suffixation.[18] As nouns, gentilic substantives mark [+HUMAN] entities, whereas as adjectives, like the adjectives in 2.3, they can refer to any feature, human or non-human. The plural inflection of masculine nouns is also different from that of adjectives: masculine nouns are pluralized by adding the suffix *-im* (N-*i* + *-im* > N-*im*), while masculine adjectives are pluralized by the addition of the suffix *-iyim* or *-iim* (ADJ-*i* + *-im* > ADJ-*iyim~iim*), e.g. *yapáni-yapánim* ‘Japanese.N.M.SG-PL’, *yapáni-yapániyim~yapániim* ‘Japanese.ADJ.M.SG-PL’.[19]

Most gentilic substantives are penultimately stressed, as indicated in G6, because they are generally derived from non-Hebrew place names or loan words, and are thus linguistically non-integrated (Melčuk & Podolsky 1996; Schwarzwald 1998b).

G6: Gentilic substantives with suffix *-i* are penultimately stressed.

The feminine suffix for most gentilic substantives is *-t* as exemplified in (25–26) below. For a small number of gentilic words with a final stressed *-i* in the masculine, the feminine noun is not formed in the same way as the feminine adjective: the noun indicating a woman is formed using the suffix *-a*, whereas an adjective takes the suffix *-t* in the feminine,[20] as in (23) (Rosén 1956: 236; Podolsky 1981; Faust 2017). No gentilic word with an unstressed *-i* takes the feminine suffix *-a* in Hebrew.[21] The word *carfati* ‘French’ was derived from Hebrew *carfat* 'France', and *yehudi* ‘Jew’ was derived from *Yehuda* ‘Judah’, both of which differ in their feminine forms. A few more examples of this type of noun are listed in (24).

(23) a. *ha-carfati yošev* *ha-carfatiy-a yošév-et*

[17] Faust (2017) prefers the name *denomyns* for gentilic substantives.

[18] The loan suffix *-nik* is occasionally used in substandard Hebrew to form gentilic substantives, e.g. *xúl-nik* 'foreigner, someone who came from abroad' (*xul* 'acronym of ***xu****c* ***l****a'arec* 'out of the country'), *ašdod-nik* 'someone from Ashdod', *gúr-nik* 'one who belongs to the Hassidic cycle of Rabbi Gur'. However, the suffix *-nik* is also used in slang expressions to form other substantives, not necessarily gentilic, e.g. *kolboy-nik* 'deep bowl set on a dining table for leftovers; a man of multiple capacities' (*kol-bo* 'all in-it'), *šipuc-nik* 'renovator (of houses and apartments)', and see *klumnik* in example (2) above.

[19] The feminine forms are identical: *yapánit-yapániyot* 'Japanese.f.sg-pl'.

[20] A glide *y* is inserted between *i* and *a*.

[21] Faust (2017) has a theoretical explanation for the choice of *-a* as the feminine formation suffix, but his analysis does not account for the exceptions stated below.

	the-French.M sit 'The Frenchman sits'	the-French-F sit-F 'The Frenchwoman sits'
	léxem carfati bread.M French 'Baguette, French bread'	*nešika carfati-t* (no **carfatiya*) kiss.F French-F 'French kiss (mouth to mouth)'
b.	*hu yehudi* he Jew 'He is a Jew'	*hi yehudiy-a* she Jew-ess 'She is a Jewess'
	ha-minhag ha-yehudi the-custom.M the-Jewish.M 'the Jewish custom'	*ha-'emuna ha-yehudit-t* (no **yehudiya*) the-faith.F the-Jewish-F 'the Jewish faith'

(24) *'italki-'italkit, 'italkiya* 'Italian', *germani-germanit, germaniya* 'German', *dati-datit, datiya* 'religious', *'ivri* 'Hebrew', *'aravi* 'Arab', *polani* 'Polish', *co'ani* 'gypsy', *rusi* 'Russian', *teymani* 'Yemenite', *xiloni* 'secular', *sfaradi* 'Spanish, Sephardi', *romani* 'Romanian', *'aškenazi* 'Ashkenazi', *parsí* 'Iranian'

However, not all gentilic words with a final stressed *-i* have two feminine forms, as the examples in (25) demonstrate.

(25) a. *lita'í-lita'ít* 'Lithuanian.M-F' (*líta* 'Lithuania'; no **lita'iyá*), *vina'í-vina'ít* 'Viennese' (*vína* 'Vienna'), *roma'í* 'Roman' (*róma* 'Rome'), *xeyfa'í* 'man from Haifa'

b. *yisre'elí-yisre'elít* 'Israeli.M-F' (no **yisre'eliya*), *pariza'í-pariza'ít* 'Parisian.M-F' (no **parisa'iya*), *pliští* 'Philistine' (no **plištiya*), *'ašurí* 'Assyrian', *'aramí* 'Aramaic', *mizraxí* 'oriental'

The nouns from which gentilic words are derived in (25a), end with the vowel *a*, and this seems to have produced a phonological restriction on the formation of two feminine forms, nominal and adjectival. However, the other examples in (25b) do not form any phonological group. Moreover, the examples in (26b) show that the suffix *-i* is not stressed although some place names end in *a* (*súrya, 'amérika*, *kánada*, *'áfrika*).

Many other penultimately stressed gentilic nouns and adjectives, only utilize the suffix *-t* to signify the feminine form for both nouns and adjectives, e.g.

(26) a. *yapáni-yapánit* 'Japanese.M-F' (*yapan* 'Japan'; no **yapaniya*), *lúbi-lúbit ~ lúvi-lúvit* 'Lybian' (*luv* 'Lybia'), *síni-sínit* 'Chinese' (*sin* 'China'), *holándi* 'Dutch' (*hóland* 'Holland')

b. *súri-súrit* 'Syrian.M-F' (*súrya* 'Syria'; no **suriya*), *bélgi* 'Belgian' (*bélgya* 'Belgium'), *kanádi* 'Canadian' (*kánada* 'Canada'), *'afriká'i-'afriká'it* 'African.m-f' (*'áfrika* 'Africa'; no **'afrika'iya*), *'ameriká'i ~ 'amerikáni* 'American' (*'amérika* 'America')

c. *brazilá'i-brazilá'it* ‘Brazilian.M-F’ (*brazil* ‘Brazil’; no **brazíliya*)

d. *meksiká'i-meksiká'it ~ meksikáni-meksikánit* ‘Mexican.M-F’ (*méksiko* ‘Mexico’; no **meksika'iya*, **meksikaniya*), *maroká'i* ‘Moroccan’ (*maróko* ‘Morocco’)

The examples in (26a) include gentilic substantives in which the suffix-*i* is added to the noun after a consonant and the stress stays on the stem; in (26b) the suffix *-i* is added to the final vowel *a* of the country's name, similar to (25a), but the stress is penultimate rather than ultimate. In the examples in (26c) and (26d) *-a* is added before *-i* (as *-á'i*), by analogy to the forms in (25a) and (26b), either after a consonant in (26c) or after a vowel in (26d), but without the stress on the suffix *-i* (cf. 25a). The choice of suffix *-i* (or *-á'i*) seems arbitrary and cannot be accounted for by any phonological rule. The stress in these cases is penultimate.

Some of the nouns in (24) and (25) have been inherited from classical Hebrew, but not all of them. New gentilic nouns are only penultimately stressed and the nominal and adjectival feminine forms are indistinguishable. Ultimate suffix stress was typical of classical Hebrew and featured heavily in the formation of new gentilic words during the early days of Hebrew speech revival.[22] Words were ultimately stressed if inherited or if coined at the end of the nineteenth century and the beginning of the twentieth century when normative standards were set by the classical language (Mor 2017). While some of these words stuck to the normative classical distinction of *'ivriyá* ‘Hebrew.F.N’ and *'ivrít* ‘Hebrew.F.ADJ; the language name’[23] as a precedence, most of them did not, not only because they did not appear in any classical sources (the Bible, the Mishna, etc.), but also because they were perceived as loan words which are typified by non-final stress. In this way, non-linguistic, sociological and historic factors determine most gentilic word stress assignment and feminine inflection.

Gentilic words of the types presented in (23–25) must be lexically assigned for the stressed suffix *-i* with the following lexical rule L2 which is different from the grammatical rule G6 stated above:

L2: Gentilic substantives are stressed on the final syllable.

The feminine nominal and adjectival distinctions presented above can be accounted for by the minor lexical rule L3 which states:

L3: Gentilic substantives with final stress in the masculine take the feminine suffix *-a* to form a [+HUMAN] noun, and *-t* to form an adjective.

This rule only applies to words such as those given in (23) and (24). The words listed in (25a) as well as those demonstrated in (26c) and (26d) must be lexically marked for their formation and need further consideration elsewhere.

2.5 *Affection*

Expression of affection determines a diminutive formation in which the palatal front high vowel is used, as in many other languages (Jespersen 1965; Bauer 1996; Dressler & Merlini

[22] An exception is the biblical word *kuší* ‘Ethiopian, from Kush’, which is penultimately stressed in Modern Hebrew, *kúši*, and has been changed to derogatively mean ‘black’, and therefore belongs to group (e) (see 2.5 below).

[23] The suffix *-it* is typical of language names, e.g. *yapánit* ‘Japanese (language)’, *swahílit* ‘Swahili’. It is also typical of adverbials, e.g. *'ekronít* ‘in principle’, *téxnit* ‘technically’.

Barbaresi 1994, 2001). The suffix *-i* is never stressed in this function as the following examples show; the stress remains intact on the stem.[24]

(27) *oríti* ‘Orít~Órit’, *matóki* ‘sweetie’ (*matók* ‘sweet’), *'ésti* ‘Esther’ (*Éster~Estér*), *xamúdi* ‘cutie’ (*xamúd* ‘cute’), *tipšóni* ‘silly little fool’ (*tipšon* ‘little fool’), *xáni* ‘Hanna’ (*Xána*), *li'óri* ‘Lior’, *Siváni* ‘Sivan’, *Yósi*, *Yoséfi* ‘Joseph’ (*Yósef~Yoséf*), *nó'ami~no'ámi* ‘Noam’ (*Nó'am*)

Unlike the four functions of the suffix *-i* listed above which existed in classical Hebrew and continue to be used productively in the Modern language, this affectionate suffix is a new phenomenon and seems to have been adopted from a similar usage in various European languages (see for instance Beard & Volpe 2005; Stump 2005; Dressler & Merlini Barbaresi 1994; Jespersen 1965: 215–220). Moreover, it is mainly attached to personal names which are non-integrated Hebrew words. Its stress follows the non-Hebrew stress pattern, in a similar way to other non-integrated Hebrew words (Schwarzwald 2013).

The unstressed suffix *-i* in this function is different from the original diminutive-affectionate Hebrew suffixes *-on* and *-it* which retain the stress on these suffixes (e.g. *sus-ón* 'small horse' [<*sus* 'horse'], *dugm-ít* 'small sample' [<*dugma* 'example, sample']). The stress in words with suffix *-i* which mean affection and diminution behaves in a similar way to the other penultimately stressed loan suffixes given in (4) above.

3. Conclusion

The distribution of the homonymic suffix *-i* shows that the unmarked stress is ultimate for the following functions: two inflectional suffixes (second person feminine in verbs and first person in non-verbs), derivational adjectives, and some derivational gentilic substantives. The suffix *-i* is always unstressed in the formation of affectionate nouns and in most gentilic new nouns derived from loan place names where the stress remains on the stem.

The inflectional second person feminine singular verbal suffix *-i* is mostly stressed ultimately; the penultimate stress can be determined by a morphological rule, which is common to other vowel suffixes in the verb system. The inflectional first person suffix *-i*, either possessive or pronominal is always stressed, excluding the prepositions *kmo* and *min* and four rarely used closed class words in which an *-n-* is inserted before the suffix in which the stress is penultimate. These few exceptions are residues of historical processes which are no longer active in Modern Hebrew and therefore need to be lexically marked and memorized by rote.

The stress of the two derivational *-i* suffixes, adverbial and gentilic, are determined by different factors. The adjectival *-i* suffix stress is determined by stem profile: the suffix *-i* is stressed in Hebrew stems which are the majority; in non-integrated stems, mainly loan words and acronyms, the stress remains on the stem. Namely, words are recognized by the speakers as original and non-original according to phonological, morphological and syllabic considerations (Schwarzwald 2013).

[24] Diminution and affectionate formation involves other phonological phenomena such as vowel alternations or vowel truncation which are discussed elsewhere, e.g. Benjamin [binyamin] > *béni*, Abraham [avram] > *'ávi* or *rami*, Esther ['ester] > *'esti*, Shoshana < *šóši* (Bat-El 2004).

The gentilic final stressed *-i* typifies both nouns and adjectives inherited from classical periods of Hebrew or formed before the 1950s. Modern gentilic nouns derived from names of foreign countries or place names are penultimately stressed as are other loan words. Stress assignment is therefore determined by historical and sociological factors. Some ultimately stressed gentilic nouns in their feminine form must be accounted for by a lexical rule (L3).

Homonymy in affixes is not exclusively typical to Hebrew, as it occurs in other languages as well. For example, the suffix *-s* in English resembles the suffix *-i* in Hebrew in that it serves as an inflection for different word classes (Plag, Homann & Kunter 2017).[25] It functions as the plural noun marker, as the third person singular in the present tense, and as noun possession. Two characteristics distinguish these suffixes in English and in Hebrew:

a. The English suffix *-s* behaves phonetically in the same way for all its inflectional functions, e.g. *dogs*, *gives*, *dog's* [*-z*]; *cats*, *cuts*, *cat's* [*-s*]; *bosses*, *judges*.V, *boss's* [*-əz* or *-ɪz*].[26] The Hebrew suffix *-i* behaves differently for each inflectional function because its stress varies for different reasons, as demonstrated above in sections 2.1 and 2.2.
b. The English suffix *-s* is only used for inflection whereas the Hebrew suffix *-i* is also productively used for derivation.

The suffixes *-ed* and *-ing* in English bear more resemblance to the Hebrew suffix *-i* in that they are both used for inflection and derivation (Jespersen 1965: 377–380, 425–435). Although they are part of the verbal system, independent adjectives and nouns can be derived from them, e.g. *headed*, *coded*, *specified*, *shaped*, *attested*; *interesting*, *setting*, *findings*, *meeting*, *meaning*, *building*. Whereas these English suffixes are always unstressed, the suffix *-i* is primarily stressed in Hebrew.

The suffixes *-ness*, *-ity*, *-ment*, and *-ation* form abstract nouns in English, the suffixes *-er* and *-or* create agent nouns, and the suffix *-al* creates adjectives. Hebrew uses a variety of devices in order to form words, the two most prominent of which are (a) consonantal root and template, and (b) stem linear derivation with an affix. Verbs are only formed by root and template. Nouns and adjectives are produced by either way as well as by compounding, blending, and base formation. As mentioned earlier, adjectives, diminutive and affectionate expressions are formed in a variety of ways in addition to the stem+*-i* formation.

Multiple functions of specific affixes are not unique to any language. It is the stress assignments and its occurrence in each case that is special and intriguing to Hebrew.

References

Avioz, Chagit. 2004. Modern Hebrew Number and Possessive Inflection in Nouns. Ramat Gan: Bar Ilan University. (Ph.D. Dissertation.) [Hebrew]

[25] Plag, Homann & Kunter (2017) added an acoustic study of the non-morphemic abbreviation of *is* and *has*.to the above suffixes.

[26] Cases like *Jesus' house* are excluded because the suffix is not added.

Bat-El, Outi. 1993. Parasitic metrification in the Modern Hebrew stress system. *The Linguistic Review* 10. 189–210.

Bat-El, Outi. 2004. Parsing forms with identical consonants: Hebrew reduplication. In Ravid, Dorit & Bat-Zeev Shyldkrot, Hava (eds.), *Perspectives on Language and Language Development*, 25–34. Dordrecht: Kluwer.

Bat-El, Outi. 2005. The emergence of the trochaic foot in Hebrew hypocoristics. *Phonology* 22. 1–29.

Bat-El, Outi. 2018. Hebrew stress: Back to the future. *Acta Linguistica Academica* 65. 3–27.

Bauer, Laurie. 1996. No phonetic iconicity in evaluative morphology. *Studia Linguistica* 50. 189–204.

Beard, Robert & Volpe, Mark. 2005. Lexeme–Morpheme Base Morphology. In Štekauer, Pavol & Lieber, Rochelle (eds.), *Handbook of Word-Formation*, 189–206. Dordrecht: Springer.

Bolozky, Shmuel. 1999. *Measuring Productivity in Word Formation: The Case of Israeli Hebrew*. Leiden: Brill.

Bolozky, Shmuel. 2013. Diminutive. In Geoffrey Khan et al (eds.), *Encyclopedia of Hebrew Language and Linguistics*, Vol 1, 731–738. Leiden: Brill.

Dressler, Wolfgang U. 2005. Word-Formation in Natural Morphology. In Štekauer, Pavol & Lieber, Rochelle (eds.), *Handbook of Word-Formation*, 267–284. Dordrecht: Springer.

Dressler, Wolfgang U. & Merlini Barbaresi, Lavinia. 1994. *Morphopragmatics: Diminutives and Intensifiers in Italian, German, and Other Languages*. Berlin Mouton de Gruyter.

Dressler, Wolfgang U. & Merlini Barbaresi, Lavinia. 2001. Morphopragmatics of diminutives and augmentatives. In Kenesei, István & Harnish, Robert M. (eds.), *Perspectives on Semantics, Pragmatics, and Discourse*, 43–58. Amsterdam: John Benjamins.

Dubnov, Keren. 2000. Synthetic and analytic possessive pronouns related to nouns in Spoken Hebrew. *Hebrew Linguistics* 47. 21–26. [Hebrew]

Efthymiou, Angeliki. 2016. On the polysemy of the Modern Greek prefix *para-*. In Körtvélyessy, Lívia et al. (eds.), *Word-Formation across Languages*, 91–107. Newcastle upon Tyne: Cambridge Scholars Publishing.

Faust, Noam. 2017. The effect of a null pronoun on morpho-phonology: Israeli Hebrew denomyns. *IATL 2017 Proceedings*. https://noamfaust.files.wordpress.com/2017/01/adjectives-iatl-proceeds-post_formatting.pdf (Accessed 2018-10-9)

Graf, Dafna & Ussishkin, Adam. 2003. Emergent iambs: Stress in Modern Hebrew. *Lingua* 113. 239–270.

Hitin-Mashiah, Rachel. 2005. The use of directive -a and directive l- in Modern Hebrew. *Hebrew Linguistics* 55. 15–29.

Jespersen, Otto. 1965. *A Modern English Grammar*, Part VI: *Morphology*. London & Copenhagen: Allen & Unwin & Munksgaard.

Lieber, Rochelle. 2005. English Word-Formation Processes. In Štekauer, Pavol & Lieber, Rochelle (eds.), *Handbook of Word-Formation*, 375–427. Dordrecht: Springer

Melčuk, Igor & Podolsky, Baruch. 1996. Stress in Modern Hebrew Nominal Inflection. *Theoretical Linguistics* 22. 155–194.

Mor, Uri .2017. Prescriptivism, nation and style: The role of nonclassical elements in the stylistic stratification of Modern Hebrew. *Sociolinguistic Studies* 11. 1–20.

Ornan, Uzzi. 1997. The structure of a dictionary and the place for the adjective. *Helkat Lashon* 25. 13–33. [Hebrew]

Plag, Ingo & Homann, Julia & Kunter Gero. 2017. Homophony and morphology: The acoustics of word-final *s* in English. *Journal of Linguistics* 53(1). 181–216.

Podolsky, Baruch. 1981. Stress as a morphemic factor in Modern Hebrew. *Lešonénu* 45. 155–156. [Hebrew]

Podolsky, Baruch. 1991. The problem of word accent in Modern Hebrew. In Mukarovsky, Hans G. (ed.). *Proceedings of the Fifth International Hamito-Semitic Congress*, 2 (Beiträge zur Afrikanistik, 41), 277–281. Wien: Unstitute für Afrikanistik und Agyptologie de Universitat Wien.

Ravid, Dorit & Chahana-Amitay, Dalia. 2005. Verbal and nominal expressions in narrating conflict situations in Hebrew. *Journal of Pragmatics* 37. 157–183.

Rosén, Haiim B. 1956. *Our Hebrew*. Tel Aviv: Am Oved. [Hebrew]

Schwarzwald (Rodrigue), Ora 1998a. Adjective Formation, Derivation and Inflection. *Helkat Lashon* 27. 9–19. [Hebrew]

Schwarzwald (Rodrigue), Ora. 1998b. Word Foreignness in Modern Hebrew. *Hebrew Studies* 39. 115–142.

Schwarzwald (Rodrigue), Ora. 2001. *Modern Hebrew*. München: Lincom Europa.

Schwarzwald (Rodrigue), Ora. 2004. Modern Hebrew consonant clusters', In Ravid, Dorit & Bat-Zeev Shyldkrot, Hava (eds.), *Perspectives on Language and Language Development*, 45–60. Dordrecht, Kluwer.

Schwarzwald (Rodrigue), Ora. 2013. The Typology of Nonintegrated Words in Hebrew. *SKASE Journal of Theoretical Linguistics* 10(1). 41–53. (http://www.skase.sk/Volumes/JTL22/index.html)

Stump, Geoffrey T. 2005. Word-Formation and Inflectional Morphology. In Štekauer, Pavol & Lieber, Rochelle (eds.), *Handbook of Word-Formation*, 48–71. Dordrecht: Springer.

Werner, Fritz. 1983. *Die Wortbildung der hebräischen Adjektiva*. Wiesbaden: Otto Harrassowitz.

Ora (Rodrigue) Schwarzwald
The Department of Hebrew and Semitic Languages
Bar-Ilan University
Ramat Gan, Israel 52900
Ora.schwarzwald@biu.ac.il

In SKASE Journal of Theoretical Linguistics [online]. 2018, vol. 15, no. 3 [cit. 2018-13-12]. Available on web page http://www.skase.sk/Volumes/JTL38/pdf_doc/06.pdf. ISSN 1336-782X

Nomina loci in Beserman Udmurt: elicitation, corpus study, and experiments[1]

Maria Usacheva
Russian Academy of Sciences/Pushkin State Russian Language Institute, Russia

This paper discusses the advantages and disadvantages of elicitation, corpus study, and experimental study in field research. -on'n'ig-forms in Beserman Udmurt which function both as nomina loci (place names) and as converbs are described with the focus on methodology of gathering the data. It is demonstrated that in the case of productive derivation hypotheses formulated on the base of corpus study can more reliably be checked during an experiment with non-verbal stimuli than via elicitation. As for morphosyntactic properties of regular inflected forms, it is easier to study them on the base of corpora data and elicitation. However, experiments provide both natural examples and interesting data for future research.

Keywords: *field linguistics, elicitation, experiments, corpus study, Udmurt*

The most popular and widespread methods of linguistic fieldwork are elicitation and recording spontaneous texts (typically narratives). Transcribed, glossed and translated texts can be organized into a corpus. On the base of natural examples taken from the corpus, a linguist can formulate a hypothesis and later check it during the fieldwork via elicitation. For investigations of grammar it seems to be sufficient. Nevertheless, in recent decades field linguistics and typology have begun to gather the data with the help of non-verbal stimuli. A wide series of projects based on experiments has been realized in Max Planck Institute of Psycholinguistics. Experiments were conducted in order to gather comparable data from different languages. The projects were devoted to space (Bowerman & Pederson 1992; Bowerman et al. 2004), to time in space (Boroditsky et al. 2008), to expressives (Tufvesson 2007), etc. The idea to study encoding of spatial relations with the help of non-verbal stimuli is very fruitful. It was realized in typological projects (Fortis et al. 2009) and in descriptions of endangered languages (Arjava 2016). Experiments are conducted during field studies of syntax (Polinsky: preprint) and semantics (Arunachalam & Kothari 2011; Majid 2012). Apart from spatiality, non-verbal stimuli are relatively frequently used in studies of possessives (Klumpp 2017; Eisenbeiss 2017a), NP structure (Kozlov et al. 2016; Arkhangelskiy & Usacheva to appear) and ergativity (Longenbaugh & Polinsky 2017). There are special linguistic courses in universities devoted to experiments in the field (Eisenbeiss 2017b).

Why do experimental methods attract more and more attention in field linguistics and typological studies? The main reason seems to be obvious: every method has its limitations, and both elicitation and corpus study are not exceptions in this respect. To demonstrate it, we will describe *-on'n'ig*-forms in Beserman Udmurt.[2]

Beserman is an unwritten dialect of Udmurt. It is spoken by ethnic Besermans who live in North-Western Udmurtia (Russian Federation). According to the 2010 Census, there are ca.

[1] The work was supported by RSF grant 18-18-00462 *Communicative-syntactical interface: typology and grammar* realized in Pushkin State Russian Language Institute.

[2] The present work is devoted to methodological account of gathering data in the field. More information about particular morphological and syntactic properties of the *-on'n'ig*-forms can be found in Usacheva & Serdobolskaya (2015).

2200 people identifying themselves as Beserman. There are 10 predominantly Beserman villages and 41 villages with a mixed population. Most of Besermans are bilingual in Russian and are familiar with Standard Udmurt. The dialect differs from the literary language in phonetics, morphosyntax and lexicon. In particular, the form described in the present article is attested only in Beserman and is used in neither of the other Udmurt dialects (Kelmakov 1998: 297).

All the examples cited in the present article were collected in the village of Shamardan (Yukamenskoye district, Udmurtia) in 2003–2018. Most of the data were taken from the Beserman corpus of oral speech (ca. 67000 tokens), the Multimedia Beserman corpus (a corpus with sound and video; ca. 38000 tokens) and from the Corpus of examples from the on-line Beserman-Russian dictionary (ca. 83000 tokens). The Beserman corpus of oral speech and the Multimedia Beserman corpus contain examples of spontaneous and quasi-spontaneous speech. Spontaneous texts include recordings of every-day communication (dialogues and polilogues), narratives, receipts, interviews, etc. Quasi-spontaneous speech was recorded during experiments of different kinds: referential communication tasks, retellings of cartoons' plots or of texts, word games, and serial reproduction. The corpus of examples from the on-line Beserman-Russian dictionary is predominantly composed of sentences got during lexicographic work. These sentences were mostly provided by Beserman speakers when we asked them to create an example where a concrete lexeme from the dictionary had to be used. Some of them are the result of a translation from Russian into Beserman, but these translations were made with a focus on lemmata from the dictionary, not on grammar.

When we searched for *-on'n'ig*-forms, we found 19 sentences in the oral corpora and 74 sentences in the corpus of examples. These examples gave us much information about the semantics and syntax of the forms in question. As for semantics, corpora examples show that the *-on'n'ig*-forms can denote place of an action (1)[3] and a process which is parallel to the action denoted by the main verb (2):

(1) *Korka-z-ə̂* *kwaška,* ***ul-on'-n'ig-z-ə̂*** *evə̂l*.
house-P.3-PL be.ruined.PST live-VN-NLOC-POSS.3-PL be.NEG
'Their house has been ruined, they have no place to live.'

(2) *Mon* ***n'umal*** ***kušman-ez*** ***bud-ə̂t-on'-n'ig-ə̂n***
I.NOM sweet beet-ACC.3SG grow-CAUS-VN-NLOC-LOC
baš't-i *pervoj* *mesta*.
take-PST.1SG first place
'I have won the competition of growing sugar beet.'

The forms in question are also used in *when/while*-clauses if the places of the two actions coincide:

(3) *So* ***dugd-on-n'ig-a-z*** *mon-e* *l'oga-m*.
that stop-VN-CVB-LOC-POSS.3 I-ACC trample.down-PST2[4]
'He drove over me while he was stopping.'

[3] All the examples cited in the article belong to Beserman Udmurt.
[4] PST2 is a label used for a past tense which also expresses a range of meanings like perfectivity or evidentiality.

(4) *Gožj-an* *mân-e* ***č'up-iš'k-on'-n'ig-e***
draw-VN go-3SG.PRS make.blockhouse-VN-NLOC-ILL
gožja-nâ.
draw-INF
'A tap is used for drawing while making a blockhouse.'

Corpora examples also show that the *-on'n'ig*-forms demonstrate both nominal and verbal properties. It is very common for converbs, but quite surprisingly for place names. One of verbal properties is the ability to attach spatial adjuncts:

(5) *Ul'č'a-jân* *so-os-len* *traktor-z-â,*
street-LOC that-PL-GEN1[5] tractor-POSS.3-PL
es-t'i ***pot-on-n'ig-ân,*** *ken'er es* *uš't-emân.*
door-PROL go.out-VN-NLOC-LOC fence door close-RES
'Their tractor is on the street, on the exit through the door, the fence's door is closed.'

Next, *-on'n'ig*-forms conserve direct object which can be unmarked or bear an accusative marker depending on the referential status of the object. Such split is typical for direct objects of verbs in Beserman (as in other Finno-Ugric idioms). Thus, objects in (6) and (7) are non-specific (generic) and therefore unmarked, whereas a specific definite object in (8) bears the accusative marker:

(6) ***N'an'*** *vož'-on'-n'ig-a-d* *n'an'* *kema*
bread store-VN-NLOC-P.LOC-POSS.2 bread long
vož'-i-d=ke, *pâk* *zân* *lu-e.*
store-PST-2.SG=if musty smell become-PRS.3.SG
'If you store bread in the bread basket too long, musty smell occurs.'

(7) *Piroški* *pâž-on* *dâr-ja* ***ul'*** ***n'an'***
patty bake-VN time-ADV unleavened dough
leš't-on'-n'ig-e *pun-iš'ko-d* *vu,* *piž',* *slal.*
make-VN-NLOC-ILL put-DETR-2.SG water flour salt
'While baking pattys, while making unleavened dough put water, flour, salt.'

(8) ***Č'orog-ez*** *pot-on'-n'ig-a-m* *žilka-ez*
fish-ACC go.out-VN-NLOC-P.ILL-POSS.1.SG line-POSS.3.SG
č'ig-i-z.
rip-PST-3.SG
'While I was pulling out the fish (mentioned before), the fish line ripped.'

The forms which head *when/while*-clauses agree with nominal subject in possessive:

[5] In Permic languages case marking of nominal dependents in noun phrases is defined by syntactic position of these phrases. Dependents of nouns in direct object position attach markers of case labeled here as GEN2. Dependents of nouns in other syntactic positions are marked by GEN1. Genitive 2 is a former spatial case which now performs different other functions like marking material, the point of comparison etc.

(9) *Č'igan'ka* ***tunač'k-on'-n'ig-a-z***
gypsy.woman tell.fortunes-VN-NLOC-P.ILL-POSS.3.SG
so-os-lâš' *luška-z* *kon'd'on-z-e*.
that-PL-GEN2 steal-PST.3.SG money-3SG-ACC
'A Gypsy woman stole their money while telling fortunes.'

As one can see from (4) above, there is no agreement with generic zero subject. Thus, in (4) there is no possessive marker on the *-on'n'ig*-form:

(4) *Gožj-an* *mân-e* ***č'up-iš'k-on'-n'ig-e***
draw-VN go-3SG:PRS make.blockhouse-VN-NLOC-ILL
gožja-nâ.
draw-INF
'A tap is used for drawing ile making a blockhouse.'

Examples (3) and (9) above show that *-on'-n'ig*-forms which head *when/while*-clauses attach both nominal and pronominal subjects in nominative. In phrases headed by nouns a pronominal dependent must be in genitive.

As for nominal properties, one can see that *-on'-n'ig*-forms which denote places of actions (10) and processes (11)–(13) can be used as adnominal dependents or dependents of relational nouns:

(10) *Mil'am* *korka-n* ***už'-on'-n'ig***=*no* *gur až' pal*
we.GEN1 house-LOC dream-VN-NLOC=ADD kitchen
st'ena *jun* *vekč'i*.
wall very thin
'The wall between the bedroom and the kitchen in our house is very thin.'

(11) *Č'ašja* *imen'n'ik* *nunal* *kalâk-jos* *vel't-âl-i-z-â*
forest birthday day people-PL go-ITER-PST-3-PL
š'i-jon-en-ju-on-en *č'ašja-je,* ***kuja-š'k-on'-n'ig***
eat-VN-INS-drink-VN-INS forest-ILL throw-DETR-VN-NLOC
č'ašša-ja-z *kert-âl-i-z-â* *basma* *pum*.
forest-ILL-POSS.3.SG tie-ITER-PST-3-PL cloth rag
'In Forest's day people used to go to the forest with food and drinks, they tied a rag of cloth in a sacred forest (lit. in the forest where one throws sacred objects).'

(12) *Trak* *š'er-ân* ***život*** ***kuj-an'-n'ig*** ***mesta***.
big.road behind-LOC livestock throw-VN-NLOC place
'Behind the big road there is a cattlegrave.'

