Table of Contents

Resyllabification in Standard Arabic: A Constraint-Based Approach

Mousa A. Btoosh, Fahad Bin Sultan University, Saudi Arabia

This study provides data-driven insights and analyses on lexical and post-lexical resyllabification in Standard Arabic within the framework of Optimality Theory. More particularly, the study is devoted to examining the impact of prosthesis, voweled letters, affixation, and vowel epenthesis and shortening on resyllabification processes within words and across word boundaries. Results show that Standard Arabic typically makes use of prothesized non-phonemic segments, voweled letters and epenthesis to avoid inadmissible clusters. Findings also reveal that in certain cases this variety shortens long vowels to avoid lexical and post-lexical trimoraic syllables. Moreover, the data provided have well proven that complex codas resulting from the deletion of word-final short vowels or nunation utterance finally do not often adhere to the Sonority Sequencing Principle.

Keywords: *Standard Arabic, constraints, resyllabification, lexical, post-lexical*

1. Introduction

Grouping and organizing successive sounds into effective entities such as syllables is not a new idea as commonly thought of. Rather, this notion dates back to the momentous reconstruction of the behavior of Indo-European sonants in the 19th century (Goldsmith 2011: 1). A similar but more explicit reference to syllables is found in Erdmann & Dodge (1898: 185)'s attempt to examine the psychological process of reading. Based on the findings of this experimental study, the authors argue "...we never find a letter-to-letter reading process in the sense that attention is paid to individual letters in succession. In this case, the reading process rather operates on letter groups of different extent which are organized into speaking syllables or other 'grammatical groups of letter sounds." Nonetheless, the concept of syllable wasn't used in its current sense until the mid-1950s and the dawn of the generative phonology age launched by Chomsky & Halle (1968). On the whole, this era witnessed unprecedented advancement on the research theories and methodological approaches employed in the study of phonology.

Despite the lack of a consensus on not only the syllable-internal structure but also on whether or not the abstract notion of syllable has a phonetic correlation (cf. Yip 2002; Kenstowics 1994), research on prosody shows that there has been an increasing interest in the role a syllable can play towards a better understanding of phonological representation and analysis. For Haugen (1956: 216), syllable is "the most convenient framework for describing the distribution of phonemes." Fudge (1965) points to the significant role of syllable as a unit of phonological analysis. Bolinger (1975: 56) argues that the syllable owes much of its obviousness to the role it plays in rhythm. Liberman et al. (1974) contend that syllables are the first linguistic units that appear in the course of language acquisition. For Mehler (1981), syllable is the most important prelexical level of representation.

In addition to its significant role in phonological representation and analysis, syllable has shown an ever-increasing powerful presence and dominant capacity in accounting for various phonological phenomena. More particularly, syllable forms the basis of many phonological

processes such as prosthesis, syncope, apocope, degemination, epenthesis, and vowel shortening and lengthening; these alternations or processes, which depend on the syllable structure in a word, occur in order to avoid the surfacing of marked syllable structures. All the more, syllable has shown exceptional capabilities in accounting for three approaches that dominated the phonological scene during the past six decades.

Goldsmith (2011) points to three generalizations on syllable structure that were modeled into three approaches, namely, the sonority approach, the syntax-based immediate constituent approach and the finite-state approach. The sonority approach "is based on the view that each segment in an utterance has a sonority value, and that there are crests and troughs, or peaks and valleys of sonority in the speech chain, with peaks coinciding with vowels and syllable nuclei, and troughs coinciding with boundaries between syllables" (2011: 166). Accordingly, the nucleus elements are more sonorant than the marginal ones. The rises and falls in the sonority of an utterance give insights on whether or not the consonants are in the onset or in the coda. So, consonants that appear in a context of rising sonority occupy the onset position while the ones with falling sonority appear at the end of a syllable. Unlike the sonority approach, the syntax-based immediate constituent approach, which uses phrase structure rules to describe possible sequences, itemizes the constituents of structure. The third approach, on the other hand, is devoted to the permissible sequences in a language. Thus, it appears that the importance of the syllable stems from the roles it plays in phonological generalizations including phonotactic patterns, phonological processes, suprasegmentals and syllabification.