Corpora data also show that *-on'-n'ig*-forms denoting place names attach nominal subjects in nominative or in genitive. Case marking of the subjects depends on their referential status. The phrase *kureg* 'hen' in (13) is non-specific and hence unmarked. *Kârban* 'a Beserman fest devoted to the end of sowing' in (14) is marked by genitive because it is opposed to other Beserman fests in the context, and this opposition makes it specific:

(13) *Mânam kureg gid'-ân* ***kureg*** *pukš'-on'-n'ig-e* *wan'*.
I.GEN1 hen shed-LOC hen sit-VN-NLOC-POSS.1.SG be.PRS
'There are three henroosts (lit. places where hen sit) in my coop.'

(14) *A so Kârban dâr-ja,* ***Kârban-len***
and that Kârban time-ADV Kârban-GEN1
kuč'k-on-n'ig *pal-a-z* *lu-âl-i-z* *pop.*
begin- VN-NLOC side-ILL-POSS.3.SG become-ITER-PST-3.SG priest
'And just before the beginning of that Kyrban [fest] the priest came.'

The same condition of referential status determines grammatical marking of dependents in phrases headed by nouns.

While studying particular grammatical phenomena we usually also take into account the data from corpora and grammars of genetically related idioms, of languages of the area and (working with dialects) of the literary variety. In the case of the *-on'n'ig*-forms, searching the corpora of other Udmurt varieties gave no results since such forms are attested neither in other dialects nor in Standard Udmurt (Kelmakov 1998: 297). There is a suffix of nomina loci *-on'n'i-*, but it does not have process meaning. *When/while*-clauses in other Udmurt varieties are formed by forms on *-ku*, *-onja-*/*-on'n'a-* (Georgieva 2017; the *-on'n'a*-forms also exist in Beserman), in Sharkan Udmurt also by forms on *-onna-* (Timofey Arkhangelskiy, p.c.). We also have not found the strict analogues of the *-on'n'ig*-forms neither in genetically related Komi varieties nor in Turkic languages (Tatar and Chuvash) spoken by people who live close to or together with Besermans and Udmurts. In Zyrian *when/while*-clauses can be formed by converbs with suffix *-ig-*(*en*) which gives the base for suggestions about the etymology of the *-on'n'ig-* complex (Usacheva & Serdobolskaya 2015).

However, the corpus data do not contain all the information about morphological and syntactic properties of the *-on'n'ig*-forms. It is a typical situation, especially for the corpora of endangered unwritten idioms because such corpora are very small in comparison with the corpora of idioms with a literary tradition: usually they contain no more that several hundred thousands of tokens. For example, we can say nothing about ability of the *-on'n'ig*-forms to syntactically behave like nouns when used as converbs on the base of corpora. We also do not know if the forms in question can attach the nominal plural marker, if they can be modified by adverbs and/or by adjectives, etc. For gathering this information we either have to wait while the constructions we need will occur in the corpus or have to use the elicitation method.

There are also questions that can be answered neither by means of corpora analysis nor by elicitation. They are:

1. How productive are the forms in their function of denoting place names? Do they denote only fixed locations where the process usually takes place (i.e. are they lexicalized or not neccesarily)?
2. Are the *-on'n'ig*-forms which head *when/while*-clauses preferred to other converbs also used in this function (to *-on'n'a*-forms, for instance) when the attention of the speaker is focused both on the location and on the process?

In order to answer these questions, I conducted an experiment with non-verbal stimuli. I formulated the following hypotheses based on the corpus data:

(i) *-on'-n'ig*-forms functioning as nomina loci are productive. The complex *-on'-n'ig-* can attach to a verbal stem denoting any process.
(ii) *-on'-n'ig*-forms which head *when/while*-clauses are preferred when the place of the process is focused.

When a field linguist is planning an experiment, it is very useful to take into account the genre of the texts he wants to get as a result. It is known that texts of different genres may contain different grammatical forms and constructions. Historical present tense in English, for example, is used in narratives (both oral and written) and not in dialogues (see Wolfson (1989) and the literature cited there); Russian imperfective aspect is used to encode general validity primarily in genres other than narratives (Sitchinava 2011), etc. Before we have shown that *-on'n'ig*-forms can denote places where something happens or processes; they can also head *when/while*-clauses. Place names occur in texts of different genres whereas names of processes and especially *when/while*-clauses are more likely to be found in narratives. Therefore, we decided to use one of the most popular techniques of stimulating narratives – a series of pictures which represent a story. In order to avoid problems with items from a culture the speakers are not familiar with we took a plot of a Russian folktale which Beserman speakers definitely know – *Ivan tsarevitch, the firebird, and the grey wolf*. In order to check the hypothesis 1, we included both images of locations which are usually described by *-on'n'ig*-forms (the stick hens are sitting on, the place in the yard where horses are usually tied, etc.) and images with actions which are not associated with fixed locations (spinning wool, coloring fence, etc.). We also tried to attract the attention to locations where the actions take place by drawing the roads leading to people or animals performing different activities. Here are some of the pictures we used during the experiment:

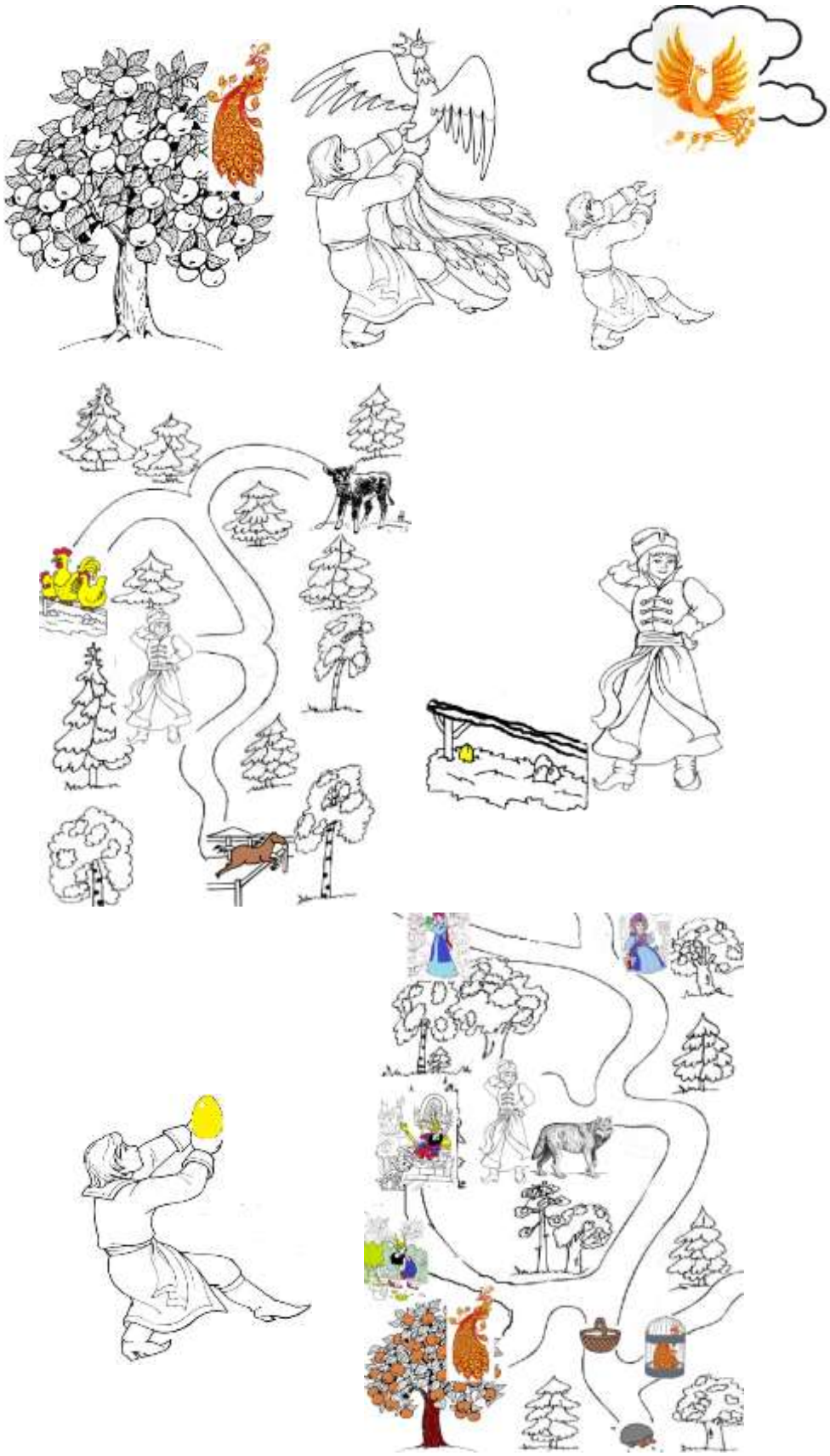


Figure 1: The pictures used in the experiment with cards

We put the cards in front of Beserman speakers and asked them to tell a story represented on the pictures. If hypothesis 1 was correct, we would get *-on'-n'ig*-forms of verbs denoting activities not associated with concrete locations. If hypothesis 2 was correct, we would get much more *-on'-n'ig*-forms denoting processes than in corpora.

Actually, the typical problem with experiments in linguistics is the following. If the experiment is successful (and well-designed), it proves or declines the initial hypothesis which was intended to test. But if the experiment fails, it tells nothing about the hypothesis because there can be many different reasons for this particular experiment to be unsuccessful. For the experiment we conducted with cards there were at least three such reasons. First of all, the pictures were drawn badly. Some of the probationers could not even recognize the story and the heroes. It would be much better if the pictures were drawn by a professional artist. Secondly, the pictures with roads were too complex. There were too many actions on them, and the probationers were confused. Thirdly, speakers of Permic languages I work with have problems with telling stories based on a sequence of cards representing events. Every card is taken as referring to a single event independent from others, and it is very difficult for my probationers to apprehend a sequence of cards as representation of a coherent story. Therefore, as the experiment failed and I could neither prove nor reject my hypotheses, I decided to try a different design. During the second experiment I asked Beserman speakers (five women and one man, all under forty years old) to play a board game designed by me and drawn by the professional artist Tatyana Panova. As one can see from Figure 2, the game represented traditional Beserman activities, ceremonies, fests and frequent problems of life.

Figure 2: The board-game designed for the experiment

Figure 3: The figures and the dice used during the experiment with the board game[6]

The game was organized as follows. There were two speakers playing. They were asked to comment what they were doing. During the game the two figures moved along different paths. One of the probationers took the figure of a girl, the other took the figure of a boy. The probationers were throwing a dice by turns and then moved their figures according to the number that appeared on the dice. When a figure stood on a circle with a red or a dark-blue number, the probationers had to put a figure on a picture with a special activity (gathering butter mushrooms, catching pigeons to heal stuttering, etc.). Then the figure was returned to the number it stayed before.

During the experiment we recorded four texts (two texts were recorded from the same pair of speakers in different days). The texts in sum last ca. 4 hours; the sub-corpus we made on their base includes approximately 21000 tokens. This experiment cannot be treated as fully successful because instead of descriptions of what the speakers were doing while playing we got narratives about traditional Beserman activities. Nevertheless, for forms functioning as nomina loci the experiment was successful. It has proved that hypothesis 1 is correct because experimental texts contain examples like (15):

(15) *Odig pol mânam podruga-je lâkt-i-z. ǯ'eč'era-nâ*
one time I.GEN1 friend-POSS.1.SG come-PST-3.SG teeter-INF
pukš'-i-mâ. ǯ'eč'era-š'ko-mâ tare olo-malâ mi
sit-PST-1PL teeter-DETR-1PL then INDF-why we.NOM
kâk-na-mâ ***ǯ'eč'era-n'-n'ig-iš't â-mâ*** *uš'-i-mâ.*
two-COLL-1PL teeter-VN-NLOC-P.EL-1PL fall.down-PST-1PL
'Once my friend came to me. We began teetering, and then we both fell down from the place where we were doing it.'

[6] Picture 3 was taken by Nickolay Philippov.

The speaker is talking about Easter. During the celebration of this fest, Besermans of Shamardan (the village where the experiment was conducted) make big teeters and are teetering on them half of the day. The teeters are made every year in different locations and then removed, so the *-on'n'ig*-form used in (15) definitely refers to the place of the action described. It cannot refer to the action itself (to the process of teetering) because in this function the form can attach only illative marker – not the elative, as in (15). For the teeters itself there is a special word different from the *-on'n'ig*-form.

In addition, Table 1 demonstrates that the share of the *-on'n'ig*-forms denoting place names in experimental texts is about 4.5 times more than the share of these forms in the spoken corpora. These forms are more frequent in the corpus of dictionary examples than in the experimental texts, but the reason is that the lexicon of the dictionary contains a good deal of lexicalized place names with several examples for each such name.

Table 1: The share of *-on'n'ig*-forms in the existing Beserman corpora and in the corpus recorded during the experiment.

	Spoken corpora	Corpus of examples	Experimental corpus
-on'n'ig- NLoc	9 of 105000 (0.86 for every 10000 tokens)	57 of 83000 (6.87 for every 10000 tokens)	8 of 21 000 **(3.81 for every 10000 tokens)**
-on'n'ig- process	10 of 105000 (0.95 for every 10000 tokens)	17 of 83000 (2.05 for every 10000 tokens)	2 of 21000 (0.95 for every 10000 tokens)

Therefore, we treat the experiment as successful for the *-on'n'ig*-forms denoting place names. As for these forms functioning as converbs, in experimental texts there are only two examples like (16) containing them:

(16) *Mon žad'-i* ***iskil'l'a-n'-n'ig-a-m***
I.NOM get.tired-PST.1SG sledge-VN-NLOC-P.ILL-POSS.1.SG.
pâr-i *muč'o-je* *šâdeč'ik-ânâ* *âžât* *otân*
enter-PST.1.SG bath-ILL have.rest-INF for.a.while there
puk-o.
sit-FUT.1.SG
'I got tired while sledging and entered the bath. I'll sit there for a while and have a rest.'

But data from all the corpora show that another strategy used in *when/while*-clauses – constructions formed by verbal nouns with suffix *-on* and the postposition *dârja* 'while' – is much more frequent than that of using constructions with *-on'n'ig-* (see Table 2).

Table 2: Absolute and relative frequency of *-on'n'ig*-converbs and alternative strategies in the existing Beserman corpora and in the corpus recorded during the experiment

	-on'n'ig 'while'	*-on + dârja* 'while' (the same subject)	*-on'n'a-* 'while'	*ku* 'while' (Standard Udmurt)
Experimental corpus	2 (20%)	**6 (60%)**	1 (10%)	1 (10%)
Spoken corpora	10 (30%)	**18 (55%)**	4 (12%)	1 (3%)
Corpus of examples	17 (9%)	**139 (77%)**	24 (14%)	0 (0%)

Three possible explanations can be offered for this situation: either hypothesis 2 is wrong; or the *-on'n'ig*-forms are pushed out by alternative constructions and are therefore too rare to be captured during experiments; or the design of the experiment is unsuitable for stimulating *when/while*-clauses. In order to choose the right one we computed the frequency of the occurences of *-on + dârja* 'while' in all the corpora and in the experimental texts. We took into account only the cases when the subjects of the two actions coincide because it is obligatory for the *-on'n'ig*-forms in converbal function. One can see from Table 2 that the share of construction *-on + dârja* in the experimental corpus is higher than in spoken corpora, but not very much: 2.85 vs 1.7 for every 10000 tokens. It means that the design was bad, and to test the hypothesis 2 one should make another experiment to get descriptions of parallel actions.

In addition, we would like to say that even unsuccessful experiments are very good for a field linguist because, as a rule, they result in long vivid texts which are not always simple to record in the field. The narratives we got during the experiment with a board game are also very interesting from the ethnographic point of view. Besides, experimental texts often give interesting data concerning different grammatical phenomena. For example, just one 45-minutes dialogue recorded during the experiment described above contains 11 examples like (16) – with combinations of iterative and resultative suffixes in a verb:

(17) - *A vot mil'am abi-len…* *vera-l'l'a-z*
and here we.GEN1 grandmother-GEN1 tell-PST2-3.SG
tod-iš'ko-d=a *mugor-a-z* *so-len* *lu-âl-i-z*
know-PRS-2.SG=Q body-LOC-POSS.3.SG that-GEN1 become-ITER-PST-3.SG
marâm-eš' *p'atno-os* *lâz-eš'*.
HES-PL.ADJ spot-PL blue-PL
- ***Č'epel't-âlâ-mân*** *kad'*.
pinch-ITER-RES like
- *Ben, tin'* *š'uš'etka=pe* *č'epel't-em* *uj-ân*.
yes here brownie=QUOT pinch-PTCP.PST night-LOC
'- You know, our granny's… [she] said that blue spots appeared on her body. - Like pinched several times. - Yes, the brownie supposedly pinched her at night.'

There are only 20 examples of that kind in the oral Beserman corpora, and they do not include classes of verbs represented in texts recorded during the experiment described. Elicitation also did not give reliable results for possibility of such verbs to attach combinations of resultative

and iterative. Therefore, the experiment with the board game gave very important data concerning another, a fully unrelated grammatical theme.

In the present article we tried to show that all three methods used in field linguistics – elicitation, corpora and experiments – have their limitations. It seems rational to combine them. The contribution of each method depends on particular language phenomena. Experiments are useful even if they are unsuccessful, but they must be designed accurately, and the genre of intended texts must be taken into account since different techniques should be used to stimulate texts of different genres. I propose the following workflow for field linguists. First, one can search the corpus of the idiom in question and the corpora of idioms which are genetically or geographically close to it. On the base of the data taken from the corpora, one can formulate a hypothesis and try to test it by experiments. After it, the gap in the data can be filled with the help of elicitation.

References

Arjava, Heini. 2016. Experimenting on spatiality: Elicitation with three-dimensional toys in a field study of Erzya. In Shagal, Ksenia & Arjava, Heini (eds.), *Mordvin languages in the field* (Uralica Helsingiensia 10), 319–338. Helsinki.

Arkhangelskiy, Timofey & Usacheva, Maria. To appear. Case compounding in Beserman Udmurt. *Eesti ja soome-ugri keeleteaduse ajakiri* [Journal of Estonian and Finno-Ugric Linguistics] 9.

Arunachalam, Sudha & Kothari, Anubna 2011. An experimental study of Hindi and English perfective interpretation. *Journal of South Asian Linguistics* 4. 27–42.

Beserman corpus of oral speech http://beserman.ru/corpus/search/?interface_language=ru.

Boroditsky, Lera & Gaby, Alice & Levinson, Stephen C. 2008. Time in space. In Majid, Asifa (ed.), *Field manual* 11, 52–76. Nijmegen: Max Planck Institute for Psycholinguistics. (http://fieldmanuals.mpi.nl/volumes/2008/time-in-space) (Accessed 2017-11-04.)

Bowerman, Melissa & Pederson, Eric. 1992. Topological relations picture series. In Levinson, Stephen C. (ed.), *Space stimuli kit 1.2: November 1992, 51*. Nijmegen: Max Planck Institute for Psycholinguistics. (http://fieldmanuals.mpi.nl/volumes/1992/bowped/) (Accessed 2017-11-04.)

Bowerman, Melissa & Gullberg, Marianne & Majid, Asifa & Narasimhan, Bhuvana. 2004. Put Project: The cross-linguistic encoding of placement events. In Majid, Asifa (ed.), *Field manual* 9, 10–24. Nijmegen: Max Planck Institute for Psycholinguistics. (http://fieldmanuals.mpi.nl/volumes/2004/put-project) (Accessed 2017-11-04.)

Eisenbeiss, Sonja. 2017a. *Course on possessive constructions at the University of Cologne*. (https://experimentalfieldlinguistics.wordpress.com) (Accessed 2018-08-28.)

Eisenbeiss, Sonja. 2017b. *Experimental linguistics in the field: Course at the University of Cologne*. (https://experimentalfieldlinguistics.wordpress.com/course-psycholinguistics-and-fieldwork/) (Accessed 2018-08-27)

Fortis, Jean-Michel & Grinevald, Colette & Kopecka, Anetta & Vittrant, Alice. 2009. *Trajectoire*. (http://www.ddl.ish-lyon.cnrs.fr/trajectoire/TUL_TRAJ_141209.pps) (Accessed 2017-11-05.)

Georgieva, Ekaterina. 2017. Converbs in Udmurt – a unified class? (Paper presented at the *2nd Budapest Linguistics Conference*, ELTE, Budapest, 1–3 June 2017.)

Kelmakov, Valey K. 1998. *Kratkiy kurs udmurtskoy dialektologii* [A short course of Udmurt dialectology]. Izhevsk: Izdatelstvo Udmurtskogo Universiteta.

Klumpp, Gerson. 2017. Givenness features: Between zero expression and definite encoding of NPs (based on Komi). (Paper presented at *Conference on the Syntax of Uralic Languages*, Research Institute for Linguistics, Hungarian Academy of Sciences, Budapest, 27–28 June 2017.) (http://www.nytud.hu/soul2017/absz/SOUL2017_Klumpp.pdf) (Accessed 2017-11-05.)

Kozlov, Alexey & Privizentseva, Maria & Pleshak, Polina. 2016. Case compounding and discourse structure in Moksha. (Paper presented at Workshop on Information Structure and Discourse, at the conference on *Language Documentation and Linguistic Theory 5*, London, 2-4 December 2016.)

Longenbaugh, Nicholas & Polinsky, Maria. 2017. Experimental approaches to ergative languages. In Coon, Jessica, & Massam, Diane, & Travis, Lisa deMena (eds.), *The Oxford handbook of ergativity*, 709–736. Oxford: Oxford University Press.

Majid, Asifa. 2012. A Guide to Stimulus-based Elicitation for Semantic Categories. In Thieberger, Nicolas (ed.), *The Oxford Handbook of Linguistic fieldwork*, 54–71. Oxford: Oxford University Press.

Polinsky, Maria. Experimental syntax and linguistic fieldwork. Preprint. https://www.academia.edu/35419250/EXPERIMENTAL_SYNTAX_AND_LINGUISTIC_FIELDWORK

Sitchinava, Dmitri. 2011. Nesovershennyj vid [Imperfective]: Materials for the project of corpus description of Russian grammar (http://rusgram.ru). Moscow. (Manuscript.) (http://test.rusgram.ru/taxonomy/term/48#313) (Accessed 2018-08-26.)

The Multimedia Beserman corpus http://linghub.ru/beserman_multimedia/search.

The Corpus of examples from the on-line Beserman-Russian dictionary http://linghub.ru/beserman_examples_corpus/search.

Tufvesson, Sylvia. 2007. Expressives. In Majid, Asifa (ed.), *Field Manual* 10, 53–58. Nijmegen: Max Planck Institute for Psycholinguistics. (http://fieldmanuals.mpi.nl/volumes/2007/expressives/) (Accessed 2017-11-05.)

Usacheva, Maria & Serdobolskaya, Natalia. 2015. Grammatikalizacia novoj deeprichastnoj formy na -onniga- v besermjanskom udmurtskom [Grammaticalization on a new converb on -onniga- in Beserman Udmurt]. In Lyitikova, Ekaterina & Zimmerling, Anton & Konoshenko, Maria (eds.), *Tipologia morfosintakticheskih parametrov. Materialy mezhdunarodnoj konferencii* "*Tipologia morfosintakticheskih parametrov 2015"*. MPGU. 366–392.

Wolfson, Nessa. 1989. The conversational historical present. *LINX (Linguistique Institut Nanterre-Paris-X)* 20. 135–151.

Abbreviations
ADD – a clitic ’and’
COLL – collective numeral
DETR – detransitive
EL – elative
HES – hesitation
ITER – iterative
ILL - illative
NLOC – verbal noun which denotes location of the action
P.EL – allomorph of the elative case marker which appears before possessive suffixes
P.ILL – allomorph of the illative case marker which appears before possessive suffixes
P.LOC – allomorph of the locative case marker which appears before possessive suffixes
PROL – prolative
VN – verbal noun

Maria Usacheva
143003,
18-247 Odintsovo
Russia
mashastroeva@gmail.com

In SKASE Journal of Theoretical Linguistics [online]. 2018, vol. 15, no. 3 [cit. 2018-13-12]. Available on web page http://www.skase.sk/Volumes/JTL38/pdf_doc/07.pdf. ISSN 1336-782X.

A pragma-sociolinguistic deconstruction of 'dress', 'meet', and 'toast' in Southwestern Nigeria

Temitope Michael Ajayi, University of Ibadan, Ibadan, Nigeria

Meaning in language is a phenomenon that has attracted the attention of language scholars, particularly those in the fields of language such as semantics, pragmatics, sociolinguistics, and discourse analysis. Of central concern to these fields is the role context plays in meaning realisation. This paper examines the influence of the Nigerian socio-cultural context on the semantic realisations of three prominent words, 'dress', 'meet' and 'toast' in the Nigerian context of English language usage, within the purview of Adegbija's pragma-sociolinguistic theory. The study reveals the words, dress, meet, and toast, manifest some senses which are peculiar to the Nigerian L_2 context of English, apart from the senses that the words depict in the English L_1 context. For instance, dress in the Nigerian sociolinguistic environment manifests senses such as to adjust, to slap, and to position. Similarly, 'meet', and toast, alongside all other senses which are applicable in L_1 and Nigerian contexts of English usage, manifest some senses which are peculiar to the Nigerian sociolinguistic environment. Thus, lack of adequate knowledge of the contexts that produce these Nigerian senses of the words in utterances by L_1 and non Nigerian speakers of English could pose a problem of meaning decoding.

Keywords: *English in Nigeria, 'dress', 'meet' and 'toast', pragma-sociolinguistic theory*

1. Introduction

The dynamic nature of language has been emphasised in language studies, especially among sociolinguists, pragmaticists, and discourse analysts. This dynamism of language explains why there exists variation in the semantic realisations of words even among speakers of the same language. In a recent personal interaction with a colleague, for instance, there was a heated debate on the appropriateness or otherwise of the popular question among many Nigerian users of English 'How was your night?'. According to this colleague of mine, the English culture frowns at such a question as it is considered prying into the privacy of the fellow being asked. To my friend, the question is like asking 'how was your sex experience last night?' However, as a sociolinguist, I tried to provide a justification for possible appropriateness of the statement in the Nigerian context, given the fact that the (Nigerian) socio-cultural context that births such a question would not interpret it as asking for one's sexual experience, as it could as well be viewed within the purview of the totality of one's experience during the night, including having a sound sleep, not being visited by armed robbers, among others. This dynamic interpretation of the expression could be hinged on the indigenisation of the language in Nigeria which apparently has birthed 'Nigerian English' variety. Hence, the interpretation of certain English words in Nigeria is subject to different contextual realisations that define the socio-cultural realities in the country. In view of this development, many scholars and observers have clamoured for the recognition of the Nigerian variety of English. However, this agitation has not been pragmatically addressed. The crux of this paper, therefore, is to further engage the impact of the Nigerian socio-cultural environment on English (in the country) usage in the country. In particular, this

study, taking a cue from Adegbija & Bello's (2001) work on the semantic nuances of OK in the Nigerian context, and Adeyanju's (2011) pragma-sociolinguistic dissection of the word SEE in Nigerian English, attempts a pragma-sociolinguistic deconstruction of the words DRESS, MEET, and TOAST among Nigerians, particularly in the southwest. As shall be seen in this paper, the words 'dress', 'meet' and 'toast' are confronted with different contextual interpretations, some of which align with its L_1[1] contextual usage and in some other instances depict the Nigerian L_2 context. This is in line with Adegbija & Bello's (2001:89) submission that:

> With the transplantation of English to new contexts, especially through colonisation, the language is forced to confront new contexts which come to bear on its use in the process of encoding and decoding of meaning.

Thus, some instances of the interpretation of the words 'dress', 'meet' and 'toast' might be difficult if not totally impossible for an English speaker from another socio-cultural background, given their Nigerian peculiarity. This submission reinforces the opinion of Adegbija & Bello (2001: 89) as follows:

> Given the potential depth of all utterances and the variety of human intentions, word knowledge, experiences and motifs in any particular speech situation, it is not surprising that in the addressee's attempt to infer meaning, the actual value of a speakers' intended message may depreciate, or be entirely misconstrued.

This study, as mentioned earlier, focuses on three lexical items (verbs) in the English language. Of course, these are not the only items that manifest the Nigerian sense of the use of English. For instance, the word traffic, a noun, is commonly used in the verb sense among Nigerians as in 'trafficate'. However, the lexical items: dress, meet and toast have been purposively selected for analysis in this study because they are prominent among English lexical items often used in the Nigerian sense. Although these words are part of the lexical items examined by Igboanusi (2010), their various contextual senses, especially as explored in this study, have not been captured.

2. English and its status in Nigeria

English language remains the most enduring legacy the British colonial masters bequeathed to Nigeria, as years after the country attained independence, she has held tightly to the language, particularly in important domains such as education, administration, health, judiciary, among others. As reported by Banjo (1982), the adoption of the language as a language for wider communication dates back to 1862. This was the period when Lagos was formally established as a colony by the British colonial masters. Taiwo (2009) reports that in 1882, in an attempt by the to promote the assimilationist culture, the British colonial government introduced an act that brought education under government control and made English the medium of instruction in schools. This tradition has continued years after the country gained independence as the language remains the major language of instruction in Nigerian schools across the various levels of education in the country. The language is

[1] As captured in the Oxford Advanced Learner Dictionary (ninth edition).

recognised as the official language (see Bamgbose, 2000; Fadoro, 2012). It is also conceived as the language of upward mobility in the country as it is often believed 'to make headway in Nigeria, one must have an English head' (Ajayi, 2013).

3. Context and language

Several arguments have been put up by language scholars with respect to the role context plays in language use. However, the conclusion of such arguments, as Verhagen (1997) opines, is the fact that meanings have to be taken as *constraints* on interpretations and for processes of interpretation, features of the context of utterance, whether linguistic or non-linguistic, can also be made recourse to. In the opinion of Ervin-Tripp (1994), context permeates language, and contextual assumptions influence how human beings comprehend language. Thus, a good knowledge of the contexts of speech is vital to develop realistic theories of language (use) and language learning. This position is reinforced by Doyle's (2007) submission that the fact that context affects meaning in language is not contestable. According to van Dijk (2008), context refers to some phenomena related to text, discourse and language usage. van Dijk differentiates between linguistic context, which is the verbal context of an utterance, and social or cultural context. According to him, the linguistic context has to do with the internal relationship between the elements of an utterance, while socio-cultural context of speech refers to linking talks or discourses to cultural values or beliefs of a people. Odebunmi (2006) describes situational context as location of a speech event, in terms of the physical arrangements, the objects in the environment, the participants and the topic of discourse, etc. van Dijk's conception of socio-cultural context, as espoused by Adegbija's (1982) pragma-sociolinguistic phenomenon forms the guide for analysis and discussion in this study.

4. Analytical tool

Following Adeyanju's (2011: 26) submission that 'the study of meaning in an L_2 setting of English usage demands a comprehensive theory such as the pragma-sociolinguistic theory because it addresses all the pragmatic and sociolinguistic factors that affect the total meaning of an utterance', Adegbija's (1982, 1988) pragma-sociolinguistic theory is adopted for analysis in this study. According to the theory, decoding the meaning of words and utterances in a discourse requires the understanding of the historical, personal, environmental, socio-cultural and linguistic aspects relating to the context in which the discourse took place (Adegbija, 1982). Accordingly, the meaning of an utterance cannot be understood without recourse to factors such as:

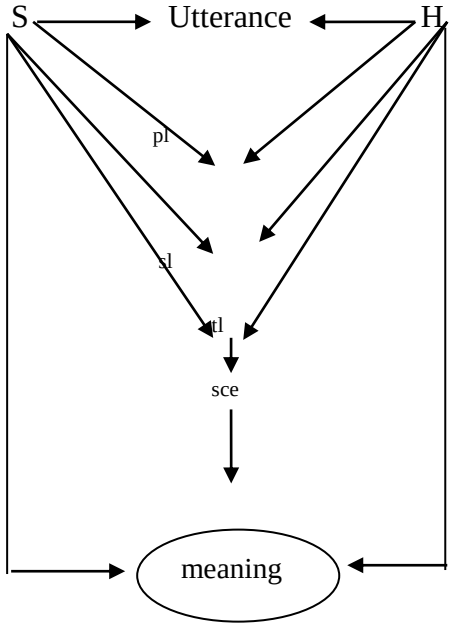
a. the history of the word or utterances;
b. the environment where the utterance is made;
c. the interlocutors (involved) and the relationship between them;
d. the socio-cultural values of the environment; and
e. the linguistic elements employed in performing the illocutionary act.

Adegbija, in his pragma-sociolinguistic theory, argues there are three layers of meaning by which utterances can be understood. These are the primary layer, the secondary layer, and the

tertiary layer (Adegbija's, 1988). At the primary level, we speak of the literal or ordinary level of meaning. It involves explaining the linguistic elements and the prosodic elements such as intonation, pitch, stress and rhythm of an utterance. The secondary layer, which handles the connotative or symbolic level of meaning (Adeyanju, 2011), deals with indirect speech acts- utterances in which one says one thing and means another; or says one thing and means what one says and also means another illocution with a different propositional content (Adegbija, 1982:32). The tertiary level of meaning involves knowledge of the different aspects of the socio-cultural environment in which an utterance is made. According to Adegbija, such aspects of socio-cultural environment most times give further information on meaning. Expatiating Adegbija's tertiary layer of meaning, Emuchay (2001: 196) submits:

> One's ethnic, religious and political background is an important part of one's identity. Thus, if the interlocutors do not share the same world view, the presuppositions they bring to bear on an utterance may differ and misunderstanding or lack of understanding may result.

The schema below, as designed by the researcher and adopted in this study, summarises the operation of the tenets of Adegbija' pragma-sociolinguistic theory.



Key:
S: speaker (s)
H: hearer (s)
pl: primary layer
sl: secondary layer
tl: tertiary layer
sce: shared socio contextual environment

Figure 1: Ajayi's Model of Pragma-sociolinguistic Theory

In the schema above, meaning, at the primary, secondary and tertiary layers, is generated by both the speaker(s) and hearer(s) who share the knowledge of the contextual environment in

an interaction. For instance, decoding the meaning of words and utterances at the literal level requires participants are familiar with the referential and linguistic elements of words and utterances. At the secondary level, for meaning to be decoded, participants are equally expected to bring their common experiences to decode indirect utterances. Similarly, at the tertiary level, knowledge of the socio-cultural environment shared by the participants plays a vital role. Thus, in this schema, as applied in this study, shared context environment (SCE) comprises all forms of knowledge- linguistic, psychological, sociological, or physical- shared by participants in interactive discourses.

5. Methodology

Given the nature of this study, an ethnographic research design was adopted. Data were gathered through participant and non-participant observation methods. Participant observation was employed in interactions in which I was personally involved, while non participant observation was used in interactions which I did not partake in but observe from a relative distance. These methods made it possible to gather data that reflected the true language practice of the participants in natural settings. Participants were observed in educational settings such as lecture rooms and hostels, social gatherings, including football viewing centres, shopping malls, parties and commercial vehicles in Southwestern Nigeria, comprising Lagos, Ondo, Oyo, Ogun, and Osun states between 2015 and 2018. A special field note was employed for documenting the observed interactions. However, for ethical considerations, the actual names of the participants where they were heard being mentioned were presented using pseudo names. Data were analysed within the purview of Adegbija's (1982, 1988) pragma-sociolinguistic theory.

5.1 *Contextual deconstruction of dress, meet, and toast*

This section of the study focuses on the contextual interpretations of the lexical items examined. My focus is on the verb forms of the words.

5.2 *Dress*

The word dress is a polysemous word whose semantic realisations are context-driven. In line with the submission of Adegbija, some of the meanings of the word can easily be decoded through the primary layer (literal) of meaning, while others can only be understood by making recourse to the secondary layer (connotative) of meaning. Some of the literal semantic senses of the word *dress* (which require the deployment of the primary layer of meaning realisation) include (1) to put on clothes, (2) to clean or treat a wound, (3) to clean and prepare meat, (4) to put an attractive arrangement in a shop window, (5) to stand in a straight line or to make soldiers stand in a straight line, among others. These senses of the use of the word dress, although require specific contexts, are basic to both the L_1 and L_2 contexts of the use of English, hence no reference to the secondary or tertiary layer of meaning interpretation is required. However, senses 6, 7, 8 and 9, as shall be presented in our analysis, are peculiar to the Nigerian socio-cultural context, hence recourse to the primary layer of meaning might not be helpful to L_1 speakers in decoding the meaning of the words in certain context-bound utterances.

Sense 6 (to adjust)
Context 1: This was an interaction between a commercial (taxi) car driver and his passengers.
Driver: *Ibo lẹ n lọ?*
Where are you going?
IP (Intending Passenger): Bodija
Driver: *Ẹ wọlé*
Come in
Driver to a sitting passenger (SP): *Ẹ jọ̀ọ̀, ẹ ba mi* ***dress*** *díẹ̀ fún wọn*
Please help adjust a bit for him/her
SP: Adjusts to create a space for IP

Context 2: This interaction took place between two undergraduate university students in a lecture room. One of the students was already seated waiting for the arrival of the lecturer. The other came in shortly before the arrival of the lecturer and this inteaction ensued.

Student A: (Already seated)
Student B: (notices an empty seat beside Speaker A) Can you please *dress* for me?
Student A: Okay (shifts)
Student B: Thank you.

In the two contexts presented above, it is observed that the word dress is worn a meaning entirely different from what is obtainable in the L_1 context of the use of English. In Context 1, the driver of the car, a semi-(il)literate Yoruba-English bilingual, pleads with the sitting passenger to adjust, that is move in a bit so as to create a sufficient space for the incoming passenger. In doing this, he employs the word *dress* instead of the verb *adjust*. This message is well understood by the passenger who quicly *dressed* in response to the plea of the driver. This mutual interpretation of the word dress by both participants as meaning 'adjustment' is a function of the shared socio-cultural background by them as Nigerians. If, for instance, the passenger were to be a native speaker of English who just finds him/herself in the country, there might have been a break down in communication, given his/her inability to interpret 'dress' as 'adjust' in that context. In Context 2, even though the interaction is between undergraduate students of a university who are assumed to have had exposure to the rules guiding the use of English in an L_1 context, the same semantic nuance is attached to the word dress by both participants. Given the socio-cultural environment speaker A shares with B, she naturally uses the word dress to mean adjust in this interaction, knowing it will be easily understood by him. This shared socio-cultural knowledge is shown by B who responds to A's request by adjusting himself on the seat to create a space for her.

Sense 7 (slap or hit)
Context 3: This interaction was a fight situation between two young men. It took place in Nigerian Pidgin (also known
A: *Na me you dey talk to like that*!
It is me you are talking to like that!
Am I the one you are addressing/talking to like that?
B: *Wetin you wan do?*
What do you want to do?

What can you do/what do you want to do?
A: *Màa dress etí ẹ fún ẹ nísin yìí*
I will dress your face for you now
I will slap you now
B: *Wà gbá ojú mi! O ò tó bẹ́ ẹ*
You will slap me! You dare not

The interpretation of *dress* in this interaction by both participants is predicated on their shared socio-cultural understanding of the word as such. It would be observed that in the statement of threat issued by Speaker A, no reference is made to the word 'slap'. He just simply says '*màá dress etí...*' "I will dress your ear". Speaker B gets the message and responds accordingly. A critical appraisal of this interaction, just like the first one, shows reference to the ordinary or literal sense of the word dress will not bring out the socio-cultural or contextual interpretation of the word. In other words, the word has assumed new semantic realisations defined by the realities that define the L_2 context in which it is used. Thus, in agreement with the tenets of Adegbija's pragma-sociolinguistic theory, some words need reference to the tertiary layer of meaning in their semantico-contextual interpretation. Such is the case of *dress* in the excerpts above where the word is used connotatively to mean 'slap'.

Sense 8 (to position) This is usually found among football lovers and fans to describe the handling of a ball at a point in time.
Context 1: The case of a footballer who is preparing to take a penalty kick in a football match.
Speaker A: *Mo jẹ́ẹ̀ri, o maa farabalẹ* ***dress*** *ball yẹn ni*
I trust him; he will take his time to position the ball
Other viewers: (Laugh) *wọ́n ti fi ṣe e*
He is being jinxed to do that
Speaker A: *À bẹ́ẹ̀ri ni*
Can you imagine!

Context 2: A footballer is seriously criticised for wasting time with the ball when he should have kicked the ball to the back of the net.

A: *Kí ló máa ń ṣe bọ̀bọ́ yìí náá?*
What is wrong with this guy?
B: *Ó ṣì ń dress ball, instead kó tètè gbá shot*
He keeps dressing the ball instead of him to quickly kick the ball
C: (Hisses in annoyance) *Bó ṣe máa ń ṣe niyẹ̀n. Idiot!*
He is fond of doing that. Idiot

In contexts 8 and 9, the word *dress* assumes another meaning other than its literal sense. In Context 8, the act of the player trying to properly position the ball in the right spot is described by Speaker A as 'dressing the ball'. The message is clearly understood by Speaker B in particular and other members of the audience in general to mean 'positioning' the ball. This mutual understanding of the word 'dress' as 'position' in this context by the participants is clearly a function of their shared socio-cultural background. The situation would be different in an L_1 context where such contextual interpretation of the word dress might be

found queer. Similarly in Context 9, the word dress is clearly interpreted as 'position' among the interlocutors. Speaker B, drawing on the socio-cultural understanding of the word to mean postion in a context like this in Nigeria, bemoans the delay on the part of the said footballer in kicking the ball. The response of Speaker C, which is apparently relevant to the comment of Speaker B, shows the trio share the mutual understanding of the word 'dress' to mean 'position' in this interaction. Although the word *dress* in the L_1 carries a meaning that is similar to the one expressed in the contexts presented above (to mean to position), it is particularly interesting to report that the sense in which it is used in the Nigerian L_2 is peculiarly different. It is never used in a football context to refer to positioning football in the L_1 context. This practice is in tandem with the claim of Adegbija and Bello (2001) that the transplantation of English to new environments confronts it with new contexts which play a major role in semantic nuances.

5.3 *Meet*

The word *meet*, very much like 'dress' examined earlier in this paper, has many senses which are both literal and socio-cultural. Some of the literal senses of the word which require recourse to the primary layer of meaning interpretation in both L_1 and L_2 contexts include (1) to see somebody at an arranged or appointed place, (2) see someone by chance, (2) see somebody for the first time, (3) come together to discuss something, (4) compete against somebody in a competition, (5) experience a problem or difficult situation, (6) to join or touch. However, in the Nigerian context of English language usage, the word 'meet' can as well carry a new meaning borne out of the socio-cultural environment of the Nigerian sociolinguistic space. In that wise, reference is made to the tertiary layer of meaning interpretation as captured in Adegbija's pragma-sociolinguistic theory. This is illustrated with the examples below:

Sense 7 (to have sexual intercourse)
Context 1: This is a case of a young man lamenting his being jilted by his girl lover to a friend.
A: I can't believe Juliet has finally left me... in spite of the love she claimed to have for me...
B: You don't know women. They can be very funny. Anyways, how many times have you *met* her?
A: Several times...
B: Ọmọ, why you dey worry yourself? You never lose now
A: Guy, you cannot understand; I love this babe...

In this interaction, the interpretation of the word 'meet' cannot be done within the primary level of meaning (the literal level). Even though the meaning of the word in the excerpt is connotative, its interpretation cannot still be handled at the secondary layer of meaning decoding. This is because the use and understanding of the word in the context of use is symbolically connotative whose semantic realisation can only be appreciated if reference is made to the socio-cultural values of Nigerians, particularly as it relates to the social norms which place certain constraints on language use. The word, as used and interpreted in this context, means to have sexual intercourse (with someone). It is obvious that, although interlocutors in this interaction are communicating largely in English, reference and reverence is shown to the socio-cultural norms of language use in the African system in

general, and the Nigerian socio-cultural environment in particular. Speakers in this interaction demonstrate their knowledge of the stance of the Nigerian sociolinguistic environment on taboo words and expressions. As defined by Oyetade (1994), taboo expressions are such expressions whose use, especially in the public space or discourse causes feelings of maladjustment among participants in a communicative activity. They are thus often avoided in embrace of their euphemistic alternatives (see Ajayi, 2017). As observed by Adeyanju (2011), sex is a taboo expression in the Nigerian sociolinguistic space. Hence, public discourses that require reference to sex are often loaded with expressions that make indirect and metaphoric reference to it.

Being aware of this socio-cultural reality, Speaker B in the excerpt above employs the use of 'meet' as a euphemistic alternative reference to sexual intercourse in order not to violate the taboo associated with direct reference to sexual intercourse and activities. This phenomenon is well understood by Speaker A who does not have to struggle to interpret the point being made by Speaker B with the use of the word 'meet'. Since neither of the primary and secondary levels of meaning can handle the interpretation of meet as used in this context, recourse to the tertiary level in decoding its meaning is imperative.

Context 2: A discussion between two brothers on a relative who keeps a diary where he keeps the record of his sexual activities.
Speaker A: *Ègbọ́n, ẹ ẹ̀ ri pe Bọ̀da Kúrúki yìí ò serious*
Brother, can't you see this Brother Kuruki is not serious
Speaker B: *Ẹni tí ò gbádùn*
One that is not okay
Speaker A: *Can you imagine, ó ń keep record iye ìgbà tó ti meet ìyàwo ẹ̀*
Can you imagine he is keeping the record of the number of times he sleeps with his wife
Speakers Aand B (Both laugh)

The situation in Context 2 is not different from what is observed in Context 1. The two siblings are commenting on an 'absurd' practice of the man at the centre of their discussion. Even though the interaction is between two brothers in a private and intimate discussion in which the socio-cultural norm of language use in the country could be relegated to the background, it is observed that the phenomenon of taboo is avoided. Speaker A, in order to avoid being obscene, resorts to the use of the word 'meet' to describe sexual intercourse in the interaction. Sharing the same socio-cultural background with Speaker A, Speaker B understands clearly the message embedded in the lexical item 'meet' in the interaction, and as such makes his contribution as relevant to the statement of Speaker A.

5.4 *Toast*

The word toast is another word that manifests different senses in the L_1 and Nigerian L_2 contexts of English language usage. Some of the senses include (1) to drink a glass of wine to thank or wish someone well, (2) to make bread or other food brown by heating it up, (3) to sit or stay near a fire to catch some warmth. However, beyond these senses that are applicable to both English L_1 and L_2 (Nigerian) contexts, the word has assumed a semantic nuance that is peculiar to the Nigerian sociolinguistic environment. This is evident in the excerpts below:

Sense 4 (to woo or ask (someone) out)
Context 1: An interaction between undergraduate students of a university in a salon
Speaker A: Eh en, I even forgot to tell you
Speaker B: What
Speaker A: You won't believe Kaka invited me for a chat last week Friday
Speaker B: (showing interest to hear more) Okay...
Speaker A: Guess what; he *toasted* me!
Speaker B: Yeh! this guy is in love
Speaker A: He said he loves me and would want to spend the rest of his life with me
Speaker B: That is serious; what did you now tell him?
Speaker A: I said I would think about it...

A critical appraisal of the deployment of 'toast' in the excerpt above reveals its meaning does not fall within the various senses that have been listed out earlier whose semantic realisation can be handled by the primary and the secondary layers of meaning realisation. Imagine if Speaker B, for instance, is a native speaker of English who is yet to familiarise herself with the sociolinguistic environment of English in Nigeria, deconstructing 'toast' in this context as 'woo' or 'ask out' would be very difficult. She would have been wondering about the appropriateness or otherwise of the use of the word, considering there is no celebration or party going on. She would also wonder if there is any loaf of bread to be toasted, especially considering the fact that the interaction takes place in a salon where the facilities to toast bread might not be readily available[2]. But, since both speakers share a common sociolinguistic environment of English usage, it is not difficult for them, and Speaker B in particular to make out the meaning of the word as used by Speaker A. In decoding the meaning of 'toast' in this interaction, both speakers make reference to the tertiary level of meaning realisation. This is because the word in this context depicts another meaning of the word, which is neither literal nor connonotative in the Nigerian sociolinguistic environment. This sense of the word is also reinforced in the interaction below:

Context 2: An interaction between a young man and a female friend about another lady whom the young man is interested in.
Speaker A: *Bambo, o ò believe pé mi ò mọ̀ bí mo ṣe máa* ***toast*** *baby yẹn*
Bambo, you can't believe I don't know how I am going to toast that lady
Speaker B: *Kí ló ṣẹlẹ̀?*
What happened?
Speaker A: *Each time tí mo bá ti ri, àyà mi máa ń já*
Each time I see her, I am always scared
Speaker B: *Ṣèbí ọkùnrin ni ẹ́*
You are a man
Speaker A: (Laughs) *o ò serious*
You are not serious
Speaker B: *Wàá dẹ́ sọ fun pé o like ẹ̀, and you will like to date her*
You will tell her you like her, and that you will like to date her

[2] Although personal experience has shown some hairdressers in Nigeria do keep cooking utensils as toasting machine, stove among others in the salon.

Speaker A: Okay o, it is just that there is something so unusual about this lady.
Speaker B: Eh en!
Really*!*
Speaker A: *Yes, kò yé mi...*but I will summon courage to meet her soon.
Yes, I don't understand...
Speaker B: *Àbí kí n bá ẹ lọ?*
Or should I go with you
Speaker A: (Both laugh) Go and sit down...

Here again, just like what obtains in Context 1, the word 'toast' assumes a meaning sense out of the senses found applicable in an L_1 context of English language usage. Both speakers orientate towards their sociolinguistic environment in decoding the actual meaning of 'toast' in the context of use. In some sociolinguistic environment other than Nigeria's, the word 'woo' or 'ask out' would be found; hence, Speaker A's statement *Bambo, you can't believe I don't know how I am going to* ***toast*** *that lady* could have been rendered '*Bambo, you can't believe I don't know how I am going to* ***woo*** *or* ***ask out*** *that lady*'. Since the interpretation of the word 'toast' in this context cannot be done making reference to the literal sense of the word, nor can it be done with recourse to the secondary layer since its use here is not connotative, making recourse to the tertiary layer of meaning interpretation becomes imperative. This is because it is the Nigerian sociolinguistic environment that has bestowed the 'woo' or 'ask out' sense on the word toast. This is line with Igboanusi's (2010) submission that 'toast' in the Nigerian context is often used to mean to 'woo' a girl or ask a girl out.

6. General discussion and conclusion

This paper has attempted a pragma-sociolinguistic deconstruction of three English words in Nigerian sociolinguistic environment. The words are 'dress', 'meet', and 'toast', with particular focus on their grammatical status as verb. As shown in our analysis, the words, apart from their meaning senses in the English L_1 context, manifest other senses that define the socio-cultural and sociolinguistic environment of English language usage in Nigeria. For instance, the word dress, depending on the context and participants, could be interpreted as 'to adjust, to position, and to slap'; the word meet connotes 'to have sexual intercourse', and the word toast can as well be decoded as 'to woo' or 'to ask out'. These Nigerian senses of the interpretation of these words, among others, are pointers to the indigenisation of English in Nigeria. This is a development to which Adegbija (2004:20) reacts thus: 'domestication, in the context of English in Nigeria, connotes 'home-grown', 'made native', adapted and tamed to suit the Nigerian environment'. As Adeyanju (2011) observes, the domestication of English in Nigeria constitutes a difficulty for L_1 users in particular and non-Nigerian speakers of English when interacting with Nigerian speakers of the language, particularly when it comes to understanding of the meaning of certain utterances in some specific contexts. To get round this problem, therefore, making recourse to the different layers of decoding meaning as encapsulated in Adegbija's pragma-sociolinguistic theory which takes in to consideration all the linguistic, pragmatic and sociolinguistic factors might be very essential.

References

Ajayi, Temitope M. 2013. Language, education and social class in Nigeria: the bridge hypothesis. *Ife Journal of the Humanities and Social Studies*, 1(1). 1-14.

Ajayi, Temiope M. 2017. Unveiling the taboo of sexual intercourse in the Beulah Yoruba/English Bilingual Parallel Bible. *Africology: The Journal of Pan African Studies* 10 (1). 309-324

Adegbija, Efurosibina. 1982. *A speech act analysis of consumer advertisements*. Indiana University. Doctoral dissertation.

Adegbija, Efurosibina. 1988. My friend, where Is Anini: Decoding the Meaning of Utterances. *Journal of Pragmatics* 12. 151-60.

Adegbija, Efurosibina & Bello, Janet. 2001. The semantics of 'Okay' (O.K.) in Nigerian English. *World Englishes* 19(1). 89-98

Adegbija, Efurosibina. 2004. The domestication of English in Nigeria. In Awonusi, Segun and Babalola, Emmanuel (eds) *The domestication of English in Nigeria: a festschrift in honour of Abiodun Adetugbo.* University of Lagos Press. 20-39

Adeyanju, Dele. 2011. A Pragma-sociolinguistic dissection of the word 'SEE' in Nigerian English. *Research in African Languages and Linguistics* 11. 23-34.

Bamgbose, Ayo. 2000. *Language and exclusion: the consequences of language policies in Africa*. Hamburg-London, Litvraig.

Banjo, Ayo. 1982. A historical view of the English language in Nigeria. *Ibadan* 28. 63-68.

Doyle, von Timothy. 2007. *The role of context in meaning and understanding*. Dissertation zur Erlangung des Doktorgrades der Philosophie (Dr. phil.) vorgelegt der Philosophischen Fakultaet der Universitaet Potsdam. (http://nbn-resolving.de/urn:nbn:de:kobv:517-opus-20691). (Accessed 2018-7-21).

Emuchay, Joyce. 2002. Cross-cultural pragmatics: me and you. In Babatunde, Sola T. and Adeyanju, Dele S. (eds), *Language meaning and society: papers in honour of Efurosibina. E. Adegbija at 50*. Ilorin: Haytee Press Ltd. 193-208.

Ervin-Tripp, Susan M. 1994. Context in language. In Dan I. Slobin, Julie Gerhardt, Amy Kyratzis & Jiansheng Guo (eds) *Social interactions, social context and language*. Hillside, NJ: Lawrence Eribaum Associates. 1-17.

Fadoro, Jacob O. 2012. Nigerian language and policy: theory or practice? In Ndimele, O.-M, Yuka, L.C and Ilori, J.F (eds.), *Issues in contemporary African Linguistics: A Festschrift for Oladele Awobuluyi*. Port Harcout. M & J Grand Orbit Communications Ltd. 29-39.

Igboanusi, Herbert. 2010. *A dictionary of Nigerian English usage*. Lit Verlag. Hamburg.

Odebunmi, Akin. 2006. *Meaning in English. An introduction*. Ogbomoso, Nigeria: Critical Sphere.

Oyetade, Solomon O. 1994. Taboo expressions in Yoruba. *Africa and Ubersee*. Band 77. 90-103.

Oxford Advanced Dictionary (Ninth edition). UK. Oxford University Press

Taiwo, Rotimi. 2009. The functions of English in Nigeria from the earliest timesto the present day. *English Today 98*, 25(2). 1-10.

van Dijk, Teun. 2008. *Discourse and context* : UK. Cambridge University Press.

Verhagen, Arie. 1997. Context, meaning and interpretation in a practical approach to Linguistics. In Leo Lentz & Henk L.W. Pander Maat (eds.), *Discourse analysis and evaluation*. Amsterdam: Rodopi. 7-39.

Temitope Michael Ajayi
Department of Linguistics and African Languages
University of Ibadan
Ibadan
Nigeria
michealtemitope@yahoo.com

In SKASE Journal of Theoretical Linguistics [online]. 2018, vol. 15, no. 3 [cit. 2018-13-12]. Available on web page http://www.skase.sk/Volumes/JTL38/pdf_doc/08.pdf. ISSN 1336-782X.

Nomination of mental illnesses in languages of different types

Alena Rudenka, Fang Xiang, Belorussian State University, Belorussia

A study of names for mental disorders in English, Chinese and Russian is the best way to reveal the national specifics of this fragment of the linguistic worldview. The naming methods of these concepts indicate the nature of this specificity. It has two causes. Firstly, these languages are typologically different: Russian is synthetic, English is analytical and Chinese is isolating one. Secondly, this specificity is explained by the mentality differences of Russian, English and Chinese speakers. However, except specifics, there are a number of common features in the nomination of mental diseases in languages of different types. These common features are in the use of native language resources for common names and borrowings for special ones; in the same metonymic transfers, in the closeness of semantic spheres of "mental illness" and 'stupidity'.

Keywords: *mental illness, nomination, semantic transfer, borrowing*

1. Introduction

The purpose of this article is to compare the nomination of mental diseases in English, Russian and Chinese. The names of mental disorders show us the diversity of naming mechanisms in different languages and national specifics of the worldview fragment which is connected with such names. For this study, the words were selected not from special dictionaries, but only from current ones, because the general tendencies of nomination and mentality specifics are connected with common nonterminological vocabulary.

2. Names of mental illnesses in current English, Russian and Chinese

The specificity and importance of the concept 'soul' in the Russian linguistic worldview is well known; it has been repeatedly and variously described (here we refer only to the conference in Warsaw, dedicated to the concept 'soul', and the book published after this conference [Kapełuś, Masłowska, Pazio-Wlazłowska (eds.) 2016]. It should be noted that only in Russian, mental illnesses received a generalized name *душевная болезнь* (literally 'an illness of soul'). Everyday common nominations, like *умственная отсталость* 'mental handicap', *слабоумие* 'dementia', etc. show that the Russian word *душа* is synonymous with the words *разум* and *интеллект* 'intelligence, mind', when we are talking about diseases. In other contexts the soul and heart are equated. The name *психическая болезнь* (literally 'an illness of psyche') is also used only in the Russian language: it is created by analogy with *душевная болезнь* on the basis of the derivative from borrowing (< Greek *psyche* 'soul').

In many other languages, the common names of mental and nervous diseases are based on the absence, lack, defect of the mind – a concept of "mind, intelligence" is their nominating basis. In particular, this is so in English: *mental affection, disease, illness, disorder, alienation, handicap, disability, deficiency*; *unsoundness of mind, illness of mind, feeble-mindedness* < *feeble-minded, weakness of mind, intellectual impairment,* etc.

Both in Russian and in English, there is the possibility of generalization: a particular concept can be denoted by the general one. In Russian, *умственно отсталый человек* 'mentally defective man' can be called simply *больной* 'a sick man'; Eng. *insane* (> *insanity*)

goes back to Lat. *insanus* ‘crazy, abnormal’ < *in* ‘not’ + *sanus* ‘healthy’ (> Eng. *sane* ‘healthy’) [Online Etymology Dictionary].

Semantics and compatibility of Eng. *soul* does not differ fundamentally from Rus. *душа*, except for the specificity of the Russian lexeme, which has been widely discussed.

Russian *душа* and English *soul* correspond to two words in Chinese: 心灵 [xinling] < 心 [xin] ‘heart’ + 灵 [ling] ‘spirit’; 精神 [jingshen], which was borrowed from Japanese 精神 [せいしん].

The history of the word with Japanese origin is as follows. In accordance with the Taoist tradition, three treasures are emphasized, and must be "protected" and "nourished" throughout human life. The first one is Jing (精), which can be translated as ‘semen’ or ‘essence’. It is a special substance (a kind of energy) that carries information about the structure and certain characteristics of the physical human body. On the basis of this concept of Taoism, Japanese created the word 精神 [せいしん] ‘soul’, which came into Chinese later.

The second treasure is Qi (氣) – vital energy (vital force) that circulates in the human body, and allows all organs and systems of the body to function.

Shen (神) is the third treasure. This term can be translated as ‘Spirit’. It is Shen that makes us reasonable, thinking, capable for self-improvement, and this is the third Chinese lexeme, which correlates with Rus. *душа* and Eng. *soul*.

From the above mentioned borrowing from Japanese we can trace the Chinese word 精神病 [jingshenbing] ‘mental illness’ < 精神 [せいしん] ‘soul’ + Chinese suffix 病 [bing] ‘sickness’[1].

However, the following nominations are more frequent. The concept of “mental retardation” is denoted in Chinese by the lexeme 智力障碍 ‘mental retardation’ [zhilizhang'ai] (simplified 智障 [zhizhang]) <智力 [zhili] ‘intellect’ + 障碍 [zhang'ai] ‘hindrance’. Chinese 痴呆症 [chidaizheng] ‘dementia’ comes back to 痴呆 [chidai] ‘stupidity’ + suffix 症 [zheng] ‘disease’. Its synonym is 失智症 ‘dementia’ [shizhizheng] < 失 [shi] ‘loss’ + 智 [zhi] ‘intelligent’ + suffix 症 [zheng] ‘disease’.

In all three analyzed languages, the names created by the native means of each language are only used for the general notion of “mental illness, mental retardation”. The designation of particular, specific mental diseases, as usual, are borrowed words. This general tendency is due to the fact that the nature and etiology of mental illness were usually ignored by the majority of the language speakers – their goal was to denote the fact of illness in general.

Here are some examples of the names of specific mental or neurological diseases in English and Russian, which come back, as a rule, to Latin or Greek roots – they are often internationalisms. As a rule, such names are medical terms, for example, Eng. *oligophreny* ‘congenital mental retardation, dementia’, different degrees of which are denoted by words *idiocy*, *imbecibility*, *debility* – its Russian equivalents are *олигофрения*, *идиотизм*, *имбецилизм*, *дебилизм*. This list is very long: Eng. *dementia* (< Lat. *dement-*, *demens* ‘crazy’ < *de-* + *ment-*, *mens* ‘mind’) - Rus. *деменция* ‘degradation of mental functions, resulting from brain damage’; Eng. *hypomania* – Rus. *гипомания* ‘an easy degree of mania’; Eng. *bipolar disorder* – Rus. *биполярное расстройство* (calque of the Latin expression) ‘maniac-

[1] Chinese 病 [bing] ‘illness’ is treated in this article as a suffix, because in modern Chinese it can not be used alone. Independent use of 病 [bing] is possible as a shortened variant of 疾病 [jibing] ‘illness’ or as an archaism.

depressive psychosis'; Eng. *kleptomania* – Rus. *клептомания*, and many others. There are also a series of international disease names which origins can be traced back to proper nouns, for example, Eng. *Lou Gehrig's disease* – Rus. *болезнь Лу Герига* (Henry Lou Gehrig was a famous American baseball player who suffered amyotrophic lateral sclerosis) [Merriam-Webster's Collegiate Dictionary]. Eng. *Munchausen syndrome* – Rus. *синдром Мюнхгаузена* is named after the literary character of Baron von Munchausen, who liked to draw long bow. Eng. *Parkinson's disease* – Rus. *болезнь Паркинсона* (> Eng. *parkinsonism* – рус. *паркинсонизм*) is named after English physician James Parkinson. The French neurologist Jean Charcot suggested this name in honor of the British doctor and author of "An essay on the shaking palsy", whose work was not properly appreciated during his life.

In Chinese, the general trend of denoting specific concepts by borrowing remains, but it is realized differently – often through loanwords from Japanese. For example, Chinese 白痴 [baichi] 'idiocy' was borrowed from Japanese.

There are many other Japanisms among Chinese medical terms. Chinese (simplified) 洁癖 [jiepi], traditional 潔癖 'misophobia' = 'a pathological fear of contamination and germs, avoidance of contact with surrounding objects' < Japan. 潔癖 症 [けっぺきしょう]. The hieroglyph 洁 [jie] has the meaning 'cleanliness', and 癖 [pi] means 'eccentricity, strangeness'. The productive Chinese suffix 癖 [pi] with the semantics of 'eccentricity, strangeness' is often used for the terminology of mental illnesses; it often appears as a part of terms like 窃盗癖 'kleptomania', 藏书癖 'compulsive hoarding', or 'hoarding disorder', 异食癖 'picacism' = 'the desire to eat something inedible'.

Chinese (simplified) 神经病 [shenjingbing] 'psychosis', traditional Chinese 神經病 'psychosis' consists of word 神经 [shenjing] 'nerve' + 病 [bing] 'sickness'. The meaning of 'psychosis' was borrowed from Japanese 神 経 症 [しんけいしょう] 'psychosis'. The very word from the traditional Chinese 神經 can be traced back to the Chinese 神 'the god, the spirit' + 經 'a book, a text', and in archaic Chinese it meant 'the mysterious book or text', i.e. it was not related to medicine. The medical meaning of 神經 'nerve' was created by Japanese doctor Sugita Genpaku (Japan. 杉 田 玄 白 [すぎたげんんぱく]) by means of a semantic translation of Dutch z*enuw* [**黃河清** 1996]. In the early twentieth century, the word with this new meaning came into Chinese medical terminology – at that time the dictionary of the Chinese language was actively replenished with Japanese lexemes. A number of other Chinese psychiatric terms are also connected with the word 神经 [shenjing] 'nerve'.

Chinese 神 经 官 能 症 [shenjingguannengzheng] 'neurosis' is a compound word consisting of 神经 [shenjing] 'nerve' (semantic borrowing from Japanese) + 官能 [guanneng] = 功能 'function' + Chinese suffix 症 [zheng] 'syndrome / illness'. The term 神经官能症 'neurosis' is most often used in Taiwan, Hong Kong and Macao, and on the mainland of China they prefer to use the word 神经症 [shenjingzheng] 'neurosis', which has the same meaning.