World languages considerably diverge with regard to their syllable types. Some (such as Hawaiian) allow just simple syllable structures while most others permit more complex ones (e.g. English and Arabic) (Davis 2002). Yet, irrespective of the permissible syllable structure in a language, it should be made obvious that organization of segments into syllables is governed by the well-formedness notion, which is, in turn, constrained by universal principles. This, of course, should not by any means conceal the existence of language variations, which are accounted for by language-specific rules or different rankings of universal constraints as proposed by Optimality Theory (OT, henceforth).

The emergence of OT, whose tenets were laid out by Prince & Smolensky (1993) and McCarthy & Prince (1993), led to the evolvement of several methodological and linguistic assumptions. Consequently, many of the doctrines and principles previously dominating the linguistic scene were changed and others came into prominence. For instance, language-specific rules have been replaced by universal constraints, where just constraint rankings are language particular. Also, a constraint violation within the OT framework does not render the form ungrammatical as it was the case under the umbrella of the earlier approaches. Moreover, the inviolable linear rules were replaced with violable hierarchically organized constraints. Still more, all candidates are considered in parallel. So, OT features no repairs and no step-by-step derivations (Kenstowics 1994).

Based on the OT principles, the choice of the winner or optimal candidate depends on the satisfaction of the higher-ranked constraints. However, in a more recent account, constraint ranking and parallel evaluation of the OT (Prince & Smolensky 1993) were challenged by the Harmonic Grammar and Harmonic Serialism theories (McCarthy & Pater 2016); based on the

tenets of these theories, constraints are weighted rather than ranked and multiple violations of the lower-ranked constraints outrank a violation of the higher-ranked ones.

Irrespective of the variations between the two competing versions, the selection of the optimal candidate is governed by structural and faithful sets of constraints. Structural constraints favor the unmarked forms over the faithful ones while faithfulness constraints favor the true-blue ones.

This study uses a constraint-based approach (Prince & Smolensky 1993) to investigate lexical and post-lexical resyllabification in SA. To this end, the study incorporates mappings between the input and output of phonological and morphological elements. It also involves an investigation into and accounting for the impact of the voweled letters, epenthesis and deletion of the non-phonemic segments, assimilation of the definite article /l/ to the following coronal consonant, and vowel shortening on lexical and post-lexical resyllabification processes.

2. Standard Arabic: An overview

SA refers to the High and ancient variety of written Arabic codified and recorded by Arab grammarians and philologists at the dawn of Islam. This variety is very rich in terms of inflections, where mood, case, number, gender, and most grammatical functions are marked by short vowels. It also distinguishes between three numbers (singular, dual and plural) and is almost completely uniform across the Arab world (Zughoul 1980). Like other languages, SA has constraints that govern the number, type and arrangement of sounds in syllables. Despite the variations among languages in this regard, it is crucial to highlight that the distribution of sounds in SA is, to a large extent, governed by universal tendencies and principles rather by than language-specific rules, as clarified through the examples given below.

As far as syllable types are concerned, six canonical patterns have been attested in SA; these types include: CV, CVC, CVV, CV(X)C, CVVCC. Unlike English, the onset node is obligatory in SA. Nonetheless, underlyingly branching onsets are banned from surfacing in this variety. As far as the nucleus is concerned, SA does not allow more than two slots linked to a monophthong or diphthong. Coda, on the other hand, is optional and ranges from zero to two sounds. Based on the illustrative examples below, branching codas are permissible just in utterance-final positions.

(1)

a. CV	shariba	/ʃa.ri.ba/	'he drank'
b. CVC	yalˤabu	/jal. ˤa.bu/	'he is playing/plays'
c. CVV	qa:la	/qa:.la/	'he said'
d. CVVC	ja:r#	/dʒa:r/	'neighbor'
	ma:t#	/ma:t/	'he died'
e. CVCC	sarj#	/sardʒ/	'saddle'
f. CVVCC	ˤa:mm#	/ˤa:mm/	'general'
g. CVCC	sadd#	/sadd/	'dam'

Unlike the first three syllable types, which can occur in any position, the fourth type (CVVC) may occur word internally or in a pre-pause position as a result of the deletion of word-final short vowel or nunation utterance-finally. The occurrence of other patterns (CVCC, CVVCC, and CVC_XC_X where C_XC_X stands for geminate consonants) is apparently restricted to a pre-pause position as a result of the deletion of word-final short vowels or nunation utterance finally. A close look at the examples in (2) below shows that consonant clusters resulting from the deletion of the word-final short vowels or nunation are permissible irrespective of whether or not such clusters result in a violation of the Sonority Sequencing Principle (SSP, henceforth).