Simplified Chinese 神经过敏 [shenjingguomin] 'hyperesthesia' = 'high sensitivity', traditional Chinese 神經過敏 < 神經 [shenjing] 'nerve' + 过 [guo] = 过于 [guoyu] 'excessively, overly, overmuch' + 敏 [min] = 敏感 [mingan] 'sensitive' is borrowed from Japanese. In modern Chinese 神経過敏 is polysemantic: the first meaning of it is the medical term 'hyperesthesia', and the second one is 'skeptical (person)', 'paranoid (person)'.

Chinese 神经衰弱 [shenjingshuairuo] ‘neurasthenia’, consisting of the word 神经 [shenjing] ‘nerve’ + 衰弱 [shuairuo] ‘weak’, also was borrowed from Japanese 神経衰弱 [しんけいすいじゃく].

The name of the concept “autism” was borrowed from Japanese, too. Simplified Chinese 自闭症 [zibizheng] ‘autism’, traditional Chinese 自閉症 ‘autism’, which includes characters 自 [zi] ‘self’ + 闭 [bi] ‘close’ + 症 [zheng] ‘disease’ go back to Japanese 自閉症 [じへいしょう] ‘autism’.

Chinese 恐高症 [konggaozheng] ‘acrophobia’ = ‘irrational fear or phobia of heights’ is not related to Japanese. The word goes back to the Chinese character 恐 [kong] ‘fear’ + 高 [gao] ‘height’ + 症 [zheng] ‘syndrome / disease’, it is the semantic calque of the internationalism *acrophobia* (< Greek ἄκρον ‘peak, summit, edge’ and φόβος ‘fear’ + Chinese suffix 症 [zheng] ‘syndrome / disease’).

There are also some special names of mental illnesses in Chinese that can be traced back to the traditional Chinese medical terms, for example, 癔病 [yibing] ‘hysteria’ = 癔 症 < 癔 [yi] ‘hysteria’ (the term from traditional Chinese medicine) + suffix 病 [bing] ‘sickness’. The other term from traditional Chinese medicine is 谵妄 [zhanwang] ‘delirium’ < 谵 [zhan] ‘to rave’ + 妄 [wang] ‘ludicrous, absurd’, i.e. the disease is named metonymically according to the characteristic symptom – the patient’s incoherent speech.

Other metonymic transfers are also commonly used: mental diseases are nominated not only as problems with the mind and intelligence (Russian ***сума****сшедший* ‘crazy’), but also as problems with organs ‘containing’ the intelligence: Russian expressions *на голову больной* ‘with a sick head’, *на голову хромает* (verbatim ‘his head is limp’), *с головой не дружит* (verbatim ‘not friendly with one’s head’); English *go out of one’s head, crack one’s brains*. Obviously, this is due to the metonymic nomination of the mind through the name of the organ that is “responsible” for it.

Often a mentally ill person or a person with inadequate behavior or a fool is designated by the same lexeme: Rus. *сумасшедший* ‘crazy’, colloquialism *псих* ‘psychopath’, Eng. *crazy* ‘mentally ill person’ → ‘person with inadequate behavior’. Examples of the same semantic development also exist in Chinese: 神经不正常 [shenjingbuzhengchang] ‘nervous, troubled, abnormal (person)’ can be traced back to 神经 [shenjing] ‘nerve’ + 不正常 [buzhengchang] ‘abnormal’. The Chinese colloquialism 疯子 [fengzi] ‘madman, loco, loony’ contains both meanings: ‘mentally ill person’ and ‘person with inadequate behavior’ (< 疯 [feng] ‘crazy’, ‘inadequate behavior’ + 子 [zi] Chinese suffix without semantics).

Semantic transfer ‘mentally ill person’ → ‘a fool’ is represented in Russian names of patients suffering from specific types of mental illnesses: *идиот* ‘idiot’, *дебил* ‘moron’; Eng. *idiot, imbecile*; Chinese 傻子 [shazi] ‘imbecile’, 傻瓜 [shagua] ‘moron’. In Chinese, this semantics may also be transferred by lexemes like 白痴 [baichi] ‘idiocy (sickness)’, ‘stupid’ (Japanism); 弱智 [ruozhi] ‘mental handicap (illness / disability)’, ‘stupid’ (< 弱 [ruo] ‘weak’ + 智 [zhi] ‘intelligence, mind’); 智力障碍 [zhilizhang'ai] (often abbreviated as 智障 [zhizhang]) ‘mental handicap’, ‘retard, moron’ < 智力 [zhili] ‘intelligence’ + 障碍 [zhang'ai] ‘hindrance’. The shortened word 智障 [zhizhang] ‘mental handicap, intellectual incapacity’ (→ ‘a fool’) came into Chinese through semantic translation from English *intellectual disability*.

3. Conclusion

There are several common trends in the nomination of mental illnesses in languages of different types.

Firstly, the native nominative means of a particular language are used for the nomination of mental illness in nonterminological names, but when naming specific mental diseases, usually, borrowing is preferred. This statement is not as true for English as for Russian and Chinese, because some of English general terms for mental disorders are loanwords from Latin (*insane*) or French (*disease*).

Secondly, in all analyzed languages there are such semantic metonymic transfers as: 1) the genus-species synecdoche – mental illness can be nominated through naming a concept of "disease"; 2) mental illness can also be nominated as a problem with the organ which is responsible for intelligence.

Thirdly, there are other frequent semantic transfers: ‘mentally sick person’ → ‘person with inadequate behavior’; ‘mentally ill person’ → ‘fool, stupid’.

References

Kapełuś, Magdalena, Masłowska, Ewa, Pazio-Wlazłowska, Dorota (ed.). 2016. Świat oczyma duszy. Antropologiczno-językowe wizerunki duszy w perspektywie międzykulturowej. Tom 2. Warszawa: Wydawnictwo Instytutu Slawistyki Polskiej Akademii Nauk.

Lees, Andrew J. 2007. Unresolved issues relating to the shaking palsy on the celebration of James Parkinson's 250th birthday. Movement Disorders. Vol. 22. Suppl. 17. P. 327–334.

Masłowska, Ewa, Pazio-Wlazłowska, Dorota (ed.). 2016. Dusza w oczach świata. Antropologiczno-językowe wizerunki duszy w perspektywie międzykulturowej. T. 1. Warszawa: Wydawnictwo Instytutu Slawistyki Polskiej Akademii Nauk.

Merriam-Webster's Collegiate Dictionary. 2008. 11th ed., Springfield, Merriam-Webster.

Online Etymology Dictionary. (https://www.etymonline.com. Accessed 2018-03-210).

李養正 [Li, Yangzheng]. 2011. 道家精氣神學說 [A theory of Taoist ‘Jing’, ‘Qi’ and ‘Shen’]. 國際道家學術總會 [The international association of Taoist Studies, Taipei]. http://www.etaoist.org/taoist/index.php/2011-08-22-02-11-46/2011-09-13-16-05-29/ 1865-2012-08-01-10-48-03. Accessed 2018-03-210).

黃河清 [Huang, Heqing]. 1996. 神經考源 [The origin of the term Nerve]. 学术集林 [Research collections], Yuanghua, Wang (ed.). Vol. VII. Shanghai: Shanghai Far Eastern Press. (http://www.huayuqiao.org/ articles/huangheqing/hhq04.htm. Accessed 2018-03-210).

Alena Rudenka
Per. Kazarmenny
5a-49, Minsk, 220030
Belarus
alena@rudenka.com

Xiang, Fang
Ul.Sverdlova
34-218, Minsk, 220004
Belarus
antonvonschaan@gmail.com

In SKASE Journal of Theoretical Linguistics [online]. 2018, vol. 15, no. 3 [cit. 2018-13-12]. Available on web page http://www.skase.sk/Volumes/JTL38/pdf_doc/09.pdf. ISSN 1336-782X.

A Phasal Account of Arabic Passive Constructions

Mohammed Odeh AL-Shorafat, Yarmouk University, Jordan

The main concern of this paper is to describe and analyze Arabic passive constructions within the most recent phase-based syntactic theory of Chomsky (I998, 1999, 2001, 2005, 2006).As far as I know Arabic passive constructions haven't been investigated within this recent framework. Describing and analyzing those constructions within Chomsky's recent phase-based framework will be of great significance to Arabic as well as to Chomsky's phase-based theory of syntax. Applying recent syntactic theories to the structure of Arabic will enable this language to assume its deserved position among the languages of the world and will attract the attention of more modern linguists to render more service to this distinct non-indo-European language by studying different aspects of its structure. As for Chomsky's phase-based theory, if the structures of Arabic can be handled within this theory, then this will lend further support to the universality of Chomsky's phasal theory. It is also important to determine if there is movement involved in the formation of Arabic passives as is the case in English or not. Besides, if there is movement involved where does the moved DP land? Further, is it an A-movement or an A-bar movement?

Keywords: *Passivization, Standard Arabic, Minimalist, Phases, A-movement, A-bar movement*

1. Introduction

Passive voice in the languages of the world, especially in English, has received a great deal of attention in the linguistic literature. There have been quite a number of studies that tackled this topic (Chomsky 1977, 1981; Jaeggli 1986 & Baker, Johnson & Roberts, 1989, among others).

As far as English is concerned I will mention here three major studies that have dealt with Passive constructions. The first study is that of Chomsky (1981:124) in which he considers passive not the result of a passive transformational rule as advocated in his earlier work (Chomsky 1977: 41) but the result of the interaction of two basic properties: the subject NP (DP) does not receive a θ-role in a passive structure and the object NP does not receive an Acc case. As for Chomsky (1981), passive is considered a case of NP movement within the universal Move α-module (cf.Ouhalla 1994:79). Thus, within the Government-Binding framework (Chomsky 1981), case assignment is withheld until the object moves to the empty position vacated by the subject where it is assigned Nom case.

In a similar fashion, Jaeggli (1986: 587) argues that passive involves the 'absorption' of the external role of the verb and it is prevented from assigning objective case. Jaeggli explains the two passive properties referred to above by assuming that the passive morpheme -en, since it is in the government domain of the verb, can receive objective case and θ-role. This forces the complement NP of the verb to move to subject position to receive Nom case.

Baker et al. (1989) have developed the idea proposed by Jaeggli further by considering the passive morpheme -en as an argument. As an argument, according to them, it should receive θ-role and case. However, the passive account proposed in the previous two studies does not fare well in the minimalist framework. Simply because the government notion, prevailing in the GB framework and under which syntactic operations including passive formation, is no longer functioning within minimalism. As a matter of fact, Chomsky (1995: 176) states that,

"the concept of government would be dispensable, with the principles of language restricted to something closer to conceptual necessity...." One more problem with the previous studies, especially Baker et, al's where it is assumed that the VP determines the θ-role of the external argument; in the minimalist framework the light verb v (or a complex of V+v) determines this and the external argument is base- generated in Spec-v.[1]

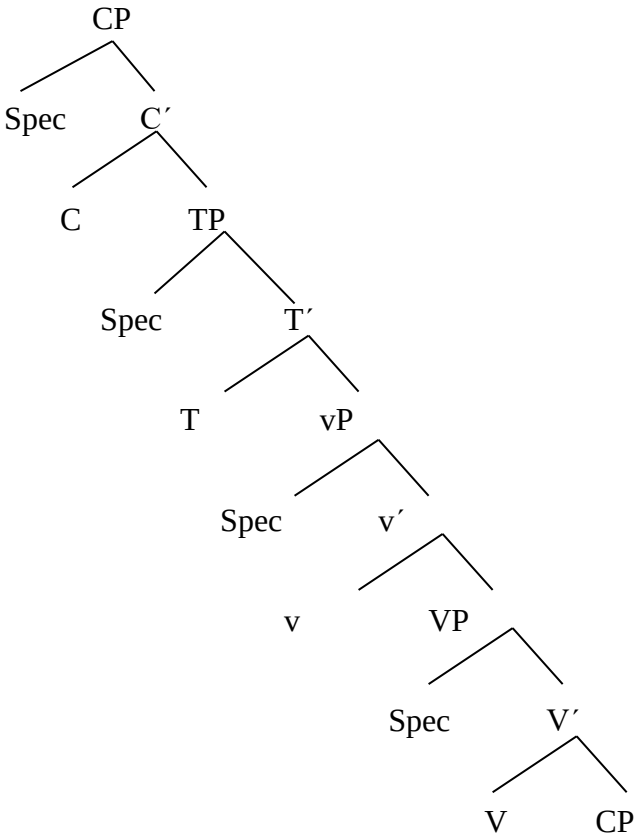
2. A Phase-based syntactic theory

Before delving into Arabic passive constructions, I think it is only appropriate to present a brief overview of Chomsky's most recent phase-based framework. As is well known, Chomsky adopts the split VP analysis where a VP is split into two projections: an outer vp shell and an inner VP core. Following Rizzi (1997), he also assumes that even root clauses are introduced by CPs in order to express the force of the clause whether it is declarative, interrogative, imperative or exclamative. The introduction of root clauses with a CP specifier and a C head has not been a common practice in previous linguistics studies. Chomsky (1999: 9) also considers CP and transitive vp (which he denotes as v*p) as phases. His reason for considering CP and vp as phases is that CP is a complete complex with a force marker and that vp represents a full thematic complex with an external argument. He further maintains that C and v are phase heads and that syntactic operations involve an agreement relation between a Probe and a local Goal (Chomsky1998, 1999, 2001).He posits that T and v are Probes which enter the derivation with unvalued φ-features (person, number and gender).When they enter into an Agree relation with a DP that carries a complete set of φ-features, their features are valued and deleted. Further, merger operations apply before any probing can take place. He also proposes that defective TPs[2] and vps that lack an external argument are not phases. This idea goes back to what has been known in the literature as Burzio's Generalization (1986:179,185) which roughly states that a verb which lacks an external argument cannot assign an Acc case and cannot theta mark an external argument.

Chomsky further assumes that probes can probe either simultaneously or sequentially with some choices converging while others crashing (Chomsky 2006:17). Further, when all syntactic operations in a given phase have been completed, the complement of the phase head becomes impenetrable to further syntactic operations which Chomsky (2001) terms the Phase Impenetrability Condition. His reason for why the complement or domain of the phase becomes impenetrable is that once the syntactic computation in a given phase has been completed, the complement or domain of this phase undergoes transfer simultaneously to the PF and LF components to be assigned the appropriate representation. Furthermore, the architecture of the clause structure in a phase framework can be represented by the following tree-diagram (modified from Radford, 2009:357):

[1] For more on this, see Tucker (2007).
[2] Defective TPs are clauses that lack a CP layer like ECM clauses.

(1)



It can also be represented by the following labeled bracketing:

(2) [CP[Spec[C´[C[TP[Spec[T´[T[vp[Spec[v´[v[VP[Spec[V´[VDP]]]]]]]]]]]]]]]]

Having laid out a brief sketch of the phasal framework of Chomsky's recent work, I now move to consider passive constructions in SA.

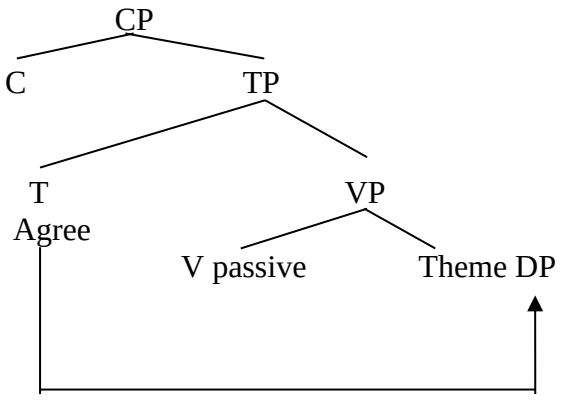
3. Passive constructions in Standard Arabic

As far as Standard Arabic (SA) is concerned, there is paucity of recent studies on passive formation.[3] However, it is important to mention here that a recent study (a doctoral dissertation) conducted by Soltan (2007) presented a short section for agreement in Arabic passives. Soltan (2007:96ff) assumes that there is no movement of the internal argument in a passive construction in SA.He also maintains that the internal argument appears with a
Nominative case and shows gender agreement as is the case with post verbal subjects. He further adds that since there is no movement involved in passive formation, there is no A-movement. For a passive sentence as in (3) below, he posits the clause structure in (4).

[3] I am indebted to a reviewer who pointed out to me two references: Hallman's (2002) comprehensive study on Passive in Arabic and English presents a concise table of the ten possible verbal forms of perfect and imperfect of active sentences and their passive counterparts. However, his paper mainly deals with the morphology and valency of passive and active, and was not carried out within the minimalist framework. Bubenik (1979) was conducted within the 'Functional Sentence Perspective' of the Prague School. It is not relevant to my research.

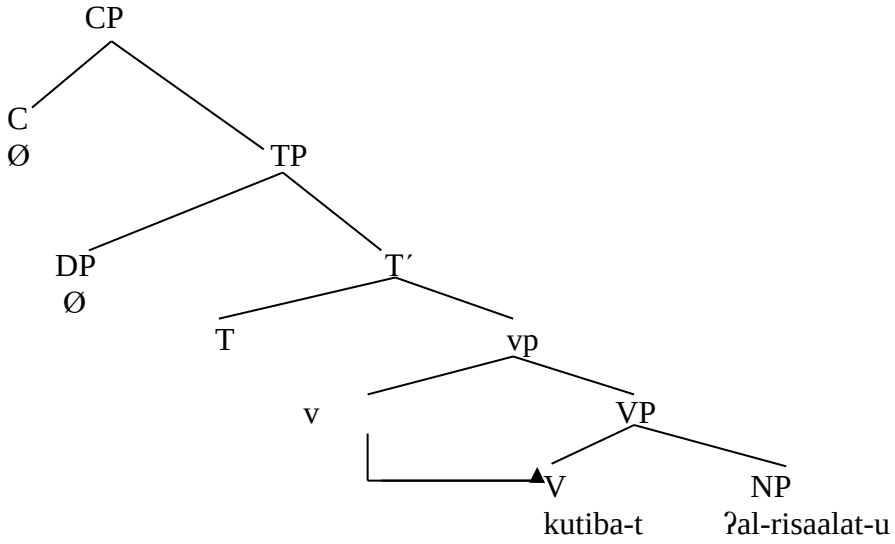
(3) *kutiba-t* *ʔal-risaalat-u*
written 3sg –fem the-letter-No
'The letter (was) written'

(4)



However, this clause structure is not compatible with a phase-based framework as represented in (1) above and partially repeated below in a tree diagram:

(5)



In the phase-based framework, the light v in the tree above has no external argument; it has no agreement features, nor case assigning ability. Hence, it is not a Probe and there is not a vP phase. But since it is affixal, it triggers movement of the verb and the affixal v will add the passive morpheme to the verb, appearing in the PF as the perfective form (*kutiba-t*, 'written').The question at this point is: how does the NP *ʔal-risaalat,* 'the letter' get its Nom case and does this NP move or stay in situ? And if it moves; is the movement A-movement or A´-movement? The answer to these questions will be the topic of the following sections. But before going into these questions let me present a brief description of passives in SA.

Arabic grammarians have dubbed Arabic passivization, 'al-mabni li-lmajhu:l' (literally built structure of the unknown) or 'apophonic' passive. As a matter of fact, a traditional Arabic grammarian, AL-Astrabadi (1996: 28; quoted in Maalej 1999: 4) calls passive, "a verb built for the object whose agent has not been named." In addition, passivization in Arabic is a productive process, i.e., any transitive verb can be passivized. Moreover, Arabic passivization is performed by a "melodic overwriting" (cf. McCarthy 1981: 399) in which the vocalic pattern of the transitive verb changes in one of two ways: If it is perfective like *kataba* 'wrote', it changes into *kutiba* 'written', i.e., a----a → u-----i, if it is imperfective like *yaktibu* 'write(s)', it changes into *yuktabu* 'written', i.e., a------i → u---a.[4] Check more examples in (6) below:

(6)

Perfective		Imperfective	
kasara	kusira	yaksiru	yuksaru
broke	(was) broken	break(s)	(is) broken
ʔarsala	ʔursila	yursilu	yursalu
sent	(was) sent	send(s)	(is) sent
ʔistaqbala	ʔustiqbila	yastiqbilu	(is) yustaqbalu
received	(was) received	receive(s)	(is) received
ʔakala-t	ʔukila-t	t-ʔakulu	t-ʔukalu
ate-fem	(was) eaten-fem	fem-eaten	(is) fem-eaten

Having shown how the vocalic pattern of perfective and imperfective verbs changes from active into passive. We now move to consider some active voice sentences and their passive voice counterparts

It has been standard practice among Arabic grammarians to consider the (7a, 8a, 9a) sentences below as active voice sentences, whereas their (b) counterparts as passive voice sentences:

(7) a. *kataba* *l-walad- u* *risaalat-an*
wrote the-boy-Nom letter-Acc
'The boy wrote (a) letter'

b. *kutiba-t* *risaalat –un*
written-fem letter-Nom
'(A) letter (was) written'

(8) a. *kasara* *l-walad-u* *l-zujaaj-a*
broke the-boy-Nom the-glass-Acc
'The boy broke the glass'

b. *kusira* *l-zujaaj-u*
(was) broken the-glass-Nom
'The glass (was) broken'

[4] Check Maalej (1999) for more on this.

(9) a. *yaktibu* *r-raajul-u* *qisʔat-an*
write(s) the-man-Nom story-Acc
'The man write (s) (a) story'
b. *t-uktabu* *qisʔat-un*
(is) written story-Nom
'(A) story (is) written'

If we examine the above examples, we realize certain properties that differentiate between the passive sentences and their active counterparts. The subject of the active sentence does not surface in the passive sentence. Instead the object which has the thematic role of patient surfaces in the subject position which has been called in Arabic 'a deputy subject'. Since the active subject does not surface and the patient object takes its place, the Arabic grammarians called this object a deputy subject (not the real agent). On minimalist assumptions, there has been 'de-linking' between passive and active constructions which means that passive and active have different derivations partly because the meaning is not recoverable across the two constructions (Chomsky1995). In other words, since the subject does not surface in Arabic, it is not known who is the agent is. It is a fact of passive syntax that the object has the thematic role of patient in both sentences which conforms to Baker's (1988) Uniform Theta Assignment Hypothesis (UTAH).The jest of Baker's hypothesis is that if two arguments fulfill the same thematic function with respect to a given predicate, they must occupy the same position in syntax.

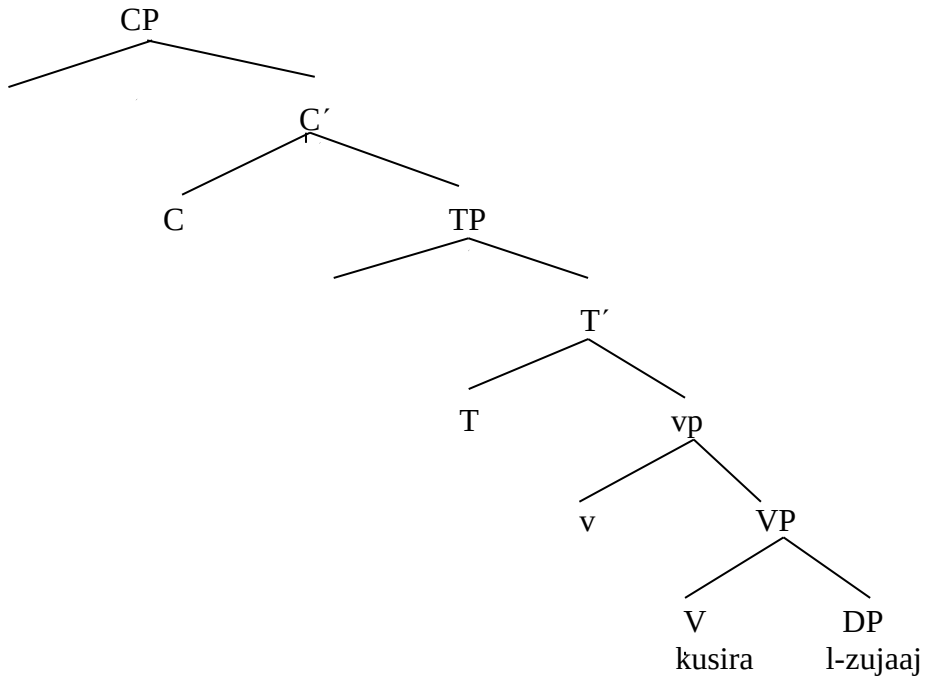
If we adopt Baker's UTAH, this means that the passive subject should occupy the same position as the active complement. Baker developed (UTAH) to show that if passive subjects have the same θ-role as active objects, it is possible to assume that passive subjects originate in V-complement position as objects. Syntacticians have assumed that Universal Grammar principles correlate thematic structure with syntactic structure in a uniform fashion. Thus, in sentence (7b) above the passive subject *risaalat-un* occupies the same complement position of the object in (7a) because both have the same thematic role patient. The question this paper tries to answer is this: If the passive (deputy) subject originates in a complement position, how does it end up in a subject position with a Nom case? Before attempting to answer this question, I have to lay down some ground work. I will start with a tree diagram[5] of the architecture of the passive clause structure in the minimalist phasal framework (10).

The main difference between an active sentence and a passive sentence is that the outer vp shell has no specifier position to host the agent because the agent is absent. Another difference is that the v is not an active Probe because it lacks a thematic agent argument and the only active Probe available in (10) above is the T. The T is an active Probe by the fact that it has unvalued/uninterpratable φ-features and a Nom case to assign. However, let us consider how the passive sentence, *kusira l-zujaaj* 'the glass (was) broken' is derived within Chomsky's phasal framework.

The verb *kusira* 'broken' enters the derivation either fully inflected with the passive vocalic pattern or as in Chomsky's (1999) recent work the verb gets its passive marking from the light affixal verb v after the verb *kusira* moves and joins to it. The complex v+V moves to T to satisfy its tense requirement. The complement *l-zujaaj* 'the glass' enters the derivation with a full set of φ- features but no case. However, within Chomsky's recent phase-based theory T is a composite of features. In VS order in SA, finite T has a set of unvalued φ-features (person, number and gender), a D-feature a'nd a Nom case. Usually the φ-features in T in VS order are "...default number and person features and... gender features" (Al-Horais 2009:18). The

relevant feature in T is the gender feature. The masculine agreement morpheme is null in SA (Al-Horais 2009:2). However, nouns with feminine gender enter the derivation with a morphological marker-t or they are lexically specified, especially with vocalic patterns such as *fa^{c}laaʔ* (*sahraaʔ* 'desert') or f*u^{c}laa* (*ḥublaa* 'pregnant').

(10)



As mentioned before, the light affixal v enters the derivation with no φ-features and no case assignment ability, (but with a passive morpheme which Chomsky names PRT, i, e, (participial, ibid) because it lacks an external argument. The T enters the derivation with unvalued φ-features, especially in this case the Gender feature and a Nom case feature. Now the derivation of the passive sentence proceeds as follows.

The verb *kusira* merges with the complement DP *l-zujaaj* to form a VP and not a V´ because as stated in Radford (2009:294), "it is a property of passive verbs that they project no external argument" The VP merges with the light v to form the vp and the vp merges with the T to form a T´ and the T´ merges with its Specifier to form a TP and the TP will merge with the null C marker to form a C´ and the C´ will merge with its Specifier to form a CP which is marked as having declarative force. Since the v is affixal in nature, it triggers the movement of the verb *kusira* to adjoin to it. It has been mentioned before that the V *kusira* joins to the light v, and gets its passive vocalic layer from the v as suggested by Chomsky. At this point it should be pointed out that the v has no φ-features to check and no Acc case to assign. The DP *l-zujaaj* is in a complement position. Though the complement has a thematic role of patient, it has an agreement feature (Gender) and no case. Even if we assume that the verb enters the derivation in its root form (*k s r*) and in this case it can project an external argument because, following Chomsky, it will receive its passive morpheme after it moves and joins to the light v. The DP complement cannot move to the Specifier of VP because it will violate a Universal principle which disallows phrase internal movement. Boeckx (2007:110) formulates this principle as a constraint:

[5] I have to point out here that this tree represents my own understanding of the passive clause structure; it might not be the standard representation of passive structures in minimalism.

(11) Antilocality constraint:

Movement internal to a projection counts as too local, and is banned.

Since vp is not a phase, the only active Probe available is T. However, T has unvalued agreement features, D-feature and a Nom case and the complement is in a position where it cannot receive a Nom case; it is a position of an Acc case. Thus the complement has to move to a subject position where it can receive a Nom case. The only position available where the complement DP can move into is Spec-T in (10) above. However, the DP cannot move to this position because if it does it will be in a position higher than the Probe T and an Agree relation cannot be established.

In the previous discussion I have argued that the null light v does not project a Specifier position if it lacks an agent or an experiencer argument. However, this does not mean that vp cannot project a Specifier position; it can. For example in English, Spec-v can be occupied by the expletive *there* (Radford 2009: 384) though v is not a phase head. I will assume at this stage of the derivation that the vp will project a Specifier position to cater for the complement DP to move into. It is the only landing site for the complement DP. Now, T, being active by virtue of having unvalued/uninterpretable agreement features and a Nom case, starts searching for a local active goal and locates the DP *l-zujaaj* in the Specifier of vp where an Agree relation is established and the features of T are valued and deleted, and the DP gets its Nom case. The complex V+v moves to the T node in order to satisfy its tense feature. The DP does not move to the Specifier of T and stays in situ, i.e., Spec-v. The question that might arise here is why the DP does not move to Spec-T.

There has been a vast body of linguistic literature on the position of subjects in VSO languages (Emonds1981; Sproat1985; McClosky1986; Ouhalla1994; Bobaljik & Carnie 1996, *inter alia*). However, what concerns me here is that traditionally the EPP feature which triggers the (deputy) subject's movement is considered to be lacking (or weak) in VSO languages which means the subject stays in Spec-v (Al-Horais 2009:16). Another reason for the non-movement is that in VSO orders the agreement between the T and the DP is not rich (Fassi Fehri 1993), i.e.; it is not strong enough to trigger movement. As a matter of fact in our passive example the only agreement on T is Gender; Person and Number can be considered as a default case in this example. Moreover, if the subject is forced to move to Spec-T we will get *l-zujaaj-u kusira* which is an SVO order or a Topic vp clause. Of course, this will take us into a completely complex different line of analysis which is not within the main focus of this paper. Furthermore, Tucker (2007) argues that the subject in Spec-v should move to Spec-T even in VSO order. He bases his position on what he wrongly assumes three pieces of evidence. First, he assumes that the 'comfortable' canonical position for manner adverbs in Arabic is to the right of the subject. He cites the following example:

(12)	*?? yı-d^{ʔ}rıb*	*l-wðlad-u*	*al-qaadi*	*^{c}adıdan*
	hit	the-boy	the-judge	often
	'The boy hit the judge often'			

He comments by saying that if the subject were in Spec-v, the adverb would appear in the position shown in (12) above (his own example), i.e., between the verb in T and the subject within the vp layer and adds but" this is not the case" and that is why he marks the sentence with double ?? to indicate that it is ungrammatical (Tucker, 2007: 27). In fact this is the

canonical position of adverbs in Arabic, i.e., they appear in clause final position. As a matter of fact, Fassi Fehri lists a number of Arabic examples where it can be seen that the 'comfortable' position of adverbs is clause final position. Consider the examples below:

(13) *kataba* *r-rajul-u* *r-risaalat-a* *ʔamsi*
wrote the-man-Nom the-letter-Acc yesterday
'The man wrote the letter yesterday'

(14) *yaktubu* *r-rajul-u* *s-saaaᶜata*
write the-man-Nom now
'The man write(s)/is writing now'

(15) *y-aktub-u* *r-rajul-u* *ɣadan*
write the-man-Nom tomorrow
'The man write(s)/is writing tomorrow' (Fassi Fehri 1993: 145)

It can be seen from the examples above that the canonical position for adverbs in SA is clause final position, contrary to Tucker's claim. One more thing *ᶜadıdan* is not a manner adverb; it is a time adverb meaning 'many times'.

His second piece of evidence is sentential negation in Arabic. He, following Ouhalla (1994), tries to establish that Neg morphemes must be situated between TP and vp and brings in examples from Arabic dialects, not from SA, but the situation is completely different with SA. I quote here Fassi Fehre (1993: 166) where he states that, "Arabic Neg morphemes select a clausal inflectional projection, not a bare VP..." This means that Negation in VSO order appears in clause initial position and not as claimed by Tucker. Concerning his third piece of evidence which he calls 'vp Ellipsis', he does not fare any better. The dialectal examples he brings in are completely ungrammatical. Hence his argument has to be rejected. One final point to mention here is that although Tucker conducts his research within Chomsky's recent minimalist syntax, he doesn't adopt Chomsky's phase-based approach. Hence; there is a theoretical difference between his work and mine.