(2)

      a. CVCC   fajr#   /fadʒr/   'dawn'
      c. CVCC   haml#   /ħaml/   'pregnancy'

Based on the principles of metrical phonology and weight, SA distinguishes between four types of syllables. A syllable with a non-branching nucleus and rime is light (CV) while the one with a branching nucleus or rime (CVC and CVV) is heavy. A syllable with a branching nucleus or coda (CVVC or CVCC) is superheavy whereas the one with a branching nucleus and coda (CVVCC) is extra superheavy.

Since the second half of 1960s, Arabic varieties have been a subject of ongoing research. Much of the research on Standard and dialectal Arabic has been devoted to major phonological aspects related to syllable structure, sonority scale, stress placement, syllable weight, syncope, apocope, and syllabification (Abdo 1969; Al-Ani 1970; Brame 1970; McCarthy 1979; Abu-Salim 1982; Alghazo 1987; McCarthy & Prince 1990; Abu Mansour 1995; Adra 1999; Mobaidin 1999; Watson 2002; Abuabbas 2003; Btoosh 2006; Dickins 2007; Rakhiya 2009; Al Tamimi & Al Shboul 2013; Youssef 2013; Heselwood & Watson 2013; Al Mashaqba 2015; Hwaidi 2016; Btoosh 2018, to name just a few). A careful examination of such studies shows that Arabic varieties diverge considerably with regard to syllable structures, syllable-based phonological processes and syllabification.

Like other Arabic varieties, SA does not allow for onsetless syllables. Yet, it still varies from other varieties in terms of the branching onset. That is, branching onsets, which are admissible in many Arabic varieties, are impermissible in SA. As far as the nucleus and coda are concerned, SA exhibits some common features with almost all other Arabic varieties. For instance, a nucleus dominates no more than two timing slots linked to a monophthong or a diphthong. Likewise, a coda may dominate up to two slots linked to consonants or glides.

3. Syllabification and resyllabification in SA

Despite the phonotactic variations among languages, syllabification, the process of dividing a word into its constituent syllables, is constrained by numerous universal principles. Chief among them are the Legality Principle, SSP, and Maximum Onset Principle. The Legality Principle constrains the types and number of segments that can begin and end syllables. Put differently, segments to the left and right of the nucleus are subject to language-specific rules (Steriade 1999; Goslin &

Frauenfelder 2001). However, this principle has a serious shortcoming as it is unable to impose a sole syllabification algorithm. For instance, the word *security* could be syllabified as [sə.ˈkjʊ.rə.ti], [sək.jʊ.rə.ti] or [sə.ˈkjʊr.ə.ti] since this principle imposes restrictions on the segments that can be tolerated in the onset or coda positions, but unable to provide the appropriate parsing of segments into syllables.

The universal phonotactic principle 'SSP', on the other hand, governs the permissible sequences of consonants within syllables in terms of sonority. Accordingly, syllable-initial consonants must rise in sonority while codas fall in sonority. Yet, it should be reiterated that languages' adherence to the SSP is never absolute. Thus, it seems apparent that both Legality and Sonority Sequencing principles are meant to dictate the permissible onsets and codas in a language. Yet, neither one adequately provides instructions on the optimal parsing of syllable-boundary segments into syllables.

Unlike the Legality Principle and SSP, which have failed to adequately account for the parsing of syllable-boundary segments into syllables, the Maximum Onset Principle displays exceptional and straightforward abilities in this regard. Based on this principle, intervocalic consonants are maximally assigned to the onsets of syllables in conformity with universal and language-specific conditions. However, appropriate syllabification of segments requires the involvement of several other universal constraints as well.

Within the OT framework, syllabification is governed by two sets of constraints. Faithfulness constraints ban any change of the input structure while well-formedness (markedness) constraints enforce input modification so as to avoid marked structures. So, syllable types that surface in a language emanate from the interaction between well-formedness and faithfulness constraints pertaining to syllable structure (Kager 2004).

Based on Comrie (2006), markedness of a construction is determined by its regularity, stability, and centrality to the core of a particular language as well as by cross-linguistic generalizations about construction types. According to this study, there are several markedness criteria, including frequency, complexity, and distribution. So, markedness constraints are devoted to evaluating output well-formedness and banning (the surfacing of) all marked features, segments and structures.

An extensive review of the onsets in SA shows that any consonant can occur as an onset either word-initially or word-internally. It is also important to reiterate that complex onsets are disallowed in this variety. As a consequence, SA, like all other Arabic varieties, resorts to epenthesis or resyllabification in order to avoid the surfacing of onsetless syllables or complex onsets. So, the interaction between faithfulness constraint (DEP-IO), which requires correspondence between input and output, and markedness constraints (ONSET) and *COMPLEX^{ONS}, ends in favor of the last ones, as shown in Table (1). However, any attempt to introduce a new syllable by inserting a vowel inside the root morpheme will render the candidate suboptimal as it fatally violates the M-O-CONTIG constraint, which prevents morpheme internal epenthesis.

(3) ONSET

*[σ V 'Syllables must have onsets.' (Prince & Smolensky 1993)

(4) NoCoda

      *C]σ   ('Syllables are open.') (Kager 2004)

(5) Dep-IO

      Every segment in the output has a correspondent in the input (McCarthy & Prince 1995)

(6) * Complex^{ONS}

      *[σ CC   ('Onsets are simple.') (Kager 2004)

(7) M-O-Contig ("No M-internal insertion")

      The portions of the output standing in correspondence and belonging to the same M form
      contiguous strings - Where M ∈ {morpheme, stem} (Landman 2003)

(8) Input: /bnah/ 'daughter'

Table 1: Onset, *Complex^{ONS}, M-O-Contig >> NoCoda, Dep-IO

/bnah /	Onset	*Complex^{ONS}	M-O-Contig	NoCoda	Dep-IO
a. /bnah/		*!			
b. /ib.nah/	*!				*
c. ☞/ʔib.nah/					**
d. /ʔi.bnah/		*!		*	**
e. bi.nah			*!	*	*

Being faithful to the input inevitably entails a violation of *Complex^{ONS} since branching onsets
are disallowed in SA. This violation constitutes the sole reason for the elimination of candidate
(1a). Irrespective of whether or not it has an onset underlyingly, the Onset constraint strictly
stipulates that under no circumstances can an onsetless syllable surface in this variety. This justifies
the immediate elimination of the second candidate. So, the epenthesis of the short vowel with the
aim of avoiding the fatal violation of *Complex^{ONS} results in ultimate failure since Onset and
*Complex^{ONS} are not ranked with each other. Candidate (1c) illustrates a successful attempt to
satisfy the *Complex^{ONS} constraint through the epenthesis of elidable non-phonemic segments
consisting of a glottal stop followed by a high short vowel. By resyllabifying the first consonant
in the impermissible consonant clusters as the coda of the epenthetic nucleus [i], the third candidate
emerges as optimal in spite of incurring a double violation of the low-ranked faithfulness constraint
Dep-IO. An input containing consonant clusters word- or syllable-initially will never be able to
surface owing to the violation of the higher ranked *Complex^{ONS}, which is responsible here for
ruling out candidate (1d). The last candidate, on the other hand, is excluded as a result of the
insertion of a high short vowel inside the root morpheme. Though candidate (1c), the optimal,
shares with candidate (1e) the violation of the low-ranked constraint Dep-IO, it is the epenthesis
locus that makes the difference between them. That is, morpheme external insertion (at left edge
or morpheme juncture) is allowed while morpheme-internal epenthesis is not. This elucidates that
morpheme internal insertion in (1e) has constituted a serious breach of the top-ranked constraint
M-O-Contig and led to the elimination of this candidate from the race.

In the context of the non-phonemic epenthetic segments (/ʔ/ and /V/, where V means a 'vowel'), traditional Arabic grammarians distinguished between two types of hamza 'glottal stop' (Ryding 2005). The first type is called hamzatul qatiᶜ, 'non-connecting or non- elidable glottal stop' while the second one is known as 'hamzatul wa<u>s</u>l', the connecting or elidable glottal stop. The non-elidable glottal stop, which is also known as the strong glottal stop, is phonemic and occurs word initially, medially and finally. Having known that no Arabic word can begin with a vowel, then it is obligatory that the word-initial ʔ*alif (long vowel: aa)* needs to be always a seat for glottal stop in order to avoid onsetless syllables. In Arabic orthography, the non-elidable glottal stop is placed over the ʔ*alif* (إ) if it is followed by *fatħa* (*a*) or *ḏˤammah* (*u*), and under the ʔ*alif* (!) if it is followed by *kasrah* (*i*).