Now, I go back to peruse the derivation of the sentence *kusira l-zujaaj-u*. After movement of the complex V+v to the T node, the whole TP (the domain or complement of CP) is transferred to the PF and LF components to receive the appropriate phonemic representation and semantic interpretation. At the end of the derivation, the remaining constituents, the null declarative C marker and its projection, the CP undergo transfer as well. The movement of the DP *l-zujaaj* to the Spec of vp is an A- movement because it is a movement from an argument position (the complement position) to another argument position (the Spec of vp). In this section I have tried to show how a passive clause in SA can be derived in a phase-based framework. I have also shown that movement is involved in the derivation of a passive clause in SA.

4. Conclusion

In this paper I have attempted to describe and analyze a sample of passive constructions in SA within Chomsky's phase-based framework of syntax. I have tried to show that the complement

DP of the passive verb has to move to Spec-vp in order to receive its Nom case. I have also indicated that the movement to Spec-vp is an A- movement.

This paper has also shown that Chomsky's minimalist phasal framework can be applied successfully to Arabic passive constructions which is very significant for both SA and the recent Minimalist Theory. I believe researchers interested in SA structures would greatly benefit from this study. Future research in this area would be expected to describe and analyze more complex Arabic passive constructions such as double object constructions in a minimalist phasal framework. Hopefully, these constructions and others would be the topic of my future research within the recent minimalist theory.

References

Al-Horais, Nasser.2009. A Minimalist Approach to Agreement in Arabic.*New Castle Working Papers in Linguistics* 15. 1-20.

Baker, Mark.1988. *Incorporation*. Chicago: University of Chicago Press.

Baker, Mark & Johnson, Kyle. & Roberts. Ian. 1989. Passive Arguments Raised. *Linguistic Inquiry* 20. 219-251.

Bobljik, Jonathan & Carnie, Andrew. A Minimalist Approach to Some Problems of Irish Word Order. In Borsley,Robert & Roberts, Ian (eds.), *The Syntax of the Celtic Languages*, 223-240. Cambridge: Cambridge University Press.

Boeckx, Cedric. 2007. *Understanding Minimalist Syntax.* Oxford: Blackwell.

Bubenik, Vit.1979. Thematization and Passivization in Arabic. *Lingua* 49. 395-313.

Burzio,Luigi1986. *Italian Syntax: A Government Binding Approach* .Hingham, MA: Kluwer.

Chomsky, Noam .1977. On Wh-Movement. In Culicover, Peter & Wasow, Thomas & Akmajian, Adrian (eds.), *Formal Syntax*, 71-132. New York: Academic Press.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Foris: Dordrecht.

Chomsky, Noam. 1995. *The Minimalist Program.Cambridge*. MA: MIT Press.

Chomsky,Noam.1998. Minimalist Inquiries: The Framework. MIT *Occasional Papers in Linguistics*, 15 (also published in Martin, R. &.Michaels, D. & Uriagerka, J. (eds.), *Step by Step Essays on Minimalism in honor of Howard Lasnik*, 89-155. Cambridge, MA:MIT Press.

Chomsky, Noam.1999. Derivation by Phase. MIT *Occasional Papers in Linguistics,* 18 (also published in Kenstowicz, Michael, (ed.), (2001). *Ken Hale: A Life in Language*, 1-52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. *Beyond Explanatory Adequacy*. MIT: Unpublished Manuscript.

Chomsky, Noam.2005. *On Phases*. MIT: Unpublished Manuscript.

Chomsky, Noam.2006. *Approaching UG from Below*. MIT: Unpublished Manuscript.

Emonds, Joseph. 1981. Word Order in Generative Grammar. *Journal of Linguistic Research* 1. 33-54.

Fassi Fehri, Abdulkader. 1993. *Issues in the Structure of Arabic Clauses and words*. Dordrecht: Kluwer.

Hallman, Peter. 2002. Passive in Arabic and English. In Bendjaballa, Sabrina & Dressler, Wolfgang U. & Pfeiffer Oskar & Voeikova, Maria (eds.), *Morphology* 2000: *Selected papers from the 9th Morphology Meeting,Vienna,24-28 February 2000*.149-160.

Jaeggli, Osvaldo.1986. Passive. *Linguistic Inquiry* 17, 4. 587-622.

Maalej, Zouhair A. 1999. Passives in Modern Standard and Tunisian Arabic. *MAS-GELLES, Nouvelle series*. 1-24.

McCarthy, John J. A Prosodic Theory of Nonconcatenative Morphology. *Linguistic Inquiry* 12. 373-418.

McClosky, James. 1986. Case, Movement and Raising in Modern Irish. In Goldberg, J. & MacCkaye, S. & Wescoat, M. (eds.), *Proceedings of WC-CFL 4*.190-205.

Ouhalla, Jamal. 1994. *Introducing Transformational Grammar: From Rules to Principles and Parameters*. Londen: Edward Arnold.

Radford, Andrew. 2009. *Analyzing English Sentences: A Minimalist Approach*.Cambrigde: Cambridge University Press.

Rizzi, Luigi. 1997. The Fine Structure of the Left Periphery. In Hageman, Liliane (ed.), *Elements of Grammar*. Dordrecht: Kluwer. 281-337.

Soltan, Usama. 2007. *On Formal Feature Licensing in Minimalism: Aspects of Arabic Morphosyntax*. University of Maryland: College Park. Ph.D.dissertation

Sproat, Richard. 1985. Welsh Syntax and VSO Structure. *Natural Language and Linguistic Theory* 3. 173-216.

Tucker, Matthew A. 2007. *Subject Positions in VSO Languages: Evidence for a Higher Ground*. (Paper presented at the 1st Annual Cornell University Undergraduate Research Colloquium. Cornell University, Ithaca, New York. 11 March 2007.)

Mohammed Odeh AL-Shorafat
Department of English Language & Literature
Faculty of Arts, Yarmouk University
Irbid-21163, Jordan
mshorafat@yu.edu.jo

In SKASE Journal of Theoretical Linguistics [online]. 2018, vol. 15, no. 3 [cit. 2018-13-12].

Available on web page http://www.skase.sk/Volumes/JTL38/pdf_doc/10.pdf. ISSN 1336-782X.

Regular and Irregular Aspects of Grammatical Constructions

Konrad Szcześniak, Palacký University, Olomouc, Czech Republic
University of Silesia, Sosnowiec, Poland

This study focuses on two grammatical forms, the Incredulity Response Construction and the je vidět / vidieť *construction. By analyzing their form and usage, I intend to weigh in on the debate concerning the importance of regularity and unpredictability inherent in grammatical constructions. The debate, initiated by Chomsky's dismissal of idiosyncratic forms as peripheral, has gained momentum especially in recent years, with cognitive linguists uncovering more unpredictable aspects of language forms and claiming that language is primarily idiosyncratic, not regular. As a result, current cognitive linguistic descriptions of constructions highlight what is unpredictable and all but dismiss the regular aspects of constructions. The present study argues that the two sides—the predictable and the form-specific—coexist and thus have obvious implications for models of mental representations of language forms. By extrapolation, a recognition of a balance of the regular and irregular properties is relevant to our theorizing on how speakers learn and use constructions: just as the presence of idiosyncratic properties makes it necessary to memorize special cases of forms, the presence of regular properties can be assumed to ease the learning burden. An accurate description of any construction should attempt to detail both those aspects that can be predicted by rule and those that are endemic to a given construction and must be memorized.*

Keywords: *grammatical construction, idiosyncrasy, regularity, IRC,* je vidět / vidieť *construction*

1. Introduction

The notions of irregularity, idiosyncrasy and unpredictability have received widely different treatments in major theoretical frameworks. They were not given much room in Chomsky's generative approach, interested as it was mainly in those phenomena—elegantly referred to as 'core'—that could be accounted for by means of a small number of general rules. Any idiosyncratic elements were famously dismissed as periphery, an area conceived to house "phenomena that result from historical accident, dialect mixture, personal idiosyncrasies, and the like" (Chomsky 1995: 20).

Cognitive linguistics, on the other hand, has taken the idiosyncratic and irregular as the essence of linguistic competence. First, peripheral phenomena have turned out to be too numerous to dismiss, and as Culicover and Jackendoff (2005: 25) put it, "'periphery' tends to become a tempting dumping ground for any irregularity one's theory cannot at the moment explain." There has been a growing realization that periphery is not just a minor appendage to an elegant core, but the very substance of what language really is. Authors working in cognitive linguistic frameworks have begun to seriously consider the hypothesis that a solid command of a language involves, first and foremost, knowing an immense number of unruly facts, idiomatic forms that cannot be accounted for by straightforward generalizations. As Taylor put it,

> the idiomatic reaches into every nook and cranny of a language, so much so that it might not be outrageous to claim that just about everything in a language is idiomatic to a

> greater or lesser degree and in some way or other. If anything, it is the fully regular, the fully compositional, that is exceptional (Taylor 2012: 71-72).

This new belief has also affected how the study of language is conducted. For example, in Construction Grammar (CxG), all language forms are treated as grammatical constructions defined as "learned pairings of form with semantic or discourse function, including morphemes or words, idioms, partially lexically filled and fully general phrasal patterns" (Goldberg 2006: 5). What cognitive linguists emphasize is that the form-meaning pairings have to be learned because of their inherent unpredictability. It is claimed that memorization from input is necessary even for those constructions that turn out to be "fully regular … and thus predictable" (Goldberg 2006: 12-13).

These two approaches, Generative Grammar and Construction Grammar, are being singled out here because they represent two extreme opposites: while generative grammarians dismiss what is irregular, construction grammarians are not very interested in the regular. These opposing attitudes are illustrated below, first in the discussion of a construction (the Incredulity Response Construction, IRC) claimed by various constructionist authors to be predominantly irregular. I intend to demonstrate that an excessive emphasis on the idiosyncratic has the undesirable consequence of obscuring what are otherwise perfectly regular properties. I then turn to another construction (*je vidět / vidieť* construction), which has been claimed to be a fairly regular instantiation of common patterns found in a language. Here too, although its regular properties are evident, the construction exhibits unusual characteristics that cannot be accounted for by straightforward reference to general rules. Put another way, depending on how they are looked at, and crucially, who looks at them, the regularities and idiosyncrasies in question reveal themselves in different proportions: those interested in regularities predictably stress regularities and authors concerned with the idiomaticity of language focus their attention on why the constructions need to be learned.

In what follows, I attempt to balance the two approaches in analyzing the IRC and the *je vidět / vidieť* construction. The reason behind this choice is that these two forms represent examples of linguistic structures that, at first glance, elicit rather foreseeable reactions: they are held up as perfect examples of either extreme idiosyncrasy (the IRC by construction grammarians) or of forms built around regular patterns commonly found in a given language (*je vidět / vidieť*). However, upon closer reflection, both reveal properties inconsistent with the preliminary description. These properties are identified below based on examples of attested usage of the two constructions, most of which are sentences taken from works of fiction in a number of languages (English, Slovak, Czech, Ukrainian and Polish). Some examples used below are constructed but each has been verified by native speakers.

2. The Incredulity Response Construction (IRC)

2.1 *Irregularities of form and meaning*

In this section, I consider a grammatical construction studied by Akmajian (1984), who referred to its instantiations as *Mad Magazine Sentences*. Uses of the construction such as those in (172) have been held up as examples of idiosyncrasy of both form and meaning. They feature an oblique subject followed by a non-finite verb phrase, typically realized as an infinitive and, sometimes, as a present participle.

(1) a. *What? Him, work? That's a good one!*
b. *Me, mow the lawn? Yeah, right.*
c. *Him, be mature? Good luck.*
d. *Them, working together?*

As these examples illustrate, the construction is used to express emphatic incredulity. More specifically, the construction serves to question the validity of a comment made in preceding discourse:

(2) -*You worry about yourself.*
-*What? Me worry?* (Stephen Coonts, *Deep Black: Biowar*)

Taylor describes it as a means of dismissing a preceding proposition as "absurd, unrealistic, preposterous" (Taylor 2012: 86). In a pre-CxG study of the construction in French, Bally (1905: 8) characterizes the communicative content of the construction as that of surprise and indignation.

Not only is the meaning portrayed as being unique, but as Barðdal and Eythórsson (2012) argue,

> …the semantics of the construction as a whole cannot be derived from either the semantics of the parts or from their form. This particular semantics of disbelief towards a proposition cannot in any way be derived from the fact that there is an oblique argument as a subject, a verb in the infinitive, and a complement, e.g. a noun phrase or a prepositional phrase (Barðdal & Eythórsson 2012: 277).

As far as the form is concerned, Michaelis remarks that the IRC "owes little or nothing to the ordinary English syntax of predication and subordination" (Michaelis 2010: 169). Similarly, Goldberg and Casenhiser (2008) say:

> The form of the construction does not obey general rules of English. For one thing, there is no verb and yet the expression stands alone as a full utterance and conveys an entire proposition. In addition, the accusative case marking is normally used for objects, and yet the initial NP would seem to act as a subject or topic argument (cf. *He's a trapeze artist*?!) (Goldberg & Casenhiser 2008: 344).

Indeed, a glance at other languages shows that there too, the construction exhibits unpredictable properties. In Slovak[1] (3-4), although the subject is realized in nominative case (unlike the oblique subject in English), it is separated from the non-finite verb by means of the conjunction *a* ('and'). This is also found in German (5).

(3) "*Rosalie, Rosalie." Pobavene krútil hlavou. "Tak začneš pracovať?"*
"*Ja a pracovať?!*" (*Kam čert nesmie, pošle Rosalie*, a fanfiction novel https://www.stmivani.eu/)
'Rosalie, Rosalie. He shook his head, amused. So you'll start work?
Me work?!'

[1] I wish to thank Lívia Lakatošová for her help with data from Slovak.

(4) *Ja* *a* *klamať?*
I and to-lie?

(5) *Oh Gott, ich und arbeiten gehen?* (Hardy Manthe, *Die Zeitreisende*)
Oh God, I and to-work to-go?
'Oh God, me and go to work?'

There should be little doubt that in these cases the English IRC and its equivalents in other languages harbor elements of irregularity and that poses a challenge for the language user, who has to learn its form and use by attending to input. The structure to be mastered does not seem to follow straightforwardly from any obvious general rules specifiable in generative fashion. This, however, is not to say that *everything* about the surface form is hopelessly idiosyncratic and thus impossible to master by recourse to any general principles. Upon closer inspection, it is possible to spot aspects of the IRC that are in fact motivated by its function. Also, as I show below, one aspect of its form, though apparently idiosyncratic, actually does follow from the general principles of English grammar.

2.2 *Regularities of form*

2.2.1 *Form of the verb*

At first glance, the non-finite form of the verb appears rather extraordinary and not readily explainable. However, Szcześniak and Pachoł (2015) show that this fairly unusual form of the verb makes sense if it is viewed as an iconic means of suggesting a disruption of the subject-predicate flow. As stressed before, the function of the IRC is to express incredulity, which can be taken to result from a sense of incongruence holding between the subject and the predicate. One way of indicating a syntacto-cognitive dissonance triggered by juxtaposing an agent and an implausible action is to suspend subject-verb agreement. For this argument to hold, it would have to be demonstrated that any uninflected form of the verb is possible in the construction. This is precisely the case: In English, the only real constraint is that the verb should be nonfinite.

(6) a. *Him, cooking? Yeah, right.*
b. *Her, bungee jump? Are you out of your gourd?*

In Slovak (7a), Czech (7b) and Polish (7c), the verb can appear in gerundive form or even as a cognate noun:

(7) a. *Oni a upratovať / upratovanie? To si robíte srandu!*
They and to-tidy up / tidying? Then si-$_{\text{REFL}}$ do kidding!
'Them tidy up? You must be kidding'

b. *Babička a lyžovat / lyžování / lyže?*
Granny and to-ski / skiing / skis-$_{\text{NOUN}}$?
'Granny ski?'

c. *Ja i grać / granie / gra w koszykówkę?*

I and to-play / playing / a game of basketball?
'Me play basketball?'

Another way to put it is that the nonfinite form of the verb is motivated by the need to express the activity in the abstract, not its particular instance performed by the subject. The speaker rejects the possibility of the subject being capable of the activity, hence the uninflected form. This way the subject and the activity considered in the abstract "are placed together in such a way that the connexion between them is at it were brushed aside at once as impossible" (Jespersen 1954: 372).

2.2.2 *Oblique subject*

The oblique argument in the subject position in the English IRC (*Me, remember passwords?*) has attracted much attention by researchers who held it up as a perfect example of idiosyncrasy. Yet the oblique subject turns out to be less bizarre when it is taken into account that it occurs without a finite predicate. When a finite predicate is missing, the accusative case is the default in English (Radford 2009: 216): Subjects with nonfinite predicates (8a) or stand-alone subjects (8b) do not appear in the nominative.

(8) a. *For her to say that is huge!*
b. *Who wants to go first? / Me!*

Interestingly, the same is the case in Danish (9), where the same accusative default applies to subjects without finite predicates. Therefore, predictably, in Danish too, the subject in the IRC takes the accusative form:

(9) *Hvad, mig, lave mad? Jeg kan ikke engang lave røræg!*
What, me, cook? I couldn't even make scrambled eggs!

On the other hand, languages that do not apply the accusative to subjects without tensed predicates do not do so in the IRC either. There is no record of nominative-default languages that would feature an oblique subject in their IRC. It is therefore safe to assume that the oblique subject is not a property of the construction itself, but a consequence of more general rules of the English syntax.

Additional evidence in favor of this view of the oblique subject comes, ironically enough, from English, whose accusative default nature is a relatively recent development. Jespersen shows that the shift from nominative to accusative default occurred after the 18th century and, predictably, uses of the construction from that period can be found in works of fiction with nominative subjects:

(10) Defoe G 44: *Why, his grandfather was a tradesman! he a gentleman!*
Austen P 333: *She a beauty! I should as soon call her mother a wit*. (Jespersen 1954: 372)

An interesting example of a subject whose form can be taken to result from questioning its relation to the rest of the sentence comes from Japanese.[2] Here the function of a lexical item

[2] Example provided by Adrian Mitoraj and Ryoichi Kato.

(such as a pronoun) is signaled by a particle (such as *-wa* often used to mark the subject position), but in a Japanese equivalent (11) of the IRC, the pronoun *ore* ('I') would most naturally occur without a particle, as if to suggest that the speaker denies a link between the pronoun and what follows it (the link is normally obvious from verb inflections and argument positions, arranged to show who is doing what to whom).

(11) なんだって? 俺? 嘘つき?
Nan da-tte? Ore? Usotsuki?
What da-COPULA-tte-PARTICLE I liar
'What? Me, a liar?'

One final point is in order considering the form of English subjects in general. As one anonymous reviewer has pointed out, the range of subjects also includes possessive forms found with gerunds (*their smoking, his whining,* etc.). It must be noted that gerunds represent a phenomenon transitional between verbs and nouns. In terms of the external syntax of *-ing* nominalizations, they are "obviously and unambiguously nominal in character" (Taylor 1996: 270). They function as subjects and objects, they can take genitive heads (*The president's handling of the issue*), and they tend to be open to pluralization (*sightings, killings*). On the other hand, phrase internally, they are "rather more verbal in character" (Taylor 1996: 270), as they allow direct object complements (12).

(12) *his handling the new Congress campaign* (Brian Stoddart, *A People's Collector*)

The main point is that such possessive subjects are not idiosyncratic, since event participants associated with nouns are normally realized in possessive form, as in *our existence, the bullet's trajectory, Lincoln's assassination* (Langacker 2008: 505). The choice of the form of a subject is a fairly direct consequence of general rules of syntax.

2.2.3 *Conjunction*
The use of the conjunction 'and' between the subject and the verb (in languages like Slovak or German) may appear to be an obvious irregularity, but it turns out to be much less idiosyncratic in light of the function that the construction is known to serve. That is, it can be argued to compound the separation between the subject and predicate. Below, the use of the conjunction is illustrated by examples from an additional three languages:

(13) a. *Hij en lezen?* (Dutch)
He and read?
'Him, read?'

b. *Mida? Mina ja varastan su rahakoti? Kuidas sa julged?* (Estonian)
What? I and steal your wallet? How you dare?
'What? Me steal your wallet? How dare you?'

c. *Tak, jasne! On i pracować.* (Polish)
Yes sure He and work
'Yes, sure! Him, work.'

While *and* normally serves to conjoin, here its unusual position makes it look more like a wedge between the two parts of the construction. This counter-intuitive effect is due to the fact that the operation of conjunction should typically involve two elements of a comparable status (e.g. noun with noun, clause with clause, etc.), and here attempting to conjoin two conspicuously disparate elements results in juxtaposition, which highlights the contrast.

2.2.4 *Echoing*

One final trace of a linguistic preference for regularity normally overlooked in cognitive linguistic analyses of the IRC is echoing. As one anonymous reviewer pointed out, the wording of the construction is a reflection of an utterance made in the preceding discourse, either by the speaker or her interlocutor. Since the very purpose of the IRC is to question the validity of a proposition introduced earlier, sentences built around the construction do not normally appear abruptly, unprompted by any clear trigger. They are invited by overt statements or at the very least by implicatures deduced from utterances produced prior to an IRC reaction. In any case, whenever the construction is employed, there is a background proposition, a sort of foil on which an IRC sentence is modeled.

Additionally, as the reviewer pointed out, what is copied from a prior utterance is not necessarily a subject and its predicate, but in some cases it can be the predicate and its complement:

(14) *Chytia ho ľahko. -Jeho a chytiť!*
Catch-3RD.PL him easily him and to-catch
'They (will) catch him easily. -Catch, him!?'

As example (14) shows, the formal match between a IRC sentence and its foil is not exact, but the differences can be explained in terms of general rules of Slovak grammar: The object of the verb appears sentence-initially (a position dictated perhaps by the IRC preference for event participants before verbs). This initial position accounts for the form *jeho* instead of *ho* (found in the foil sentence), but the form *ho* cannot occur in stressed positions. The sentence could echo its foil more closely (*Chytiť! Jeho!*), and here too the strong form of the pronoun *jeho* is selected for purposes of stress, which in itself is a regular and rather natural device used to highlight the logical contrast that the construction serves to convey. The general point is that some of the irregularities found in the construction are far from being completely intractable, and certainly not as idiosyncratic as they are claimed to be in cognitive linguistic accounts.

3. The form and use of the *je vidět* / *vidieť* construction

3.1 *Regular aspects of form*

This section turns to a construction found in Slavic languages, where it appears in comparable form, although it exhibits language-specific peculiarities in each. Perhaps the first mention of the construction under consideration focuses on its Czech variant analyzed in Bělič (1954, 1969), where it is referred to as a "construction of the type *je vidět Sněžka / je vidět Sněžku*" (Bělič 1969: 37). Here I use the name '*je vidět* construction' for short.

Unlike the IRC, the *je vidět* construction appears regular enough. Its form 0 follows an otherwise common Czech pattern VERB + INFINITIVE found in expressions like *zůstat ležet*

'remain lie-INF' or *viděl ho utíkat* 'saw him escape-INF'. This is how the construction is presented in Dušková (2004), where it is treated as a fairly regular and predictable example of Czech subject complement constructions.

(15) *Je vidět štít.*
Is to-see summit
'The summit can be seen.'

Also regular is the tense flexibility of the copula *být* 'to be', such that sentences featuring the construction can appear in any tense or mood, just like any regular Czech sentence. The following are expressions of the proposition *The summit can be seen* and in (16e) *Let the ugly truth be seen*.

(16) a. *Bylo vidět štít*. (Past)
b. *Je vidět štít*. (Present)
c. *Bude vidět štít*. (Future)
d. *Bylo by vidět štít*. (Conditional)
e. *Ať / nechť je vidět holá pravda*. (Imperative)

The form of the construction in Slovak is the same. Here too the verb *byť* 'to be' is followed by the infinitive *vidieť* (17), and it is available in all the expected tense constructions.

(17) *…na brehu jazera bolo vidieť kríže* (János Gáspár, *Slovenská čítanka*)
On bank-LOC lake-GEN was to-see crosses
'On the bank of the lake one could see crosses'

3.2 *Irregular aspects of form*

However, upon closer examination, the form of the construction turns out to be rather idiosyncratic and not so predictable from more general rules of Czech grammar: The language user needs to learn that the fixed element *být* 'to be' is only allowed in third person:

(18) a. *Na nebe je vidět mraky*.
In sky is to-see clouds
'In the sky clouds can be seen.'

b.**Na nebe jsem / jsi / jsme / jste vidět mraky*.
In sky am / are-2SG / ARE-1PL / ARE-2PL to-see clouds

Then it should be stressed that as Bělič's name of the construction suggests, the verb can take an object in either nominative (*Sněžka*) or accusative (*Sněžku*). It is also interesting to stress the inclusion of fairly fixed perception verbs in the construction. Apart from the two most frequent *slyšet* and *vidět* 'hear' and 'see', the construction is also attested with *cítit* 'feel', *poznat* 'recognize' and *rozumět* 'understand'. Crucially, many perception verbs are not found in use (e.g. *pozorovat* 'observe': *Je pozorovat *Sněžku* 'Sněžka can be observed'). The degree of unpredictability becomes more evident when data from other Slavic languages are

considered. At first glance, the Polish variant of the construction looks like a structural calque of the Czech and Slovak equivalents:

(19) …*ledwo go było widać*. (Stefan Żeromski, *Ludzie bezdomni*)
barely he-ACC was to-see
'…he was barely visible'

However, there is one striking difference, namely the fact that where Czech and Slovak use 'general purpose' *see* and *hear* verbs, Polish employs a specialized form *widać* related to the verb *widzieć* 'to see' (a cognate of the Czech / Slovak pair *vidět / vidieť*), but *widać* is not used in infinitive function outside this construction (and by the same token, the verb *widzieć* cannot be used in it).

Similarly, in Russian[3], the copula is followed not by infinitives, but by the adverb *видно* (*vidno*, lit. 'clear, bright') or *слышно* (*slyshno*, roughly 'audibly') derived out of the verbs *видеть* (videť 'to see') and *слышать* (*slyshať* 'to hear'). In Ukrainian too, the adverbs *видно* (*vydno*) or *чутно* (*chutno*, roughly 'audibly') are used:

(20) …*було видно все дерева в садку*. (Taras Shevchenko, *Mykola Dzheria*)
…bulo vydno vse dereva v sadku
…was clear all trees in orchard
'All trees could be seen in the orchard'

(21) …*було чутно шум*. (Vladislav Ivchenko, *Khymery dykoho polya*)
…bulo chutno shum
…was audible noise
'Noise could be heard'

The Ukrainian construction comes with an extra twist in that sentences expressing audible perception are possible with either the adverb *чутно* or, like in Czech and Slovak, with the infinitive *чути* (*chuty* 'to hear').

(22) *Було видно, було чути / Як реве ревучий*. (Taras Shevchenko, *Zapovit*)
Bulo vydno, bulo chuty / yak reve revuchyi
Was clear, was to-hear / how roars roaring
'One oculd see, one could hear / how the roaring roared'

Similarly, in Slovak, apart from the form *vidieť* 'to see', the construction also features the adverbial *vidno*, but conspicuously not a corresponding adverbial for expressing auditory perception.

The discussion of this brief sample should conclude with the observation that the irregularities found in Polish, Russian and Ukrainian are not beyond any constraints. While the complements of the copula may seem exceptional, especially compared to the Czech and Slovak infinitives, they can be considered small pockets of regularity in themselves. As the examples presented here show, the *see-hear* complements (the most frequent complements of the copula) come in pairs of mutually related elements. Polish does not use Czech/Slovak-style

[3] Russian and Ukrainian examples provided by Monika Hałaś.

infinitives and instead employs special-purpose forms, but it does so consistently so the complements *widać / słychać* form a tandem sharing the same suffix. In the same vein, the Russian members of the pair *видно* and *слышно* both have adverbial suffixes, which is also true of Ukrainian (with the proviso that it allows an infinitive complement *чути,* perhaps under the influence of Slovak).

4. A balance of the regular and irregular

The significance of the peculiarities of the constructions discussed above is that they can be expected to affect the learning process. The exceptional facts cannot be 'figured out' from the logic of Slovak or Czech grammar. An important point stressed by cognitive linguists (e.g. Bybee 2010, Goldberg 2006) is that language knowledge follows from language use (so called usage-based approach): a person has no choice but to closely attend to input, a rich source of details of form and meaning which are not always predictable from the general rules of grammar. Thus, in the last three decades or so, cognitive linguistic research has departed from the generative preoccupation with regularity and its hope of explaining how children learning their mother tongue may acquire more material than they can encounter in input. As a consequence, opinions like the one expressed by Pinker (1995) below are now treated as implausible, if not naïve:

> …the child would set parameters on the basis of a few examples from the parental input, and the full complexity of a language will ensue when those parameterized rules interact with one another and with universal principles (Pinker 1995: 173).

This sentiment is evident in many cognitive linguistic discussions. For example, Taylor (2012: 68) looks at the usage patterns of a number of individual words and identifies unexpected usage patterns that he says "cannot be predicted from the categorization of the word as a noun, verb, or whatever, in association with syntactic rules operating over these categories." He goes on to conclude that "knowledge of a language can *only* be attained through exposure to actual usage events, whose particularities are noted and laid down in memory" (ibid., my italics).

One important consequence of this conviction is the cognitive treatment of each construction separately and dismissal of any regular correspondences that were previously claimed to hold between constructions. Thus, while the active and the passive voice constructions were traditionally treated as related constructions, Hilpert (2014: 42) claims that it is "difficult to maintain the idea of a grammatical rule that systematically links both constructions," and he goes on to stipulate that "the Passive is a construction in its own right, a generalisation that speakers have to learn as an independent unit of grammatical knowledge." Similarly, in her critique of transformational approaches to diathesis alternations (e.g. *Mina sent Mel a book* being derived from *Mina sent a book to Mel*), Goldberg (2002: 327) expresses the view that "it is profitable to look beyond alternations and to consider each surface pattern on its own terms." She also looks at the ditransitive alternation and explains that one should focus mainly on finding generalizations applying to each surface pattern separately, and not point out any correspondences between them. Perhaps the most extreme example of this 'separationist' approach to language forms is Taylor's (2012) proposal that morphological derivations and any etymological relations between lexical items are, from the language user's point of view, irrelevant. He argues that because "Not everyone is aware of the historical

development of the words of their language" (2012: 230), knowledge of relations between words has no bearing on the speaker's ability to use them. He explains, "observing how a speaker uses a word would give us no clue as to how she judges the relatedness of its meanings. … whether she perceives them to be related is immaterial" (ibid.).

Of course, putting too much faith in correspondences, regularity and predictability is a risk, and users most likely approach the learning challenge allowing for some potential unpredictability, remaining open and receptive to exceptions. Still, it makes sense to entertain predictions about how a construction like the *je vidět / vidieť* may behave with 'audible perception' meanings, based on its behavior with the corresponding 'visual perception' function, even if such predictions have to be revised in some cases. Language users may need to confirm their expectations for each use individually, but that is more helpful than discovering all the details of their use in the two constructions completely from scratch, without any preconceived expectations. Learning new facts cannot happen without some reliance on one's prior knowledge, based on input alone. Indeed, the dismissal of prior knowledge is baffling, given how diverse its kinds are. A good example of how much background knowledge is enlisted while handling novelty can be found in Štekauer (2005):

> "the (degree of) acceptability/non-acceptability and, consequently, meaning predictability/unpredictability (conceived of as a cline) depend on the interaction of linguistic knowledge (knowledge of the meanings of morphemes, including affixes, knowledge of productive Onomasiological, Word-Formation, and Morphological Types, knowledge of acceptable Onomasiological Structure Rules, etc.) and extra-linguistic knowledge (including knowledge of real and unreal, tangible and intangible objects of extra-linguistic reality, and their place in the narrower and broader system of relations and interactions)" (Štekauer 2005: 60).

5. Conclusions

My goal here has been to show that in their opposing stances, both generative grammarians and construction grammarians could be said to be guilty of the same mistakes. While the former unfairly dismissed the irregular and the idiomatic, the latter underestimate the significance of regular aspects of language. The constructionist ambition to uncover peculiar aspects of constructions has prompted some authors not only to overemphasize the idiomatic, disregarding the regular, but also to misidentify what is regular as being idiosyncratic and unexplainable, as was the case of the oblique subject in the English IRC.

The mistake of extreme versions of the cognitive linguistic focus on idiosyncrasy and the consequent emphasis on learning from input is that little attention is paid to how language learners bring their knowledge of the system to the task of learning new language forms. It should not be too much of an exaggeration to claim that in their preoccupation with the peculiar minutia and dismissal of regularities, cognitive authors seem to be approaching language as a system with no rules, only exceptions.

References

Akmajian, Adrian. 1984. Sentence types and the form—function fit. *Natural Language and Linguistic Theory* 2. 1-23.

Bally, Charles. 1905. *Précis de Stylistique*. Genève: A. Eggimann et Cie.

Barðdal, Jóhanna & Eythórsson, Thórhallur. 2012. Reconstructing syntax: Construction grammar and the comparative method. In H. C. Boas & I. A. Sag (eds.), *Sign-based Construction Grammar*, 257-308. Stanford: CSLI Publications.

Bělič, Jaromír. 1954. *Studie a práce linguistické. K šedesátým narozeninám akademika Bohuslava Havránka*. Prague: Nakladatelství Československé akademie věd.

Bělič, Jaromír. 1969. *Kultura českého jazyka*. Liberec: Severočeské nakladatelství.

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

Chomsky, Noam. 1995. *The minimalist program.* Cambridge, MA: MIT Press.

Culicover, Peter & Jackendoff, Ray. 2005. *Simpler syntax.* Oxford: Oxford University Press.

Dušková, Libuše. 2004. *Syntactic constancy of the subject* complement 1: A comparison between Czech and English. Linguistica Pragensia 14(2). 57-71.

Goldberg, Adele. 2002. Surface Generalizations: An alternative to alternations. *Cognitive Linguistics* 13(4). 327-356.

Goldberg, Adele. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

Goldberg, Adele & Casenhiser, Devin. 2008. English Constructions. In April McMahon and Bas Aarts (eds.), *Handbook of English Linguistics*, 343-355. Oxford: Blackwell Publishing.

Hilpert, Martin. 2014. *Construction Grammar and its Application to English*. Edinburgh: Edinburgh University Press.

Jespersen, Otto. 1954. *A Modern English Grammar. Part III: Syntax*. London: Allen & Unwin.

Langacker, Ronald W. 2008. *Cognitive Grammar. A basic introduction.* Oxford: Oxford University Press.

Michaelis, Laura. 2010. Sign-Based Construction Grammar. In Berndt Heine and H. Narrog (eds.), *The Oxford Handbook of Linguistic Analysis*, 155-176. Oxford: Oxford University Press.

Pinker, Steven. 1995. Language Acquisition. In Lila R. Gleitman and Mark Liberman (eds.), *An Invitation to Cognitive Science*. Cambridge, MA: MIT Press.

Radford, Andrew. 2009. *Analysing English Sentences: A Minimalist Approach*. Cambridge: Cambridge University Press.

Štekauer, Pavol. 2005. *Meaning Predictability in Word Formation: Novel, context-free naming units*. Amsterdam: John Benjamins Publishing.

Szcześniak, Konrad & Pachoł, Małgorzata. 2015. What? Me, Lie? The Form and Reading of the Incredulity Response Construction. *Constructions* 1. 1-13.

Taylor, John R. 1996. *Possessives in English: An Exploration in Cognitive Grammar*. Oxford: Oxford University Press.

Taylor, John R. 2012. *The mental corpus. How language is represented in the mind.* Oxford: Oxford University Press.

Konrad Szczesniak
Institute of English Studies
Silesian University
Sosnowiec
Poland
konrad.szczesniak@gmail.com

In SKASE Journal of Theoretical Linguistics [online]. 2018, vol. 15, no. 3 [cit. 2018-13-12]. Available on web page http://www.skase.sk/Volumes/JTL38/pdf_doc/11.pdf. ISSN 1336-782X.

The complex sentence across written genres from native and nonnative contexts; a corpus-based study

George Kodie Frimpong, University of Ghana

This is a corpus-based study which sought to investigate the use of the complex sentence and its immediate internal clause combining mechanisms across three genres from the written components of the Ghanaian and British versions of the International Corpus of English. The study is based on the theoretical argumentation from functional register perspective that the distribution of linguistic elements across genres is functionally motivated. Texts examined were sampled from Academic Natural Science, Academic Social Science and Administrative Writing. The complex sentence received the attention of this study because it was the dominant structural sentence type across the three genres from these two native and nonnative contexts, a phenomenon we have argued in this study to be functionally motivated. An investigation of two internal clause combing dynamics among the three genres reveals that whereas the Academic Natural Science genres across Ghanaian and British corpora rely a lot on adverbial clauses, the Ghanaian and British academic social and administrative genres vary in their clause combining preferences, a phenomenon we have associated with genre-internal variability.

Keywords*: complex sentence; clause complementation; dialectal variation; register variation; written genre*

1. Introduction

Complex structures are of a particular importance in English for academic purposes and English as a second language. Whereas complex structures facilitate the expression of complex thoughts in academic texts, in second language learning, they indicate one's level of proficiency. And the complex sentence is one important complex structure which embeds most other complex structures, including clauses and phrases. The observation is that it is more difficult, even for advanced nonnative speakers of English, to combine clauses in their speech to create complex sentences (Vercelloti & Packer 2016). However, work in this area has focused most predominantly on clause types, clause combination, finite and nonfinite clauses, subordination and the phrase structural (Biber & Gray 2010; Frimpong 2015; Vercellotti & Packer 2016). The complex sentence, as a structure, has received little attention. Apart from its treatment together with other types of sentence in grammar textbooks (Quirk et al. 1985; Wardhaugh 1995), the complex sentence is under-researched, especially to explore its usage dynamics and its internal structure across genres from different cultural perspectives – an engagement this study seeks to do. It explores the complex sentence and the complexity of its internal clause combining manifestations across genres from native and nonnative contexts.

It has been observed in the literature that the sentence has generally been understudied (Simpson 2004). One of the reasons why the sentence and its subtypes have not received much empirical attention may be attributed to the disagreement about the status of the sentence in the literature. In descriptive grammar (Quirk et al. 1985: 717) the sentence is considered the highest grammatical unit, followed by the clause, the phrase, the word and the morpheme on the grammatical rank scale. Elsewhere, however, especially in functional linguistics, the argument has been raised about the concept of the sentence as a stretch "of written text bounded by full

stops or the equivalent" (Thompson 2014: 23). The argument is that this definition makes the sentence a property of written texts since the features of the sentences mentioned in this definition are limited to written texts only. Thus, proponents of this argument consider the sentence as a problematic construct since its categorization has been based on written text (Eggins 2004: 126; Fawcett 2004). The problem really, as they rightly observe it, is with describing the sentence within spoken texts. As Thompson (2014: 22) puts it, "the sentence is an idealization of the written language which it is often difficult to impose on spoken language." Based on this argument, certain grammars (c.f. Biber et al. 1999) intentionally avoid treating the sentence in any detail. However, much as this argument makes a lot of sense, it does not disqualify an empirical study of the sentence, especially when the investigation is limited to written data like the current study. The more reason why it is compellingly necessary to investigate the sentence (especially the complex subtype) particularly across written genres is that the written genre by its nature is observed to be complex (Biber 1988: 5). This means that discussions of complexity in written texts should begin with the sentence. This is because, it is partly through such investigations that we will be able to ascertain the levels and the nature of the complexity of written genres within the scope of the sentence.

One work that investigates aspects of the complex sentence empirically is Wiredu (2012). Though his study is one sided by focusing only on newspaper editorials from one newspaper from Ghana, he makes interesting observations about the relationship between language and context. His data reveals that the complex sentence is the most preferred type of sentence in the newspaper editorial genre, an observation he explains to be functional. His findings are much in line with Hopper and Traugott's (2003: 177) observation that the organization of complex structures in text is genre-based and that certain complex sentence types, are noted to be associated particularly with written genres. A comparison of the use of the complex sentence across academic genres (Academic Social Science and Academic Natural Science) and a nonacademic written genre (Administrative Writing) will enable us to contribute to the debate about structural complexity across written genres and language use across native and nonnative dialects.

2. Between dialectal variation and register variation

Linguistic variation has been explored from a variety of perspectives (Biber & Conrad 2009). Two of these perspectives are dialectal variation and functional register variation. From the perspective of dialectal variation, scholars in the new/world Englishes, especially, have argued that varieties of a language used in different geographical locations essentially reflect variable linguistic choices (Kachru 2006, Schneider 2007). In fact, Kachru was among the first scholars to make this observation. He explained dialectal variation between native and nonnative Englishes using the concentric circle according to which varieties of English around the world are classified into the inner circle (i.e. the varieties of English spoken by native speakers), the outer circle (i.e. the variety of English spoken in former British colonies where English is used as second/official language) and the expanding circle (places where English is used as an international language).

In the past decade, one of the most compelling arguments for explaining varieties of English around the world has been advanced by Schneider (2003). In his developmental dynamic model, Schneider (2003: 233) has argued that dialects of English undergo 5 phases of growth from the day of their implantation on a new soil to the day they become fully-fledged

varieties of the English language. The crux of his argument is that "consistent sociolinguistic and language-contact conditions" involving "English speaking settlers and an indigenous people and their languages" work together to shape the English language introduced to the area into a novel dialect of the English language (Schneider 2003: 233). From this evolutionary perspective, one can place the English used in Ghana at the endonomative stabilization phase, though there are still lingering signs of the nativaization phase. That is, though the complaints about falling standards of English which characterize the nativization phase have minimized, many Ghanaian users of English are not aware that they speak a Ghanaian variety of English. And yet, Dako (2003) has documented copious examples of Ghanaian local expressions used in English texts produced by educated Ghanaians. Huber (2004) has also observed grammatical features which are peculiar to the English used in Ghana. These and other linguistic features have been noted to be uniquely Ghanaian (see Sey 1973; Koranteng 2006; Huber 2014). The audacity of Ghanaian users of English to innovate with the language and use local norms in newspapers and novels as captured in the literature suggests the Ghanaian English is at Schneider's endonormative phase which "is marked by the gradual adoption and acceptance of an indigenous linguistic norm, supported by a new, locally rooted linguistic self-confidence" (2003: 249).

The observation in the literature is that language change from dialectal perspective exists not only in the creation and use of entirely new linguistic forms but also in a restructuring of the linguistic features of the old native English stock (Huber 2014: 104). Thus, Huber (2014: 103) would argue that in the area of the relative clause construction, the Ghanaian English is different from the British English only in the ways certain features of the relative clause structure has been reorganized. He observes, for instance, a variation in the way the relativizers – *which*, *who* and *that* – have been used.

From functional register perspective, on the other hand, linguistic variation has been construed as variation in linguistic choices as a result of variation in the situational context of language use (Bloor & Bloor 2004: 2). That is, certain linguistic features have been preferred in a particular text in comparison with others because those features are suitable to the situational background of the text (Biber & Conrad 2009: 33). Besides, those linguistic features enable the particular text (register) to achieve a certain communicative function. The situational context has been theorized in systemic functional linguistics into the three register variables of *field*, *tenor* and *mode* according to which a text reflects an area of activity (field), a social contact and status (tenor) and the channel of communication (mode) (Eggins 2004: 9).

From another, perhaps, more simplified register perspective, Biber & Conrad (2009: 31) explain register variation from a set of seven situational variables which include the participants of the communication, the relationship between the participants of the communication, the mode of communication, the topic, the setting, production and comprehension circumstances and the purpose of communication. These set of situational variables are deemed to constitute the functional background that influence linguistic choices across texts. As they put it

> …texts from different registers are produced in fundamentally different circumstances, for fundamentally different communicative purposes; and as a result, these texts often use completely different sets of linguistic features that are functionally appropriate for those circumstances and purposes (Biber & Conrad 2009: 256).

The bridge between dialectal variation and register variation which has not received adequate attention in the literature, is the focus of this study on register variation across genres from Ghanaian and British dialects of the English language. The argument in the literature is that texts from the same genres share similar linguistic features across dialects of a language. Biber & Conrad (2009: 12) capture this argument more cogently as follows

> [R]egardless of any dialect differences, speakers using the same register are doing similar communicative tasks; therefore, in most basic respects the characteristic language features used in a given situation are similar across speakers from different dialects.

This similarity is of course different from core grammar in the sense that whereas core grammar is a form of standard linguistic features shared across all varieties of a particular language, register similarity is about the organization of features of core grammar based on functional parameters. Our data from the three genres across the two dialectal contexts should be able to give us insights about linguistic variation at the dialect-register interface.

3. Language and function

The relationship between language and function has been explored from various perspectives in linguistics. Interestingly, each of the conceptualizations of the relationship between language and function brushes with an aspect of context. From semantico-grammatical perspective, linguistic function is considered a property of the co-text according to which a grammatical element is assigned a functional role (actor, goal, theme, instrument, senser, etc.) based on its placement in the syntagm of the sentence. From pragmatic perspective, function is based on the cultural context where meaning of an utterance is not only inferred from the co-text but also from the circumstantial background of the speech act event.

Language function in this paper is taken from the register perspective (Bibber & Conrad 2009; Halliday & Matthiessen 2004; 2014). From this perspective, the function of a linguistic form is based on the situational context. That is the situational background of a text makes certain linguistic features more suitable in a particular text than in others. For Biber and Conrad, the relationship between language and function is the association of 'pervasive linguistic features' with a language variety for functional purposes. This is what Halliday & Matthiessen from the systemic functional linguistic perspective capture as follows

> A register is a functional variety of language (Halliday 1978) – the patterns of instantiation of the overall system associated with a given type of context (a situation type). These patterns of instantiation show up quantitatively as adjustments in the systemic probabilities of language; a register can be represented as a particular setting of systemic probabilities (Halliday & Matthiessen 2004: 27, 28).

The key components of registerial function include '*patterns of instantiation' of linguistic variables* (i.e. linguistic choices), a *situational context* and *quantitative variability* of linguistic features across texts. In this study, structural sentences and the internal patterns of clause combination are identified and counted across the three genres from the two dialectal contexts. Patterns of realization are compared both across the genres and across the two dialects. It is anticipated that the shared situational contexts will induce similar patterns of distribution of

these grammatical features investigated across the same genres from the Ghanaian and British corpora.

4. The complex sentence and its immediate internal constituents

Structurally, sentences in English may be grouped into simple sentences and non-simple sentences (Quirk et al. 1985: 987). As the name suggests, a simple sentence has a simple internal clause structure, having only one independent clause. Non-simple sentences are multi-clausal sentence structures made up complex, compound and compound complex sentences (Wardhaugh 1995: 115). The major difference between the complex sentence and the compound sentence is that whereas the compound sentence combines at least two independent clauses, the complex sentence combines one independent clause with at least one dependent clause. The compound complex sentence, of course, has features of the complex and the compound sentences whose structure has at least two independent clauses and at least one dependent clause (Quirk et al. 1985: 987). Examples of these sentence types are as follows:

(1) Simple sentence: In addition algae such as Chlorella, Euglena, Oscillatoria, Ulothrix and Stigeoclonium will be abundant. (ICE-GB:W2A-021 #1)
(A MONO CLAUSAL SENTENCE)

(2) Complex sentence: For National Insurance purposes, your working life is the number of tax years from the beginning of the tax year in which you reached the age of 16 up to the end of the one before you reach pension age (or death, if earlier). (ICE-GB:W2D-004 #1)
(A MULTI-CLAUSAL SENTENCE WITH A FINITE INDEPENDENT CLAUSE AND 2 EMBEDDED CLAUSES)

(3) Compound sentence: Students listen to radio programmes, watch television and engage in conversation with friends and family. (ICE-GH:W2A-001#18)
(A MULTI-CLAUSAL SENTENCE WITH THREE COORDINATED INDEPENDENT CLAUSES)

(4) Compound complex sentence: These responses reveal a problem that deserves attention and the problem might be posed in the form of a question (ICE-GH:W2A-001#16)
(A MULTI-CLAUSAL SENTENCE WITH 2 INDEPENDENT CLAUSES AND 1 DEPENDENT CLAUSE)

Of these four structural sentence types, the complex sentence and its allied compound complex sentence are the most versatile and so have attracted my research curiosity. That is, a closer look at the sentences above reveals that whereas the compound sentence is a coordination of two equivalent clauses both of which are finite and the simple sentence is a single finite independent clause, the complex sentence is a layered grammatical structure whose complexity may be varied. In sentence (2) above, the complex sentence, for instance, has one finite independent clause which is the entire sentence and two subordinate clauses (…*in which you reached the age of 16* and *before you reach pension age*) which are rank shifted within noun phrases. Rank shifting, also referred to as embedding, is an interesting phenomenon where

clauses, which are higher in the grammatical rank than the phrase, are used to modify a phrase (Halliday & Matthiessen 2014: 382). However, in alternative perspectives, we have complex sentences in which the subordinate clauses are dependent rather on the main clause or are clause constituents. That is, it is possible for dependent clauses to be subjects, objects/complements or adjuncts as highlighted in sentences (5) to (7) below.

(5) **Expanding one's knowledge of the world and acquiring information on lecture strategies which will allow one's mind to work back and forth** seemed not to feature in these responses. (ICE-GH:W2A-001#99)
(A COMPLEX SENTENCE WITH PARTICIPIAL CLAUSE AS A CLAUSAL SUBJECT)

(6) The paper reveals **that several students have not given any thought to academic listening as a skill that can enhance their ability to derive maximum benefits from lectures during lecture hours**. (ICE-GH:W2A-001#3)
(A COMPLEX SENTENCE WITH COMPLEMENT CLAUSE AS CLAUSAL OBJECT)

(7) **If you do not receive an award from the Department of Health** you may still be eligible for an award from your LEA for certain courses. (ICE-GB:W2D-003 #15)
(A COMPLEX SENTENCE WITH ADJUNCT AS DEPENDENT CLAUSE)

This means that a complex sentence offers the possibility for four functional types of dependent clauses: the relative clause postmodifies the noun phrase; the clausal subject, the clausal object and adverbial clause are dependent clauses which function as clause constituents (Payne 2011: 334). Of these clause types, clausal subjects and clausal objects/complements, as in sentences (5) and (6), are required in their main clauses, unlike adverbial clauses. Particularly, clausal objects/complements are considered special types of embedding by which a dependent clause is embedded within a matrix clause (Payne 2011: 337). These are clauses that are used after complement-taking predicators and after copula verbs (subject complements). It is not automatic that complement clauses will be more preferred in a particular genre simply because they are required in their main clause. This is because, complement clauses compete with phrases in the structure of the sentence. But more interestingly, the fact that adverbial clauses and other adverbial structures are optional in a sentence makes their choices particularly informative. This study focused on the use of clausal objects/complements and circumstantial adverbial clauses across the two academic genres and the legal administrative genre explored in this study. Payne (2011: 344) explains the distinction between complements and adverbial clauses as follows

> 'Adjunct' is a good term since the term "Complement" implies completion, and a phrase or clause does not express a complete thought until all its Complement positions are filled. On the other hand, adverbial clauses and phrases attach to already complete clauses. The adverbial clause simply adds some additional information to what is expressed in the other clause.

It is hoped that this engagement will enhance our understanding of the variation that exists across these genres from native and nonnative English contexts. The comparative dimension of this study, it must be noted, is not an end in itself since the linguistic variables explored are

not interchangeable or competing concepts. The aim is rather to explore the stylistic or register implications in the distribution of these structures. Since adverbial structures are optional in the syntagm of the sentence the nature of their use – minimal or otherwise across a set of genres – should engage our linguistic curiosity.

Thus, the complex sentence was selected for this study not only because it is the preferred sentence pattern in our written genres but also because the internal complexity of the complex sentence offers an opportunity to explore the structural complexity of the three genres in much detail.

5. Methodology

This work is a corpus-based comparative study. The three genres investigated (Academic Natural Science, Academic Social Science and Administrative Writing) were culled from the written components of Ghanaian and British parallel International Corpus of English (ICE) corpora. In order to facilitate comparative study of the Englishes spoken around the globe, compilation of the various national ICE sub-corpora focused on comparable features: about 1 million words, featuring 15 spoken genres and 17 written genres, annotated similarly in key areas. Thus, selection of texts for comparison was not difficult.

For the purpose of this study, we selected the first five texts for each of the 3 genres investigated across the two dialectal contexts. This means that in all 30 texts were analyzed in this study. For each of the texts selected, 500 words were excerpted from the body of the text for detailed analysis. This was done by scooping away the first 200 words after which the following 500 words were gleaned for the study. This is in line with Biber & Conrad's (2009: 6) argument that for register analysis one does not really need an entire text because the grammatical features that are normally investigated are found in any part of the text. However, the caution was to guard against scraping the introductory parts of the texts studied for fear that they may not give us a true picture of the genres investigated. The excerpted texts realized a total of 625 sentences which have been described in this study.

6. Coding

The first step towards the analysis was to manually code the corpus data excerpted with the grammatical features explored. In order to relate the complex sentence with the other types of sentences, simple, compound and compound complex sentences were also coded into the data. In the same way, the internal clause constituents were identified and categorized into complement clause and adverbial clause structures.

To begin with, the original notations for the sub-genres investigated were maintained. These notations, which are used in the analysis, are captured in Table 1 below.

Table 1: Parallel Notations Across ICE GH and ICE GB

NOTATION	**INTERPRETATION**
ICE-GH:W2A-021-025	ICE Ghana, Written Corpus, Academic Natural Science, Text 21 to 25
ICE-GH:W2A-001-005	ICE Ghana, Written Corpus, Academic Social Science, Text 1 to 5
ICE-GH:W2D-001-005	ICE Ghana, Written Corpus, Administrative Writing, Text 1 to 5
ICE-GB:W2A-021-025	ICE Great Britain, Written Corpus, Academic Natural Science, Text 21 to 25
ICE-GB:W2A-011-015	ICE Great Britain, Written Corpus, Academic Social Science, Text 11 to 15
ICE-GB:W2D-001-005	ICE Great Britain, Written Corpus, Administrative Writing, Text 1 to 5

For the sentence types, since we were dealing with written data, determination of sentence types was not difficult. One observes that written texts, unlike spoken texts, are carefully produced to ensure that grammatical structures are fully realized and marked appropriately. Classification of sentence types were based on the two models by Quirk et al. (1985: 987) and Wardhaugh (1995: 115) which identify simple, complex, compound and compound complex sentences based on the number and types of clauses in them. For clausal objects/complement and adverbial clause structures, we scrutinized the nature of clause combining strategies in the complex sentence across the genres from the sub-dialectal corpora. Downing & Locke (2006: 275) argue that embedding, rather than adjunction, offers the tightest structural integration. We distinguish between clausal object/complement clause embedding and adverbial clauses based on Payne (2011) whose distinction between clausal object/complement and adjuncts is captured under section 4 above.

After these grammatical features were coded and checked by a research assistant, attested structures were compared and described using interpretive qualitative methodology. This involved relating attested grammatical patterns with the situational context within which the genres were generated (Biber & Conrad 2009: 6).

7. The Genres studied

The genres explored – academic natural science, academic social science and administrative writing – are all written genres. As written genres, they are deemed equally well planned and elaborate. However, these genres differ in several respects, making one anticipate variable patterns of distribution of grammatical elements. Essentially, these genres differ in the field, participants involved in the communication and in their communicative purpose.

In terms of the field, the areas of human endeavor covered are natural science, social science and administration. There is a sense to expect that these three registers would use language differently. In terms of participants, both natural science and social science are academic texts. These are typically texts from university textbooks and academic journals. Thus, communication across these genres is from an expert to colleague experts or to students.

Administrative writing, however, involves a professional expert communicating administrative instructions to lay people. The different orientations in participants across these genres are crucial parameters that have a potential of influencing linguistic choices.

Finally, the three genres differ in their communicative purposes. Thus, whereas the two academic genres are research/academic genres whose function is to educate and inform colleagues and students about research findings and new developments in the area, the one communicates social information and the other, hard science. Administrative writing, however, serves the communicative purpose of instructing its audience. These are the variable factors that we deem will influence distributional patterns of the complex sentences and their internal clause combination patterns across the genres.

8. Analytical methodology

This work straddles quantitative and qualitative methods. Register-based studies typically combine the two approaches. As Biber puts it

> …the two approaches have complementary strengths and weaknesses…Quantitative analysis gives a solid empirical foundation to the findings; non-quantitative analyses are required for the interpretation. Either type of analysis in isolation gives an incomplete description (Biber 1988: 52).

This has been achieved in this work by describing the situational characteristics of the genres involved in section 7, describing the targeted linguistic features in section 4 and interpreting patterns of distribution of the targeted linguistic features functionally by relating the situational features with the patterns of distribution of the linguistic features in section 9.1 and 9.2 (Biber & Conrad 2009).

9. Data analysis and discussion

Patterns of distribution of structural sentence types are presented below. The dominance of the complex sentence throughout the three genres across the two dialects studied motivated our exploration of the clause combining strategies prevalent across the three genres.

9.1 *Structural sentence types*

There is no disputing the fact that the complex sentence is the most preferred sentence pattern in written registers, as our data suggest. And as can be seen in Table 2, the dominance of the complex sentence cuts across both Ghanaian and British genres. This is part of the structural evidence that written texts are, by nature, complex and elaborated (Biber 1988). Our data reveal the patterns of distribution of structural sentence types captured in Table 2 below.

Table 2: Distribution of sentence type

	SOCGH	%	SOCGB	%	NATGH	%	NATGB	%	ADMGH	%	ADMGB	%
SIMP	53	41	26	28	32	28	35	34	18	20	31	31
CMPL	**55**	**42**	**52**	**57**	**59**	**53**	**46**	**45**	**63**	**72**	**52**	**51**
COMP	13	10	9	10	11	10	11	11	2	2	8	8
C CMPL	**9**	**7**	**5**	**5**	**10**	**9**	**10**	**10**	**5**	**6**	**10**	**10**
	130	**100**	**92**	**100**	**112**	**100**	**102**	**100**	**88**	**100**	**101**	**100**

The dominance of the complex sentence type is a good motivation for exploring the complex sentence in detail. However, looking at the patterns of distribution of the complex sentence across the three genres, one cannot but argue that registerial motivation more than anything else influenced the choices of this type of sentence. The first indication of this dominance is that the complex sentence is somehow suited to the situational context of the three genres investigated. However, the lack of regularity in the patterns of distribution of these grammatical features among genres across the two dialectal contexts explored makes it difficult to make bold generalizations.

For instance, when we combine attestations of compound complex sentences which have complex sentences within them with complex sentences, we observe irregular usage variations both cross-generically and cross-dialectally. As can be compared in Figure 1 below, usage variations for complex and compound complex sentences across the Ghanaian and British genres are between 49% and 62% for Academic Social Science, 62% and 55% for Academic Natural Science and 71% and 61% for Administrative Writing. The only consistency here is the fact that the complex and compound complex sentences are dominant across the three genres. Cross-dialectally, the observable variations are 49%, 62% and 77% for Ghanaian Academic Social Science, Academic Natural Science and Administrative Writing and 62%, 55% and 61% for British Academic Social Science, Academic Natural Science and Administrative Writing. These patterns, captured in Figure 1 below, show some similarity among the British genres, especially between Academic Social Science and Administrative Writing.

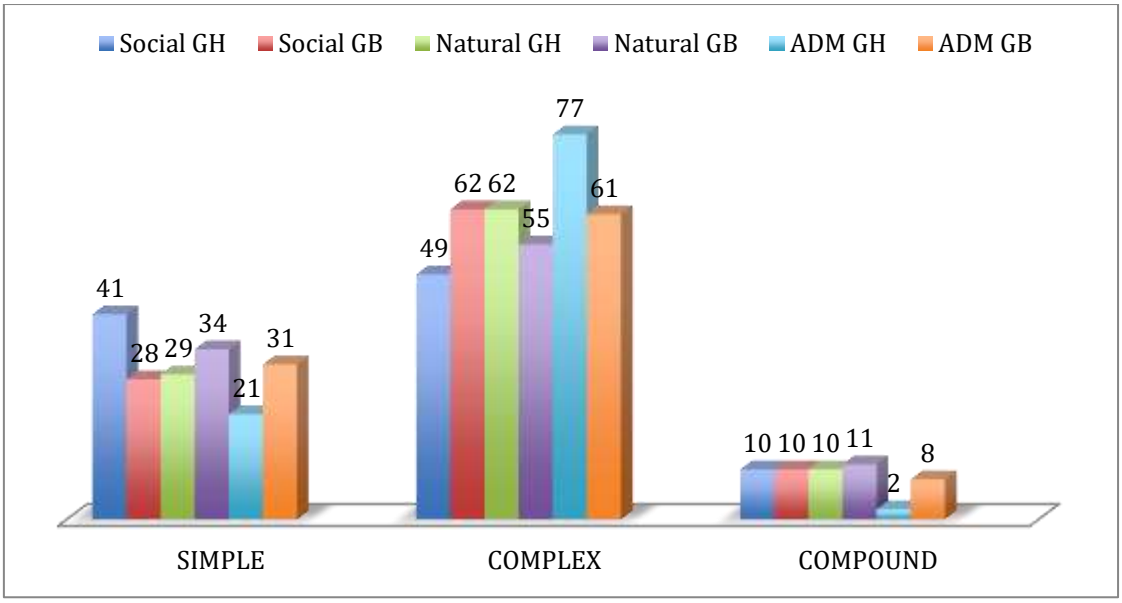


Figure 1: The dominance of the complex sentence types across the 3 genres

What these representations mean is that the complex sentence, such as the ones in sentence (8) to (11) are dominant across data from our two written corpora, though in variable patterns.

(8) Their primary role is likely to be in the removal of nutrients (nitrogen and phosphorus), although in the absence of light it is possible that they may behave heterotrophically and play a small part in BOD removal. ICE-GB: W2A-021 #2
(A COMPLEX SENTENCE WITH EMBEDDED AND ADJUNCTIVE DEPENDENT CLAUSES)

(9) From the above reasoning, we can consider the mole as the unit used by chemists to express large number of atoms, ions, molecules, and any identifiable species. ICE-GH:W2A-025 #25
(A COMPLEX SENTENCE WITH TWO EMBEDDED DEPENDENT CLAUSES)

(10) A parent or any other person who is legally liable to maintain a child or contribute towards the maintenance of the child is under a duty to supply the necessaries of health, life, education and reasonable shelter for the child. ICE-GH:W2D-004 #1
(A COMPLEX SENTENCE WITH THREE EMBEDDED DEPENDENT CLAUSES)

(11) To switch the metaphor: rather than producing a melting-pot in which differences disappear, the increased input has resulted in an everexpanding smorgasbord. ICE-GB:W2A-012 #14
(A COMPLEX SENTENCE WITH THREE DEPENDENT CLAUSES)

Apart from the fact that the complex sentence has been identified with planned discourse, our data indicate that the complex sentence is essential for the genres we investigated and so is dominant across these genres from the two dialectal perspectives. It is observed that the two academic genres have a tendency of integrating ideas from related extant literature to establish best applicable methodologies and theories. It is not surprising that the complex and compound

complex sentences should be dominant to help perform the communicative function of integration of ideas. Similarly, one observes that the instructional and directive nature of administrative writing equally involves integrating ideas. Thus, even though a close study of the types of complex sentence used across the administrative texts reveals some variability across Ghanaian and British texts, one observes at this stage that the complex sentence is the most preferred across Ghanaian and British administrative texts.

For the purpose of stylistic elegance and perhaps also for organization of ideas into logical argumentation, it makes sense that the simple and compound sentences are not the preferred types. This is because whereas a succession of simple sentences frustrates the flow of communication, especially when written, pairing of simple sentences into compound sentences limits the extent and dynamism of the logical argumentation one can make. For their communicative functions, whereas the simple sentences typically "engender frenetic or fast-paced feel" to a text, "the symmetrical nature of the connection between their units makes them [compound sentences] a favoured style in material designed for junior readers" (Simpson 2004: 60). Thus, the participants of the three texts explored (adults versus experts) the need of the authors to synthesize ideas and to argue out positions tend to attract complex sentence structures. As Halliday & Matthiessen (2014: 438) argue in their tactic relations of clauses, complex structures are issues of logical relations between ideas. Thus, the dominance of the complex sentence across the three genres from the native and nonnative Ghanaian and British contexts can be argued as a strategy that facilitates a variety of ideas to be logically expressed in sentence forms.

For example, the dynamism of the complex sentence in sentence 8 above resides in the complexity of the variety of clause combining relations within the sentence. One observes that in sentence (8) there are 4 clauses from different clause typologies which are combined using different combinatorial strategies. We observe, for instance, dependency at various levels. At the level of the clause we have *'[T]heir primary role is likely to be in the removal of nutrients* and *although in the absence of light it is possible that they may behave heterotrophically'* which exhibit a dependency relation between a main clause and a subordinate clause. A more detailed investigation reveals that even within these clauses are other clausal units such as the clausal subject (*that they may behave heterotrophically).* At the level of the phrase, we have an instance in which a phrase is being postmodified by a clause: 'to be in the removal of nutrients…' These are tactic and embedding or rankshifted relations – relations argued by Downing & Locke (2006: 275) as reflecting the degrees of the integration of events from the point of view of the speaker with embedding encoding a tighter integration.

Payne (2011:339) observes that complementation in particular "presents a major challenge for second language learners" because though all the structural types of dependent clauses can occur as complement clauses not all of them can complement any type of complement-taking predicator. One therefore expects variability to exist in the use of these structures across different texts from native and nonnative contexts. The curiosity therefore was to explore the complex sentence further to verify whether the data would exhibit regular patterns in the usage of these combinatorial strategies. That is, does the use of the two clause-combining strategies exhibit a distinction between native and nonnative language use or a registerial differentiation across the three genres from the two dialects of English?

9.2 *Clausal object/complements and adverbial clauses*

Our data reveal a mixture of realizations in the distribution of the two clause combining strategies across the three genres. We observed as recorded in Table 3 below that only the academic natural science genres across Ghanaian and British corpora had similar patterns of realization. That is, whereas circumstantial clauses were the dominant clause combining strategy for the natural science genres across the British and Ghanaian corpora, variable patterns of clause combining realizations were observed across Academic Social Science and Administrative Writing as captured in Table 3 below.

Table 3: Usage patterns of clause Adjuncts and clause complements

	SOCGH		SOSGB		NATGH		NATGB		ADMGH		ADMGB	
		%		%		%		%		%		%
COMPLEM	22	**44**	47	**70**	26	**35**	16	**27**	47	**66**	7	**9**
ADJUNCTS	28	**56**	20	**30**	48	**65**	44	**73**	24	**34**	69	**91**
	50	**100**	**67**	**100**	**74**	**100**	**60**	**100**	**71**	**100**	**76**	**100**

As observed in Table 3 above, it is only the natural science genre alone that reflects some form of registerial variability. It is observable that both Ghanaian and British natural science genres show consistency in their preference for adverbial clauses. Examples of the adverbial clauses that were dominant across Ghanaian and British natural science texts include the highlighted in sentences (12) and (13) below

(12) **To recrystallize an impure solid**, we select a solvent that will dissolve the solid **when the solvent is hot** but not **when the solvent is cold**. ICE-GH:W2A-021#88
(A COMPLEX SENTENCE WHICH HAS 3 CIRCUMSTANTIAL AND TEMPORAL ADVERBIAL CLAUSES)

(13) Isolated carbonate platforms occur in the Devonian of western Canada, **where reef development is more extensive on the windward side**. W2A-023#27
(A COMPLEX SENTENCE WHOSE DEPENDENT CLAUSE IS A LOCATIVE ADVERBIAL CLAUSE)

The results for the natural sciences support the register argument that similarities in the situational context induce similar linguistic choices even across different dialects of a language (Biber & Conrad 2009:12). The preference for adverbial clauses of Academic Natural Science from both the Ghanaian and British corpora seems to correspond with the situational context of the natural science genre. One principal communicative function of academic writing – and for that matter the Academic Natural Science - is not just to report on findings but also to integrate knowledge and it seems that these hypotactic adverbial clause structures are crucial in realizing these communicative goals.

Similarly, the consistency in the usage of clause combination patterns across the natural science texts is in support of another argumentation made by Biber and Gray (2010: 4) whose investigation of academic registers observes that the natural sciences are not as structurally elaborated as the humanities. This means, perhaps, that since the language of natural science is tilted towards structural compression rather than structural elaboration (Biber & Gray 2010:

4), complement positions are mostly filled by phrases. The problem that is till outstanding is why circumstantial notions are not the preferred type among the academic social science texts.

The academic social science and the administrative writing genres, as may be seen in Table 3, did not achieve any consistent patterns of realization. That is, whereas adverbial clauses were the preferred clause combination in the Ghanaian Academic Social Science, clause complementation is the preferred pattern in the British corpus. This though is not a strong evidence to enable one to argue that the Ghanaian users of the English language are faced with the clause complementation challenge observed by Payne (2011: 339); for, in the Administrative Writing genre, clause complementation is the dominant pattern in the Ghanaian data. What can be said now is that whereas clause combination patterns highlighted in sentence (14) are preferred in Ghanaian Academic Social Science, those highlighted in (16) and (17) are preferred in British Social Science.

(14) But **when Governor Guggisberg, one of the founders of Achimota, suggested that Achimota College should be linked with London University**, the Gold Coast Advisory Education Committee rejected the idea **because it would mean Achimota losing the opportunity to develop the indigenous system of higher education in Africa**. ICE-GH:W2A-005
(A COMPLEX SENTENCE WITH 5 DEPENDENT CLAUSES: 2 ADVERBIAL CLAUSES, 3 EMBEDDED CLAUSE)

(15) Consequently, it was suggested **that the problem of understanding codes which could not be related to jobs could be solved by having definitions for each code specific to each group**. ICE-GB:W2A-016
(A COMPLEX SENTENCE WITH 3 EMBEDDED CLAUSES)

(16) The dramatic increases in capacity planned by the Japanese producers have been paralleled by reinvestment and capacity increases by the US and European auto companies and it is estimated **that between 1985 and 1990 European production capacity will have increased by almost 20 per cent - from 11.2 million to 13.2 million units per annum.** ICE-GB:W2A-016
(A COMPOUND COMPLEX SENTENCE WITH 2 INDEPENDENT CLAUSES AND 2 EMBEDDED CLAUSES)

From dialectal perspective, these patterns of distribution have some implications. In the first place, one observes that whereas two of the British genres (Academic Natural Science and Administrative Writing) rely on hypotactic circumstantial clause combination, two of Ghanaian genres (Academic Social Science and Academic Natural Science) rely on clause complementation. What is more interesting is the disparity between the usage patterns of Ghanaian Administrative Writing and its British counterpart.

What one expected was that the participants involved in the administrative writing (expert addressors versus non-expert audience) would occasion a structural simplicity involving circumstantial relations – a phenomenon achieved somehow in the British Administrative Writing. The argument is that, as observed above, circumstantial relations are easier to generate and to comprehend than complementation (Payne 2011; Halliday & Matthiessen 2014). Thus, for the audience in the administrative writing who may not be experts to easily comprehend the message communicated, more circumstantial relations rather than

complementation patterns are required. The reality, however, is that there are more circumstantial relations (90.8%) in the British Administrative Writing but more complement clauses in the Ghanaian Administrative Writing (66.2%).

What perhaps accounts for the disparity in the distribution of circumstantial relations and clause complementation across Ghanaian and British administrative genres is that whereas texts in British Administrative Writing are mainly instructions on social service policies obviously intended for lay people, texts in the Ghanaian administrative genre are typically legal contracts between the social service provider and an expert who may intervene on behalf of a client. Excerpts in Table 4 below show the variation both in the kind of texts making up the administrative genres across these two native and nonnative contexts and in the variability of the prevalent clause combining strategies.

Table 4: Variability across Ghanaian and British Administrative genres

ADM WRITING (GB W2D-004 #15 – 17)	ADM WRITING (GH W2D-001#2 &9)
You must pay these *if you are normally self-employed, unless you have applied for and been granted exception because your earnings are below the exception limit for Class 2 contributions*. *If you are excepted* you may, *if you wish*, pay Class 2 contributions voluntarily *to keep up your right to the benefits* they provide. Class 2 contributions are paid at a flat rate – you pay the same amount *however much you earn*.	The President may, acting in accordance with the advice of the Council of State, by Proclamation published in the Gazette, declare *that a state of emergency exists in Ghana or in any part of Ghana* for the purposes of the provisions of this Constitution. For the avoidance of doubt, it is hereby declared *that the provisions of any enactment, other than an Act of Parliament, dealing with a state of emergency declared under clause (1) of this article shall apply only to that part of Ghana where the emergency exists.*

Thus, it is indicative, for example, across excerpts in Table 4 that instructions in the British administrative genre are made directly to the service consumer predominantly with the use of the second person pronoun. Besides, the dominant clause combining patterns in the British Administrative Writing are circumstantial adjunctive structures, which like all other forms of tactic relations, are relatively easy to produce (Eggins 2004: 269-270) and so by implication easy to comprehend. What it means is that, for the British administrative texts, after the main information has been provided in the main clause, which is usually not stuffed with complementation clauses, there are several adjunctive clauses to provide background information. As observable in Table 4 above, each of the main instructions in the British administrative register is backed by at least one circumstantial clause which in the excerpts in Table 4 are predominantly conditional adjuncts such as those in the examples below:

...if you are normally self-employed,
...unless you have applied for and been granted exception
...because your earnings are below the exception limit for Class 2 contributions
...If you are excepted

...if you wish
...however much you earn

For the Ghanaian administrative texts, however, language is used as though it is meant for 'whom it may concern'. As may be seen in the second column in Table 4 above, the message is not directed towards the consumer of the social service. It looks more like a legal document which is meant for an expert to interpret to the service consumer. Thus, typical of the complexity associated with legal documents, both main clauses and subordinate clauses are padded with clause complements and other embedded clauses. Although this phenomenon is not limited to the Ghanaian data, the conspicuous difference is that whereas a list of ideas typically gets expressed in the Ghanaian data through clausal complements which themselves embed other clauses as in sentence (17) below, in the British administrative texts, they are typically expressed in circumstantial structures as in sentence (18).

(17) The objectives of the policy are to:

- Promote widespread acceptance and observance of the Convention on the Rights of the Child;
- Promote the observance and enforcement of the Children's Act 1998, Act 560;
- Integrate ECCD issues into Development Planning at the community, district, regional and national levels;
- Provide pre-school education in collaboration with District Assemblies, the private Sector, Community Based Organisations (CBO), NGOs and Faith Based Organisations (FBOs);
- Provide fee-free tuition in pre-schools;
- Provide for the preparation (training) and upgrading of pre-school teachers and caregivers;
- Promote nutrition and household food security;
- Reduce the high Infant and Under-Five Mortality Rates;
- Provide information and skills to parents and other caregivers;
- Improve income earning capacity of parents;
- Develop an in-depth preparatory programme of action on ECCD and HIV/AIDS;
- Provide for the preparation and upgrading of preschool teachers and caregivers handling children with special needs. (GH:W2D-003 #34 - #55)

The communicative purpose of this single sentence with more than 10 complement clauses some of which also embed other clauses is to provide basic information. As such, it minimizes or avoids completely the use of circumstantial information. Such communication is too tight and would be more appreciated by experts or professionals.

The style of writing for the British Administrative genre, on the other hand, is a direct one, as already observed above. The addresser is directly addressing the addressee and the language somehow reflects a relationship between an expert addresser and a non-expert addressee. And to make the addressee understand the information communicated, the addresser uses a strategy which usually involves two things. That is, there is usually an instruction which is the main message such as '*You may pay these voluntarily*'. This instruction is usually given at least one background or circumstantial information such as 'to help you qualify for basic

Retirement Pension or widows' benefits', 'if you are not liable for contributions', etc. in sentence (18) below.

(18) You may pay these voluntarily **to help you qualify for basic**
MAIN MESSAGE CIRCUMSTANCIAL INFO 1

Retirement Pension or widows' benefits, if:
- you are not liable for contributions; or
  CIRCUMSTANTIAL INFO 2
- you have been excepted from Class 2 contributions; or
  CIRCUMSTANTIAL INFO 3
- your contribution record is not good enough to qualify for benefit.
  CIRCUMSTANTIAL INFO 4 (GB W2D-004 # 18)

10. Limitations to the study

This variable use of language across the two sociocultural contexts is interesting for a number of reasons. The most obvious implication relates to the compilation of the International Corpus of English. One observes that particular genres have wide coverage, having subgenres within them. The administrative genre, for instance, has within it direct instructions and legal administrative documents. It appears that whereas the Ghanaian administrative texts selected for this study were predominantly administrative legal texts, their British counterparts were predominantly instructional texts. This creates a disparity in linguistic choices which perhaps reflected across the Ghanaian and British administrative texts.

Besides, it makes sense to expect that grammatical elements should correspond to situational contexts, especially to communicative functions. However, because there is no scientific method for corresponding grammatical elements to situational contexts, register analysis is still a subjective enterprise.

11. Conclusion

This study has explored aspects of the sentence and its immediate internal constituents across three written genres from Ghanaian and British corpus data with the hope of establishing whether functional or dialectal motivation influences the distribution of these grammatical elements. The consistency of the dominance of the complex sentence across the three genres (Academic Natural Science and Academic Social Science and Administrative Writing) led us to argue that this sentence type has functional significance for the three genres. Concerning the two clause combining strategies explored in this study, our data again reveals a form of registerial consistency across the natural science genres from the Ghanaian and British corpora. It was observed that adjunctive clause combination is the preferred strategy for both Ghanaian and British academic natural science genres. The other two genres realized variable patterns of clause complementation and adjunction. Some dialectal patterns were equally detected. It was observed that two of the Ghanaian genres (Academic Social Science and Academic Natural Science) relied predominantly on circumstantial clause patterns. Similarly, two British genres

(Academic Natural Science and Administrative Writing) relied predominantly on circumstantial clause patterns.

For administrative writing, it was observed that whereas the Ghanaian administrative texts showed preference for clause complementation, the British administrative texts had dominance of circumstantial relations. This variability, we have argued, is a reflection of internal variation across the Ghanaian and British administrative genres. As our corpus data reveal, whereas the Ghanaian administrative texts are made up of legal social policies, the British counterparts are dominantly direct instructional texts. This variability is relevant to the situational context of the administrative genre. It reflects variability in participants of the discourse and communicative function.

References

Biber, Douglas. 1988. *Variation across speech and writing.* Cambridge: Cambridge University Press.

Biber, Douglas & Gray, Bethany. 2010. Challenging stereotypes about academic writing: complexity, elaboration, explicitness. *Journal of English for Academic Purposes* 9(1). 2-20.

Biber, Douglas & Conrad, Susan. 2009. *Register, Genre and Style.* New York: Cambridge University Press.

Biber, Douglas & Johanson, Stig & Leech, Geoffrey & Conrad, Susan & Finegan, Edward. 1999. *The Longman grammar of spoken and written English.* Cambridge: Cambridge University Press.

Bloor, Thomas & Bloor, Meriel. 2004. *The Functional Analysis of English.* London: Hodder Education.

Dako, Kari. 2003. *Ghanaianisms: A Glossary.* Accra: Ghana University Press.

Downing, Angela & Locke, Philip. 2006. *English Grammar: A University Guide* (2nd Edition ed.). London: Routledge.

Eggins, Suzanne. 2004. *An Introduction to Systemic Functional Linguistics* (2nd Edition). London: Continuum International Publishers.

Fawcett, Robin. 2000. *A Theory of Syntax for Systemic Functional Linguistics.* Amsterdam: John Benjamin.

Frimpong, Kodie G. 2015. *A Comparative Register Analysis of Editorials from Ghanaian and British Newspapers.* Accra: University of Ghana, Doctoral dissertation).

Frimpong, Kodie G. 2017. Subordination Across Ghanaian and British Newspaper Editorials: A Register Perspective. *Ghana Journal of Linguistics* 6(1). 72 - 115.

Halliday, M.A.K. & Matthiessen, Christian. 2004. *An Introduction to Functional Grammar* (3rd Edition.). London: Hodder Education.

Halliday, M.A.K. & Matthiessen, Christian. 2014. Halliday's Introduction to Functional Grammar (4th Edition). London: Routledge.

Hopper, Paul & Traugott, Elizabeth Closs. 2003. *Grammaticalization* (2nd Edition). Cambridge: Cambridge University Press.

Huber, Magnus. 2004. Ghanaian English phonology. Kortmann, Bern and Edgar W. Shneider (eds.). *A handbook of varieties of English. A multimedia Reference Tool.* Berlin: Mouton de Gruyter.

Huber, Magnus. 2014. Stylistic and sociolinguistic variation in Schneider's Nativization Phase: T-affrication and relativization in Ghanaian English. In S. Buschfeld, T. Hoffmann, & M. Huber (Ed.), *The Evolution of Englishes: The Dynamic Model and Beyond,* 86-106. Amsterdam: John Benjamin.

Kachru, Braj B. 2006. World Englishes and Culture Wars. In B. B. Kachru, Y. Kachru, & C. L. Nelson (Ed.), *The Handbook of World Englishes*, 446-471. Malden: Blackwell Publishing.

Koranteng, Louisa. 2006. Ghanaian English: A description of its sound system and phonological features. PhD Thesis, Accra, University of Ghana.

Payne, Thomas. 2011. Understanding English Grammar: A Linguistic Introduction. Cambridge: Cambridge University Press.

Quirk, Randolf & Greenbaum, Sydney & Leech, Geoffrey & Svartvik, Jan. 1985. *A comprehensive grammar of the English language.* Edinburgh: Longman Group Limited.

Schneider, Edgar W. 2007. *Postcolonial English: Variations around the World.* Cambridge: Cambridge University Press.

Schneider, Edgar W. 2003. The dynamics of New Englishes: From identity construction to dialect birth. Language 79(2). 233 - 281.

Sey, Kofi. 1973. *Ghanaian English: An exploratory survey.* London: Macmillam.

Simpson, Paul. 2004. *Stylistics: A Resource Book for Students*. London: Routledge.

Thompson, Geoff. 2014. Introducing Functional Grammar (3rd Edition). London: Routledge.

Vercellotti, M. L. & Packer, J. 2016. Shifting structural complexity: The production of clause types in speeches given by English for academic purposes students. *Journal of English for Academic Purposes* 22. 179 - 190.

Wardhaugh, Ronald. 1995. *Understanding English Grammar.* Oxford: Blacwell.

Wiredu, John Franklin. 2012. A Grammar of Newspaper Editorial language: the Complex Sentence. *Legon Journal of Humanities* 23. 75 - 96.

Wiredu, John Franklin. 2014. The Complex Sentence in Legal English: A Study of Law Reports. *UNILAG Journal of Humanities* 2(02). 1-27.

George Kodie Frimpong
Department of English
P. O. Box LG 129
University of Ghana
Legon, Accra, Ghana – West Africa
frimpong.kodie@gmail.com

In SKASE Journal of Theoretical Linguistics [online]. 2018, vol. 15, no. 3 [cit. 2018-13-12]. Available on web page http://www.skase.sk/Volumes/JTL38/pdf_doc/12.pdf. ISSN 1336-782X.

Rhetorical Structure of English and Czech Academic Book Reviews

Jana Kozubíková Šandová
University of South Bohemia in České Budějovice, Czech Republic

This study focuses on a contrastive analysis of linguistics book review articles written in English and Czech. The aim is two-fold: 1) to find out any variation in the rhetorical structure of these reviews; 2) to explore whether these structural differences somehow affect the way communicative goals of the genre of the book review article are achieved in the two language cultures. The theoretical framework of this study provides a modified version of Motta-Roth's (1995) taxonomy of rhetorical moves and their sub-functions in book review articles. The results show that in spite of a number of similarities between the two writing traditions, certain variation occurs attributable to different rhetorical preferences of both cultures and varied cultural expectations of the genre.

Keywords: *book review article, academic discourse, genre analysis, rhetorical structure, rhetorical move*

1. Introduction

This paper is an investigation about whether the rhetorical structure of the book review article written in English may be found in corresponding texts in different languages. Therefore, one aim of the present study is to examine any variation in the rhetorical structure of English and Czech academic book reviews. This paper also attempts to find out whether the possible differences in the internal structure somehow influence achieving communicative goals of the book review articles in the two different academic writing traditions.

Nowadays, academic review genres have taken on importance in applied linguistics literature. One of the reasons may be that the Internet abounds with various scholarly articles and scientific books so gaining access to them is much easier than ever before. As a result, it is very difficult for scholars to orient themselves in all these works and information they contain, and to distinguish between valuable contributions and those of a lower quality. Therefore, book reviews have gained significance in academia.

The genre of the book review article plays a crucial role in introducing new book titles to a specific discipline and in assessing their quality considering the latest development in the field. The book review article describes the structure and contents of the book, its purpose, the ease one can read the text, the clarity of tables or graphs, the quality or appropriateness of the corpus under study, etc. It highlights the most important parts of the book, evaluates it, and designates which given field of study the book belongs to. Therefore, such a review article belongs to discursive genres whose purpose is being descriptive, informative, and evaluative (Hyland 2000; De Carvalho 2001; Suárez & Moreno 2008, Dontcheva-Navrátilová 2018).

Even though the significance of the genre of the book review article has been clearly recognised by the academic community, not much is known about those attributes which make it a distinctive genre. Initially, genre has been a literary concept which became common when analysing non-literary discourse. In his seminal work concerning the study of genres, Swales (1990: 58) offers this definition:

> *A genre comprises a class of communicative events, the members of which share some set of communicative purposes. These purposes are recognised by the expert members of the parent*

> *discourse community, and thereby constitute the rationale for the genre. This rationale shapes the schematic structure of the discourse and influences and constrains choice of content and style.*

A similar claim has been made by Henry & Roseberry (1997: 480), who maintain that genre "groups texts together based primarily on their purpose within a social context."
Thus, the concept of genre is understood as a set of communicative events serving a specific communicative purpose and performed by various discourse communities.

In order to be able to classify a text within a particular genre, it must follow adequate conventions and have a particular structure. This has been confirmed by Swales (1990, 2004), who points out that one of the aspects relevant for a comprehensive description of a genre is the generic rhetorical structure of a text, which has been proved by current research into the academic book review (cf. Motta-Roth 1995; Gea Valor & del Saz Rubio 2000-2001; Suárez & Moreno 2008; Hyland & Diani 2009, i.a.). The segment a text is composed of is called a *move*. The division of texts into rhetorical moves has also been developed by Swales (1990, 2004), who employed this approach by the functional description of particular sections of research articles. Swales defines a move as a "discoursal or rhetorical unit that performs a coherent communicative function in a written or spoken discourse" (2004: 228-229).

In genre analysis, the concept of move is most frequently utilised to determine which regularly repeating structures appear within a particular genre. Depending on the general communicative function of the genre, moves may differ in length, stretching from one sentence to several paragraphs, but they usually consist of at least one proposition. As moves represent complete semantic and functional stretches of text determinable on the basis of their communicative function and linguistic boundaries, move analysis is very useful in genre-based approaches to discourse analysis (Connor & Mauranen 1999; Ding 2007). Thus, in a move analysis, rhetorical units are classified according to a specific communicative function each of these units perform with the aim of designating the overall communicative purposes of a text. These specific communicative functions then contribute to the general communicative intention of the whole genre. The special arrangement of the moves of a particular genre determines its specificity, which makes it distinct from other genres.

As regards the number of moves of a particular genre, it is not fixed since there is no mutual relationship between the manner by which rhetorical units of a genre are organised and its formal arrangement (Parodi 2010). In this connection it must be emphasised that not all of the moves of a particular genre are always present in a text. Some moves are *optional*, i.e. discourse participants may choose to utilise these moves to make the communication more effective but they do not influence the function of the text. On the contrary, some moves are *obligatory* because they are vital for achieving the communicative function of the genre. All these features are examined within genre analysis, which started to be developed after the most significant work of Swales (1990) concerning this type of analysis was published.

Following Swales, extensive research into various academic genres has been carried out, for example, into research articles (Dontcheva-Navrátilová 2016; Hyland 2000; Holmes 1997; Nwogu 1997, i.a.), grant proposals (Connor & Mauranen 1999), company audit reports (Flowerdew & Wan 2010), or essay conclusions (Henry & Roseberry 1997). The genre of the book review has not received as much attention. Therefore, Motta-Roth's (1995, 1998) examination of the rhetorical macrostructure of English book reviews in the disciplines of chemistry, economics, and linguistics may be regarded as a pioneering study in the field of move analysis stemming from the Swalesian tradition. The outcomes of Motta-Roth's study relating to the overall rhetorical structure of academic book reviews have been corroborated by work of other scholars, for instance, by Suárez & Moreno (2008), De Carvalho (2001), Gea Valor (2000), or Nicolaisen (2002). None of these studies reveal

any major differences in the basic organisation of book reviews, either cross-linguistically and cross-disciplinarily. This supports the claim that the book review may be considered a distinctive genre. However, as Suárez & Moreno (2008) correctly point out, subtle cross-linguistic and cross-disciplinary variations have been found out so in order to perform as accurate a description of the rhetorical structure of the book review as possible, it is necessary to reflect on these differences.

2. Material and method

For the purpose of this study, a corpus of 40 linguistics book review articles was compiled. The Anglophone sub-corpus contains 20 book reviews excerpted from distinguished linguistic journals (*Journal of Pragmatics, Journal of English for Academic Purposes, Journal of Linguistics, Applied Linguistics*, and *International Journal of English Studies*) written in English by Anglophone scholars affiliated with British or American universities. The Czech sub-corpus comprises likewise 20 book review articles, published in two distinguished peer-reviewed Czech linguistic journals (*Časopis pro moderní filologii, Slovo a slovesnost*). The authors of Czech reviews are academics of Czech origin employed by Czech universities. Even though there is a certain difference in the Anglophone and Czech journals used in this study as for the prestige and size of the target audience, the book reviews drawn from them form a representative sample for the purpose of this research.

All reviews were published between 2015 and 2018. The extent of the whole corpus is 80,237 words, the Anglophone sub-corpus containing 40,176 words, the Czech sub-corpus reaching the amount of 40,061 words. Both sub-corpora are almost identical in size, hence, they can be mutually compared.

Rhetorical move analysis was adopted to identify the individual moves of which each book review article is composed. The theoretical framework utilised in this study is a modified version of Motta-Roth's (1995) classification of rhetorical moves and their sub-functions in book reviews.

Now, let us describe the taxonomy of the book review rhetorical structure as proposed by Motta-Roth (1995) in greater detail. It consists of four basic moves, each of which is divided into several sub-functions, as apparent from Figure 1. Moves 1 and 2 belong to descriptive moves, whereas Moves 3 and 4 are evaluative. For the purpose of the present analysis, the sub-functions of Move 4 were modified in order to better differentiate between the categories of *evaluation* and *recommendation*, the latter category thus constituting a separate move.

Move 1 Introducing the book
Sub-function 1 Defining the general topic of the book and / or
Sub-function 2 Informing about potential readership and / or
Sub-function 3 Informing about the author and / or
Sub-function 4 Making topic generalizations and / or
Sub-function 5 Inserting book in the field

Move 2 Outlining the book
Sub-function 6 Providing general view of the organization of the book and / or
Sub-function 7 Stating the topic of each chapter and / or
Sub-function 8 Citing extra-textual material

Move 3 Highlighting parts of the book
Sub-function 9 Providing focused evaluation

Move 4 Providing closing evaluation of the book
Sub-function 10A Definitely recommending / disqualifying the book or
Sub-function 10B Recommending the book despite indicated shortcomings

Figure 1: Schematic description of rhetorical sub-functions in book reviews (Motta-Roth 1995: 142)

In Move 1 (*Introducing the book*), with the help of sub-function 1 (*Defining the general topic of the book*) the reviewer describes the subject matter of the book briefly. It is followed by sub-function 2 (*Informing about potential readership*) which mentions the possible target audience of the book. This sub-function is sometimes left out as part of Move 1 and the information about target readers is provided at the end of the review in the closing evaluation of the book. Sub-function 3 (*Informing about the author*) supplies information about the author of the reviewed book, such as their previous works or academic career. In sub-function 4 (*Making topic generalisations*), reviewers relate the book to the known facts and theories in the field. Sub-function 5 (*Inserting book in the field*) is more specific than the previous sub-function and discusses previously published books on the same topic or mentions the gap that the book under review could have covered.

The other descriptive move, Move 2 (*Outlining the book*), focuses on the description of the structure of the book. It may contain evaluative comments but they are quite rare since the main intention of this move is to outline the book. Sub-function 6 (*Providing general view of the organisation of the book*) explains the general structure of the book, i.e. the number of parts or chapters. Sub-function 7 (*Stating the topic of each chapter*) is more specific and describes the contents of each chapter. In some cases the reviewer is very specific, while in other cases the content is summarised in just one sentence. Sub-function 8 (*Citing extra-textual material*) mentions additional material not being a direct part of the book, such as illustrations, tables, figures, appendices, etc.

Move 3 (*Highlighting parts of the book*), together with Move 4, belongs to the evaluative moves of the book review. It has only one sub-function (sub-function 9 *Providing focused evaluation*) in which positive and negative aspects of the specific parts or chapters of the book are discussed. This move is sometimes connected with Move 2, in case the reviewer describes the particular chapters and evaluates them at the same time.

Move 4 (*Providing closing evaluation of the book*) provides a final summary of the book review and simultaneously expresses an unequivocal opinion of the general significance of the book under review. Compared to the evaluation provided in Move 3, this assessment is not that specific and relates to the book as a whole. Motta-Roth (1995) identifies two sub-functions within this move: Sub-function 10A (*Definitely recommending / disqualifying the book*) and sub-function 10B (*Recommending the book despite indicated shortcomings*). At this point, we have modified Motta-Roth's classification since we consider *evaluation* and *recommendation* to be two different rhetorical functions, which cannot be subsumed under one joint category. When evaluating the book, the reviewer expresses their subjective viewpoint of the book and assesses it in either positive or negative terms. When recommending the book, the reviewer encourages readers to act somehow, e.g. buy the book, read the book, borrow the book, etc. Therefore, Move 4 consists of sub-function 10 *Completely positive evaluation* and sub-function 11 *Positive evaluations with aspects to improve*. In case of an overall negative evaluation, sub-function 12 *Completely negative evaluation* could be added. However, it was not found in the analysed corpora.

Move 5 (*Recommendation*) comprises three sub-functions: *Definitely recommending the book* (sub-function 13), *Recommending the book despite indicated shortcomings* (sub-function 14), and *Not recommending the book* (sub-function 15). Sub-function 13 contains a straight-forward recommendation of the book being reviewed. In sub-function 14 the reviewer recommends the book, although, certain negative aspects are also highlighted. Finally, sub-function 15 consists in an outright

rejection of the book under review. The modified version of Motta-Roth's classification described in the preceding paragraphs is illustrated below (Figure 2).

Move 1 Introducing the book
Sub-function 1 Defining the general topic of the book and / or
Sub-function 2 Informing about potential readership and / or
Sub-function 3 Informing about the author and / or
Sub-function 4 Making topic generalizations and / or
Sub-function 5 Inserting book in the field

Move 2 Outlining the book
Sub-function 6 Providing general view of the organization of the book and / or
Sub-function 7 Stating the topic of each chapter and / or
Sub-function 8 Citing extra-textual material

Move 3 Highlighting parts of the book
Sub-function 9 Providing focused evaluation

Move 4 Closing evaluation
Sub-function 10 Completely positive evaluation or
Sub-function 11 Positive evaluation with aspects to improve or
Sub-function 12 Completely negative evaluation

Move 5 Recommendation
Sub-function 13 Definitely recommending the book or
Sub-function 14 Recommending the book despite indicated shortcomings

Figure 2: A modified version of Motta-Roth's taxonomy of rhetorical sub-functions in book reviews

3. Results and discussion

One aim of this paper was to explore any variation in the rhetorical structure of Anglophone and Czech linguistics book review articles. The possible differences in the use of moves and their sub-functions are indicative of distinctions connected with the specific academic culture. By rhetorical structure we mean the way a text is organised into a coherent whole in a given genre. A text, of course, does not represent a mere series of clauses but it is composed of hierarchically organised constituents being in diverse relations to each other.

According to the above-described classification, the basic rhetorical structure of book review articles consists of five moves and their sub-functions. Their occurrence in both corpora is summarised in Table 1 below.

Table 1: The absolute frequency of rhetorical moves and sub-functions in English and Czech BRs

	M1, SF1	M1, SF2	M1, SF3	M1, SF4	M1, SF5	M2, SF6	M2, SF7	M2, SF8	M3, SF9	M4, SF10	M4, SF11	M4, SF 12	M5, SF13	M5, SF14	M5, SF15
English BRs (20)	20	4	0	9	7	16	19	3	17	8	12	0	2	2	0
Czech BRs (20)	20	6	10	9	15	16	17	11	14	15	5	0	1	2	0

The figures in Table 1 indicate the absolute frequencies of the moves and sub-functions. As is apparent, some cross-linguistic differences occur. Perhaps one of the most striking variations is the absence of Move 1, sub-function 3 (*Informing about the author*) from the Anglophone sub-corpus, while in the Czech sub-corpus it occurs in 10 out of 20 reviews. It is quite common for Czech reviewers to introduce the author of the book briefly. They usually provide basic information about their affiliation and/or recent publications, often on a similar topic as the reviewed book. This background information may contribute positively to the author-reader relationship, which becomes more interpersonal and, consequently, it may raise the interest of the reader to buy or read the book under review or other publications written by the same author. Furthermore, the reader may learn about the author's expertise. Last but not least, the reviewer may be regarded as an expert who knows other authorities in the field. Anglophone reviewers do not consider such kind of information important or necessary as it may divert attention from the merits of the book under review. Below are two examples excerpted from the Czech sub-corpus.

(1) *Docentka Martina Šmejkalová, vedoucí katedry českého jazyka na Pedagogické fakultě Univerzity Karlovy*, [Associate Professor Martina Šmejkalová, Head of Czech Department at the Faculty of Education, Charles University] *je odborné veřejnosti známa jako autorka rozsáhlé monografie* Čeština a škola […]. [CMF6]

(2) *Autorkou publikace je brněnská portugalistka Iva Svobodová* [The author of the publication is the Brno-based expert in Portuguese linguistics Iva Svobodová], *jež se specializuje zejména na oblast morfologie současného portugalského jazyka (problematika členu).* [CMF2]

Another difference between the two corpora lies in the occurrence of Move 2, sub-function 8 (*Citing extra-textual material*). It appears only three times in the Anglophone sub-corpus, whereas in the Czech sub-corpus it was found in 11 reviews. This sub-function provides the readers with information about additional parts of the book, such as illustrations, appendices, figures, tables, or references. Anglophone reviewers focus more on providing readers with the contents of the book and a detailed evaluation of particular chapters. Czech reviewers give information about the extra-textual material in about half of the reviews. However, they do not add any further specific details. Reviewers sometimes emphasise the additional character of these parts by using expressions such as *at the end of each chapter*, as in Example 3, or *only at the end*, *finally*, etc.

(3) *Finally, there is an appendix and index. References are included at the end of each chapter and not as a list in the end-matter of the book.* [IJES2]

(4) *Rozsáhlá monografie, doprovázená bohatou bibliografií* [extensive bibliography] *a obsahující množství podrobných tabulek, soupisů a názorných grafů,* [comprehensive tables, lists, and illustrative graphs] *podává v rámci zkoumaných korpusových dat* […] *vyčerpávající a detailní analýzu* […]. [CMF6]

(5) *Následují závěr, bohatá sekce referencí, řada příloh, rejstřík a anglické summary*. [What follows are conclusions, extensive references, a number of appendices, index, and an English summary.][SS4]

A further variation between the Anglophone and Czech corpora, though not particularly wide, concerns Move 1, sub-function 5 (*Inserting book in the field*). This sub-function contextualises the reviewed book in the field by referring to books published previously on an identical topic or by stressing the importance of the new book which should fill a gap in the existing literature. Even though this sub-function may build the relationship and common ground between the reviewer and the audience, its occurrence in the Anglophone corpus is, with 7 occurrences, relatively low, compared to the Czech corpus where it appears in 15 out of the total of 20 reviews. In all the examples below, the reviewer places the book under review within a certain theoretical framework without making any generalisations about the topic which are the subject of sub-function 4.

(6) […], *and it brings together a collection of papers which explore topics and themes from across a range of work within the relevance-theoretic pragmatic framework*. [JP3]

(7) Faces of English Education *sets out to provide coverage of some of the key issues in current debates on English language education throughout the world* […]. [JEAP1]

(8) *Autorky se hlásí k odkazu Noama Chomského, především k jeho konceptu univerzální gramatiky z 60. let minulého století*. […the legacy of Noam Chomsky, especially his concept of Universal Grammar][SS2]

(9) *Jednotlivé stati prolíná problematika funkční perspektivy větné (FSP), pojímaná ve smyslu teorie Jana Firbase* [the theory of the Functional Sentence Perspective, dealt within Jan Firbas's theory][…] *a jeho žáků a pokračovatelů* […]. [CMF3]

Focusing on the rest of the moves and their sub-functions, there are no major differences between the two writing cultures. Sub-functions 2, 7, 9, 10, and 11 occur with a very similar frequency in both corpora, the occurrence of sub-functions 1, 4, and 6 is even identical.

All reviewers find it necessary to briefly describe the general topic of the book and its contents (Move 1, sub-function 1), see Examples 10-14 below. Apart from stating the general topic, they sometimes mention the theoretical framework employed by the author of the book (Example 10). One means of drawing the reader's attention to the book is employing the nominal phrase *this book/volume/monograph,* followed by a verb in the present tense and an object (Examples 10 and 11). Another means is stating the full title of the book in italics (Example 12) or the author's name (Examples 13 and 14).

(10) *This book re-examines this notion in the light of recent advances in Minimalism and compares the framework to alternatives that do not embrace the notion*. [JL3]

(11) *This book provides an important account of the issues surrounding institutional English language proficiency requirements and the types of institutional support needed to support the increasing international student cohort.* [JEAP1]

(12) *Publikace* Právní překlad v teorii a praxi, *jak napovídá název, je syntézou teorie překládání českých právních norem a praktických pokynů a návodů vycházejících z případového rozboru nového občanského zákoníku (NOZ).* [The publication *Legal translation in theory and practice* is a synthesis of the theory of translation of Czech legal norms and practical instructions based on a case analysis of the new Civil Code.][CMF1]

(13) *Kniha Petra Pořízky* [Petr Pořízek's book] *seznamuje čtenáře s problematikou vytváření jazykových korpusů a vyhledávání v nich.* [SS4]

(14) *Monografie Naděždy Kudrnáčové* [Naděžda Kudrnáčová's monograph] *je jedním z výsledků mnohaletého autorčina zájmu o sémantiku anglických pohybových sloves.* [CMF4]

Contrary to Move 1, sub-function 1, which is not missing from any review in both sub-corpora, only a minority of the reviewers consider it important to inform about the target readers of the book (Move 1, sub-function 2) at the beginning of the review (Examples 15-18 below). This information, if present, is more frequently provided at the very end of the review together with closing evaluation of the whole book. The potential readership is referred to explicitly, e.g. *postgraduate students, applied linguists, research community, academic support staff,* etc.; however, an implicit reference, such as *it offers an introductory guide…, it provides a basic insight…*, may occur, which means that the book under review is meant for undergraduate students or non-specialist readers. In case of a direct reference, moves belonging to this category contain phrases related to the interest of the target audience, such as *be of (great) interest to, be useful for, be of great value to, be helpful, providing rich insights that would benefit researchers from many disciplines*, etc. Sometimes there is reference to background knowledge or expertise (*introductory, general, essential*) or reference to education (*undergraduate, more advanced, graduate* (students).

(15) *As such, this book is essential reading for a range of people in the university system, from policy-makers and administrators, through to lecturers and academic support staff.* [JEAP1]

(16) […], *it is of great interest to the wider research community to learn more about this corpus resource for Turkish, and gain insights into the politeness (and impoliteness) principles of a non-Indo-European language*. [JL2]

(17) *Z tohoto důvodu je text určen především pro studenty oboru Portugalský jazyk a literatura uvedené univerzity.* [for students of Portuguese language and literature][CMF2]

(18) *Kniha je určena především vysokoškolským studentům portugalštiny* [for university students of Portuguese], *užitečné informace v ní však nepochybně naleznou i další romanisté a lingvisté* [other Romanists and linguists]. [CMF6]

Almost half of Anglophone and Czech reviewers makes specific topic generalisations (Move 1, sub-function 4). In this move, they provide further details concerning theories or facts presented in the reviewed book. They may give definitions of important terms or concepts, sometimes with examples

(Example 19). Also, important schools of thought or authorities are cited, as well as experiments or discoveries (Example 20).

(19) *In attributive usage, he distinguishes three patterns of modification: intersective attribution (such as* beautiful picture, *which denotes the intersection of the set of pictures and the set of beautiful things), intensional attribution (such as* false friend, *where the attribute augments "the semantics of head with the feature "not actual", and subsective attribution (as in* beautiful dancer, *in the sense "a person who dances beautifully*"). [JL3]

(20) *Autorky kromě výkladu vycházejícího z chomskyánského teoretického rámce* [Chomskyan theoretical framework] *popisují některé provedené psycholingvistické experimenty* [psycholinguistic experiments][…]. [SS2]

Another move with identical distribution across both corpora is Move 2, sub-function 6. With an incidence of 16 out of 20 reviews, it belongs to the most frequent ones. In this sub-function, reviewers pay attention to providing a general overview of the organisation of the book. They do it by explicitly stating how many sections or chapters the book contains, or alternatively the topics addressed in these parts. Some typical verbs expressing this sub-function are *divide*, *subdivide*, or *organise* used in the passive voice preceded by the noun *book* or *volume.* Of course the specific numeral indicating the number of chapters or parts is stated. Below are several examples of this rhetorical sub-function:

(21) *The volume is divided into four parts*. [AL1]

(22) *Recenzovaná kniha je rozdělena do osmi kapitol*. [The book under review is divided into eight chapters.][SS2]

What appears quite frequently in both corpora is the co-occurrence of sub-functions 6 and 7. Firstly, the reviewer describes the general organisation of the book while then continuing with a more specific summary of the key ideas of the particular chapters (Example 23). In some reviews, the chapters are described one after another (Examples 25 and 26), or sometimes, depending on the length of the reviewed book, they are grouped into sections and discussed together (Example 24).

As regards linguistic features employed by reviewers to express sub-function 6, they most frequently refer to individual sections of the book as *chapters*, combining this noun with a specific numeral (cardinal or ordinal number) or adjectives such as *following*, *next*, or *final*. All these means serve for a better orientation in the review itself as well as in the reviewed book since they indicate how the particular parts are arranged in a logical order. The most frequent verbs used in this sub-function are *feature, present, define, illustrate, discuss, show*, etc.

(23) *Their volume consists of seven chapters* […]. *Chapter 1 sets the tone by offering a general overview and a justification for their work* […]. *The second chapter features an excellent summary of various socioeconomic and institutional confluences* […]. *Chapter 3 features an impressively integrated history of EAP* […]. [JEAP1]

(24) *The following four chapters deal with discourse markers, but exploit different theoretical frameworks*. [JL2]

(25) *Kapitola Komunikační funkce věty a modalita obsahuje pojednání o způsobu oslovování a pozdravech, které většina dostupných gramatik v podstatě ignoruje*. [The chapter

"Communicative function of a sentence and modality includes a treatise on the way of address and greetings.][SS3]

(26) *Autoři v osmé kapitole nabízejí precizní výklad jednotlivých jevů* […]. [In Chapter 8 the authors offer an accurate explanation of the particular phenomena.][CMF4]

As is apparent from the figures in Table 1, Move 3 occurs very frequently in both corpora as well. In this move, which is realised only through sub-function 9 (*Providing focused evaluation*), the reviewers discuss positive and negative remarks of particular parts, usually chapters, of the book, which means that it is not a descriptive move, as Move 2, but rather evaluative. Judging from the occurrence of evaluative critical comments, Move 3 is the most evaluative move in both corpora.

To express their subjective attitude, both Anglophone and Czech reviewers employ lexical phrases such *in my view, z mého pohledu* [from my point of view], *in my/our opinion* (Example 29), or *in my own experience.* Assessing information is further conveyed explicitly by evaluative adjectives, nouns, or verbs, either positive or negative, e.g. *excellent, interesting, useful, weak, limited, convincing, value, strength, problem, shortcoming, succeed, overlook, lack,* etc. Also, superlative expressions such as *the best example, one of the greatest parts, the major strength* occur in this move.

(27) *Having set the standard for conceptualising the linguistic and literacy needs of the university student, Murray provides an excellent critique of pre-enrolment language testing in the fourth chapter*. [JEAP1]

(28) *The papers in this third volume, therefore, constitute a definite strength in the collection*. [JL2]

(29) *Velmi přínosná je podle našeho názoru kaptiola pojednávající o Šmilauerově činnosti lexikografické*. [Very beneficial is, in our view, the chapter dealing with…] [CMF6]

(30) *Podkapitola o členu působí kvůli svému rozsahu (celkově pouze tři strany) poněkud odbytě* [rather sloppy][…].[SS3]

(31) *Zařazení této kapitoly spíše lexikologické povahy se nám jeví jako velmi vhodné* [very suitable][…]. [CMF2]

The reviewers may also employ various positive or negative attitude markers, such as *nicely, clearly, impressively, convincingly* or *unfortunately*:

(31) *In an extensive overview* […], *the author Kathleen Currie Hall convincingly proposes a probabilistic metric of phonological relationships* […]. [JL2]

(32) *Unfortunately these are not always accompanied by substantial suggested answers* […]. [JL2]

(33) *Na s. 71 jde bohužel* [unfortunately] *o překlad zavádějící* […]. [SS4]
(34) *Nesmírně zajímavá* [particularly interesting] *je podkapitola pojednávající o Šmilauerově "slovenském období"* […]. [CMF6]

Another lexical means of expressing evaluation is validity markers, such as the modal verbs *would* or *should*. When employing them, the reviewer attenuates the force of their judgments to be not so

emphatic and to leave some room for the reader to decide about the validity of the reviewer's evaluation.

(35) *If there is to be a subsequent edition, a concluding chapter would prove helpful in this regard*. [AL2]

(36) *Rather, percentages based on small raw number can be misleading, and should be treated cautiously*. [JP2]

(37) *Domnívám se však, že by bylo ze strany překladatele vhodné uvést* [it would be appropriate to mention][…]. [SS3]

In conclusion, Move 3 stresses the positive and negative points occurring in the reviewed books, and reasons are provided why the reviewer assessed something positively or negatively, offering citations or data from the book. As is apparent from Table 1, Move 3 is not present in all the reviews, even though it provides evaluation, the key characteristic of this academic genre. The reason is that Move 4 also contains evaluation, which is present in this move in every review. The difference between these two moves is that Move 4 provides a closing evaluation of the book and summarises its most important aspects.

Thus, in Move 4, reviewers express their overall, either positive (sub-function 10 *Completely positive evaluation* and sub-function 11 *Positive evaluation with aspects to improve*) or negative (sub-function 12 *Completely negative evaluation*) opinion of the reviewed book. Move 4 is present in every review because apart from an overall assessment, it explicitly closes the review, especially in case Move 5 is absent. The evaluation may be completely positive, positive with aspects to improve or negative. Sometimes a combination of positive and negative features occurs. Lexical phrases indicating closing the review are, for example, *overall*, *altogether* or *in conclusion/summary*. Examples 38 and 39 illustrate sub-function 10, while Examples 40 and 41 demonstrate sub-function 11.

(38) *Overall, the volume forms a much needed middle ground between introductory textbooks and research papers and is suitable for a variety of students* […]. [IJES1]

(39) *V našem případě jde o studii velmi podařenou a inovátorskou, která mj. ukazuje, že lze srovánvat i více než dva jazyky* […][In our case it is an outstanding and innovative study, which shows that it is possible to compare more than two languages]. [CMF6]

(40) *Also, given the nature of the publication, it would have been very interesting to include Kachru's early work* […]. [JL2]

(41) *Podobných zdánlivě dílčích významových a formulačních nesrovnalostí je v knize celá řada. Pro případné druhé vydání knihy by proto bylo žádoucí provést revizi překladu*. [The book contains a number of similar, seemingly minor, inconsistencies in meaning and formulations.] [SS2]

The reviewer sometimes uses lexical phrases typically associated with academic life, such as *courses, libraries, shelves,* etc. (Example 42). Future applications of the reviewed book are suggested by employing the modal auxiliary *will* (Example 43).

(42) Faces of English Education *is a very useful addition to the bookshelves of those involved in TESOL programmes, especially at Masters level*. [JEAP1]

(43) […] *this is undoubtedly a very useful collection of many of Kachru's key papers, which will prove invaluable to both researchers and students of sociolinguistics and World Englishes*. [JL2]

Move 5 (*Recommendation*) does not appear much frequently in both corpora. An explicit recommendation occurs only in three book reviews altogether: in two English reviews and in one Czech review (Examples 44-46). The majority of book reviews in both corpora contain an implicit recommendation, which was included in Move 4 since the reviewer stresses the positive points but does not recommend the book explicitly.

(44) *So I recommend it wholeheartedly to the readership of this journal*. [AL2]

(45) […] *I recommend this as an addition to your personal shelves or university library if you are interested in any of those areas*. [IJES2]

(46) *Práce zaplňuje prostor synchronní lingvistické komparatistiky velmi účinně a lingvisticky zajímavě. Je bezesporu přínosná v mnoha ohledech* […]. *Lze ji proto vřele doporučit* [We may thus heartily recommend it] *každému vážnému zájemci o tuto jazykovou oblast a problematiku.* [CMF6]

What must be emphasised at this point is that the structure of some book review articles is not that homogeneous as one might assume from the description in this section. The so-called *cycling of the moves* occurs in both corpora and is related to Moves 2 and 3. This feature appears in 5 out of 20 Czech reviews and in 6 reviews written in English. In these reviews, Move 3 precedes a sub-function of Move 2, as, for instance, in Examples 47 and 48 below. Here, Move 3 is combined with sub-function 7. A particular part of the book is evaluated and outlined at the same time. This sequence is repeated several times throughout the review.

(47) *Chapter 6 presents perhaps the most positive appraisal of segments in the first section of this volume, eschewing the distinction between categorical segments and gradient representations in favour of new probabilistic models* […]. [JL2]

(48) *Sedmá kapitola (Nejstarší portugalsky psané texty) poskytuje čtenáři z mého pohledu velmi zajímavé informace o okolnostech hledání nejstaršího portugalského textu* […] [Chapter 7 (The oldest Portuguese-written texts) provides the reader, from my point of view, with very interesting information on the circumstances of searching for the oldest Portuguese text.]. [CMF6]

4. Conclusion

This paper has aimed to investigate variation in the rhetorical structure of English and Czech book review articles and to find out whether this variation in internal structure somehow influences achieving communicative goals of the book review articles in two different writing cultures.

The detailed analysis has revealed the existence of several recurring patterns present in both sub-corpora. For instance, it has shown that a group of texts of the same genre is composed of

prototypical moves, which usually appear in the same order. Opening and closing moves show a tendency to be obligatory, while central moves are more varied and optional. This is apparent, for instance, in sub-function 3 (*Informing about the author*), which does not occur in the Anglophone sub-corpus at all, whereas in the Czech sub-corpus it occurs in half of the reviews. Hence, basic features are combined in various ways in particular reviews and some moves may be missing but these texts are still considered to be book review articles because they contain obligatory moves and sub-functions designating the genre. Regarding linguistic realisation, the moves are expressed by the same or very similar language means in both language cultures. Furthermore, the sub-functions of the particular moves in English and Czech reviews correspond to each other. These findings confirm the understanding of the book review article as a genre in its own right within the two academic writing traditions. Thus, this study may contribute to a more accurate definition of this genre.

Nevertheless, the contrastive analysis of both sub-corpora has indicated a certain variation in the distribution of sub-functions across the moves. A higher incidence of sub-function 3 referred to in the previous paragraph suggests a positive contribution both to the author-reader and reviewer-reader relationship. Furthermore, Czech reviewers pay more attention to the extra-textual material in the book under review than their Anglophone counterparts. Last but not least, Czech reviewers contextualise the book within the particular field to relate it to the books on the same topic. All these three sub-functions appear much more frequently in the Czech sub-corpus, which supports the tendency of Czech reviewers to utilise descriptive moves far more frequently than evaluative ones. Besides, evaluation in Czech reviews is incorporated in the whole body of the review to a higher extent than in English reviews, where evaluation is usually found in the respective evaluative moves.

The ratio of descriptive versus evaluative moves in the Anglophone sub-corpus is more balanced. In addition, Czech reviewers are more sympathetic with the authors of the reviewed books in their overall evaluation than Anglophone reviewers. 15 out of 20 books are evaluated positively without any objections, whereas in the Anglophone sub-corpus only 8 books are rated completely positively, in 12 reviews the reviewers recommend some aspects to be improved. When Czech reviewers evaluate the book positively in all respects, such evaluation may be found in the introduction of the review as well, rather than only at the end. This is not a common practice in English reviews.

Another feature found exclusively in Czech reviews is the fact that the reviewer sometimes starts the review with information about the author of the book (sub-function 3) and after that they return to defining the general topic of the book (sub-function 1). In the Anglophone culture, it is more common to address the content of the text rather than place the scholar's background in a positive or negative light so that the text is judged on its own merits.

In conclusion, we can say that the differences mentioned in the previous paragraph reflect preferences of the two writing cultures concerning rhetorical structure of the genre of the book review article, which may also be connected with different cultural expectations and with a different way of achieving communicative purposes of this genre.

Many inquiries remain for further discussion. It would be useful to carry out a more detailed qualitative and quantitative research to explain possible sources of cross-cultural variation by focusing more on the cultural characteristics of academic writing traditions. It would be no less interesting to examine other disciplines besides linguistic book reviews and compare review articles assessing books written in the field of hard sciences with those of soft sciences.

Sources

Applied Linguistics (2015), 36 (1). [AL1]
Applied Linguistics (2016), 37 (1). [AL2]
Časopis pro moderní filologii (2015), 97 (1). [CMF1]
Časopis pro moderní filologii (2015), 97 (2). [CMF2]
Časopis pro moderní filologii (2016), 98 (1). [CMF3]
Časopis pro moderní filologii (2016), 98 (2). [CMF4]
Časopis pro moderní filologii (2017), 99 (1). [CMF5]
Časopis pro moderní filologii (2017), 99 (2). [CMF6]
International Journal of English Studies (2016), 16 (1). [IJES1]
International Journal of English Studies (2017), 17 (1). [IJES2]
Journal of English for Academic Purposes (2018), 31. [JEAP1]
Journal of Linguistics (2015), 51. [JL1]
Journal of Linguistics (2016), 52. [JL2]
Journal of Linguistics (2017), 53. [JL3]
Journal of Pragmatics (2016), 103. [JP1]
Journal of Pragmatics (2016), 98. [JP2]
Journal of Pragmatics (2018), 123. [JP3]
Slovo a slovesnost (2016), 77 (1). [SS1]
Slovo a slovesnost (2017), 78 (2). [SS2]
Slovo a slovesnost (2017), 78 (3). [SS3]
Slovo a slovesnost (2017), 78 (4). [SS4]

References

Connor, Ulla & Mauranen, Anna. 1999. Linguistic analysis of grant proposals: European union research grants. *English for Specific Purposes* 18(1). 47–62.

De Carvalho, Gisele. 2001. Rhetorical patterns of academic book reviews written in Portuguese and in English. *Proceedings of the 2nd International Linguistics Conference*. Rio de Janeiro: Universidade do Estado do Rio de Janeiro, 261–268.

Ding, Huiling. 2007. Genre analysis of personal statements: Analysis of moves in application essays to medical and dental schools. *English for Specific Purposes* 26(3). 368–392.

Dontcheva-Navrátilová, Olga. 2016. Cross-Cultural Variation in the Use of Hedges and Boosters in Academic Discourse. *Prague Journal of English Studies*, Prague: Charles University 5 (1),163-184.

Dontcheva-Navrátilová, Olga. 2018. Persuasion in academic discourse: Cross-cultural variation in Anglophone and Czech academic book reviews. In Jana Pelclová & Wei-lun Lu. *Persuasion in Public Discourse: Cognitive and Functional Perspectives*. Amsterdam/Philadelphia: John Benjamins, 227-257.

Flowerdew, John & Wan, Alina. 2010. The linguistic and the contextual in applied genre analysis: The case of the company audit report. *English for Specific Purposes* 29(1). 78–93.

Gea Valor, Maria-Lluisa & del Saz Rubio, María Milagros. 2000-2001. The coding of linguistic politeness in the academia book review. *Pragmalingüística*, 8-9, 165-178.

Gea Valor, Maria-Lluisa. 2000. *A Pragmatic Approach to Politeness and Modality in Book Review Articles*. SELL Monograph. València: Universitat de València.

Henry, Alex & Roseberry, Robert L. 1997. An investigation of the functions, strategies and linguistic features of the introductions and conclusions of essays. *System* 25(4), 479-495.

Holmes, Richard. 1997. Genre analysis, and the social sciences: An investigation of the structure of research article discussion sections in three disciplines. *English for Specific Purposes* 16(4). 321-337.

Hyland, Ken. 2000. *Disciplinary discourses. Social interactions in academic writing*. Harlow: Pearson Education.

Hyland, Ken & Diani, Giuliana. (eds.). 2009. *Academic evaluation: Review genres in university settings*. London: Palgrave Macmillan.

Motta-Roth, Desiree. 1995. *Rhetorical features and disciplinary cultures: A genre-based study of academic book reviews in linguistics, chemistry, and economics*. Florianópolis: UFSC.

Nicolaisen, Jeppe. 2002. Structure-based interpretation of scholarly book reviews: A new research technique. *Proceedings of the Fourth International Conference on Conceptions of Library and Information Science*, 123-135.

Nwogu, Kevin Ngozi. 1997. The medical research paper: Structure and functions. *English for Specific Purposes* 16(2). 119-138.

Parodi, Giovanni. 2010. The rhetorical organization of the textbook genre across disciplines: A 'colony-in-loops'? *Discourse Studies* 12(2). 195–222.

Suárez, Lorena & Moreno, Ana I. 2008. The rhetorical structure of academic book reviews of literature: A Spanish-English cross-cultural approach. In Ulla Connor, Nagelhout, Ed & Rozycki, W William V. (eds.), *Current Research in Contrastive Rhetoric: Building Toward Intercultural Rhetoric*. 147-168. Amsterdam: Benjamins.

Swales, John. 1990. *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

Swales, John. 2004. *Research genres: Explorations and applications*. Cambridge: Cambridge University Press.

Jana Kozubíková Šandová.
Department of English
Faculty of Arts, University of South Bohemia
Branišovská 31a
370 05 České Budějovice
Czech Republic
sandova@ff.jcu.cz

In SKASE Journal of Theoretical Linguistics [online]. 2018, vol. 15, no. 3 [cit. 2018-13-12]. Available on web page http://www.skase.sk/Volumes/JTL38/pdf_doc/13.pdf. ISSN 1336-782X.

Interview with
Matthew Dryer

PS
Let me start with our traditional question. Why linguistics? What motivated you to study language and to deal with it from the typological perspective?

MD
While there are many different stories about how linguists got into linguistics, mine is a bit unusual, especially for someone of my generation, in that I got interested in linguistics before I went to university. My favourite subjects in high school were Latin, Greek and Mathematics, and my father was a professor of philosophy and was aware of Chomsky and his work and suggested that I might like to be a linguist. So I started reading before going to university.

I also got interested in typology in a rather fortuitous way. My undergraduate degree was a combination of linguistics, philosophy, mathematics, and computer science, so I have a more formal background than most typologists. As a graduate student, I attend a summer LSA institute and took a course in typology from Ed Keenan. His approach was more mathematical in the sense that he looked at linguistic operations as functions and this fit in with my background. It is unlikely that I would have become a typologist if I hadn't taken that course from Keenan.

PS
You are known for, inter alia, research into the typology of word-order. What progress has been achieved in this respect since Greenbergian times? And more generally, what is the impact of Chomskyan and Greenbergian approaches upon the current typological research?

MD
I believe that there has been substantial progress in word order typology, because of work of my own and of John Hawkins, among others. The real challenge that has not fully been solved is how to know whether particular crosslinguistic patterns reflect something linguistic as opposed to being accidents of history.

I am inclined to say that current typological research IS simply the Greenbergian approach extended to the current day. The more interesting question is whether Chomskyian approaches have had any impact on typological research. On this point, I probably have a different view from that of most typologists, most of whom probably believe that Chomskyan approaches have had no impact whatsoever.

I believe that it has has some impact. First, there is a sharp contrast between the pre-Chomskyan approach of Greenberg and the pre-Chomskyan approach of structuralist approaches. In contrast to both early generative grammar and structuralist approaches, Greenberg assumed to a large extent notions from traditional grammar: most of his universals in his classic 1963 paper are formulated in terms of notions from traditional grammar. But apart from Greenberg, the dominant paradigm until the mid-1960s were structuralist and while structuralist approaches dealt well with phonology and morphology, their approaches to syntax were clearly inadequate. While typologists did not adopt the machinery of generative grammar, generative grammar did a lot to raise the level of syntactic awareness that has contributed, I think, significantly to typological approaches.

A second reflection of the way in which generative grammar contributed to typology is the fact that linguistic typology really took off in the 1970s, and the most important paper that signaled this was Keenan and Comrie's paper on the accessibility hierarchy. But that paper was clearly an offshoot of the work in the late 1960's and early 1970's by generative linguists

looking at extraction constraints, starting with Ross's 1967 dissertation and subsequent work in the early 1970's by Chomsky. Thus the widespread increase in interest in typology in the 1970's was clearly influenced by generative grammar.

A third reflection of the impact of generative grammar on typology in the 1970's is due to Relational Grammar, a particular version of generative grammar. Relational Grammar differed from other versions of generative grammar in two ways that are relevant to typology. One was that it looked at lots of languages, in a way largely unknown in generative grammar at the time, but clearly typological. The other was the emphasis on grammatical relations. There is little question that although Relational Grammar did not last long as a theoretical framework, a lot of its ideas about grammatical relations had a major impact on typology in the 1970s that continue to this day. So, there are these three ways in which the development of typology in the 1970s was influenced by generative grammar, and overall positive influence.

However, I must immediately add that generative grammar has had next to no impact on typology since the 1970s. The impact was therefore from early version of generative grammar, which were quite different from more recent approaches

PS

Recently, there has been extensive discussion of comparative concepts reflecting different approaches to this topic. What is your standpoint to this issue?

Some proponents of the view that there are no crosslinguistic categories, such as Martin Haspelmath and William Croft, cite my 1997 paper "Are grammatical relations universal?" as the starting point for this view. Although my arguments in that paper were directed specifically at grammatical relations, arguing that grammatical relations are always language-specific and arguing against the notion of crosslinguistic grammatical relations, my arguments apply to crosslinguistic categories of other sorts as well.

But I should add that my views on this reflect two features of my educational background. The most important is the fact that my original linguistic training was in American Structuralism and the view that there are no crosslinguistic categories, at least substantive ones like nouns, subjects, agreement, and case, was assumed by American Structuralism. The situation is far less clear with respect to what one might call formal categories, like words, morphemes, phrases, sentences, phonemes, affixes, inflection, compounds, assimilation, and vowel harmony. Here, I am not so sure and it may depend on the particular notion. While American Structuralists rejected crosslinguistic substantive notions, they assumed at least some crosslinguistic formal notions, like phonemes.

A second factor behind my views on these matters reflects my background in philosophy, from which I acquired a type of ontological conservatism, not positing the existence of things without there being good reason to believe they exist. There is no evidence for crosslinguistic categories, nor is there anything that positing them explains. The attraction of crosslinguistic categories to many linguists arises from the clear similarities between language-specific categories in different languages. But I believe that we can explain these similarities without recourse to crosslinguistic categories.

PS

What are the main current tasks of typological research in your view?

MD

My own view is that people should work on whatever interests them, so I don't see things in terms of there being main current tasks.

PS
Typologists heavily depend on the data from informants. How is the reliability of data guaranteed in the remarkable project of the WALS database?

MD
While some typologists depend heavily on data from informants, many others do not. In my own research, I do not use data from informants, but instead depend almost entirely on published (and unpublished) grammatical descriptions. I look on languages as holistic systems so that one cannot really understand a particular feature of a language without knowing how it fits into the overall system. When one collects data from informants, one cannot know how that data fits into the system as a whole. It is for that reason that I do not use data from informants and most of the data in the WALS database was not based on informants. In fact, the WALS editors specifically forbade contributors to WALS from using questionnaires as a way to obtain data for their chapters, again for the same reason. We did use experts on particular languages to some extent, but that is a more reliable source of data than informants.

PS
This like any other source of cross-linguistic data is biased for morphosyntax. On the other hand, some other areas, such as word-formation, are not well represented. Do you and your WALS co-editors plan to rectify this lack of balance?

MD
The selection of topics for WALS was based on what people offered to produce chapters on. Some topics were not chosen simply because nobody proposed doing a chapter on them. Although we have not closed the door to additional WALS chapters and would accept such if someone could propose one based on a large crosslinguistic sample, nobody has yet actually followed through on proposed new chapters. In fact, apart from correcting errors, no authors other than me have added data to WALS since 2005. For these reasons, the WALS editors have no plans on filling gaps in WALS, though we are open to the possibility.

PS
Do you think that typology is paid relevant attention in university curricula in the USA and other countries?

MD
Certainly not in North America. It is paid better attention in Europe and especially Australia.

Thank you very much for the interview.

Pavol Štekauer

In SKASE Journal of Theoretical Linguistics [online]. 2018, vol. 15, no. 3 [cit. 2018-13-12]. Available on web page http://www.skase.sk/Volumes/JTL38/pdf_doc/14.pdf. ISSN 1336-782X.