However, the elidable non-phonemic epenthetic glottal stop /ʔ/, which occurs when a word begins with a consonant-cluster, does not show up above or below the 'ʔalif'. Rather, it is a superscript <u>s</u>*aad* (diacritic) drawn above a line (إ). It appears rather apparent from the example below that epenthetic segments (/ʔ/ and /V/) are triggered only to avoid the surfacing of the underlyingly consonant-clusters word-initially, as shown in (9).

(9)   /nkasar-a/ ⟶ ʔin.ka.sa.ra   'it broke'

The glottal stop, along with the following high short vowel, is epenthesized here since it is impermissible in SA to start with a non-voweled consonant as this leads to the surfacing of a complex onset. Consequently, this glottal stop, together with the vowel that follows, is elided when preceded by a voweled letter (when joined to the preceding word). This explains the deletion of the epenthetic segments (/ʔ/ and /V/) when the ONSET and/or COMPLEX^{ONS} are satisfied.

(10) #wa# #ʔinkasar-a# ⟶ /wan.ka.sa.ra/   'and it broke'

However, when the non-elidable glottal stop is used in the same phonetic environment, the glottal stop, together with the following vowel remains unaffected, as it is part of the input rather than an epenthesized one. A simple example to show this is given in (11).

(11) /wa + ʔaħmad-u/ ⟶ /wa.ʔaħ.ma.du/   'and Ahmad'-Nom.

3.1 *Resyllabification within words*

Satisfaction of the ONSET constraint does not always necessitate epenthesis. On the contrary, SA resorts to epenthesis only when other choices fail to work. Therefore, this explains the resort of SA to resyllabification rather than epenthesis to avoid the surfacing of onsetless syllables, as illustrated in Table (2). Based on the examples below, resyllabification, a phonotactically-motivated repair process, is applied cyclically when a vowel-initial suffix is added. Therefore, resyllabifying the last member of the coda as an onset of the following vowel-initial syllable will result in a clear breaching of ALIGN (R).

(12) ALIGN (R)
       'Align root morpheme boundaries with syllable boundaries at both edges.' (Yip 1994)

(13) Input: /kalb+an/ 'a dog' (accusative)

Table 2: ONSET, *COMPLEX^{ONS}, M-O-CONTIG >>DEP-IO, NOCODA· ALIGN (R)

#kalb#+#an#	ONSET	*COMPLEX^{ONS}	M-O-CONTIG	DEP-IO	NOCODA	ALIGN (R)
a. /kalb.an/	*!					
b./ka.lban/		*!			*	*
c. ☞/kal.ban/						*
d. /ka.li.ban/			*!	*	**	*

The first candidate is rejected on account of its violation of the markedness constraint ONSET, which prevents onsetless syllables from surfacing. Despite satisfying ONSET, the second candidate is ruled out due to incurring a violation of the markedness constraint *COMPLEX^{ONS}. The optimal candidate (2c), on the other hand, wins the race by accomplishing a total satisfaction of all constraints except for the rightmost (lowest-ranked) one ALIGN (R). In spite of being in complete harmony with the top-ranked markedness constraints (ONSET and *COMPLEX^{ONS}), the last candidate is eliminated as it fatally violates the top-ranked faithfulness M-O-CONTIG constraint.

In addition to the onset-motivated resyllabification process, SA resorts to onset-motivated epenthesis when a word begins with a non-voweled letter. The structure in (14), which is the outcome of applying resyllabification cyclically, presents examples of both onset-motivated resyllabification and onset-motivated epenthesis processes. Onset-motivated epenthesis is triggered by the complex onset, which is banned in this variety. Consequently, the first element of the complex onset is resyllabified as a coda of the epenthetic CV (where C refers to the glottal stop and V to the high short vowel). Onset-motivated resyllabification, on the other hand, is triggered by the newly formed onsetless syllable resulting from affixation [-u].

(14)
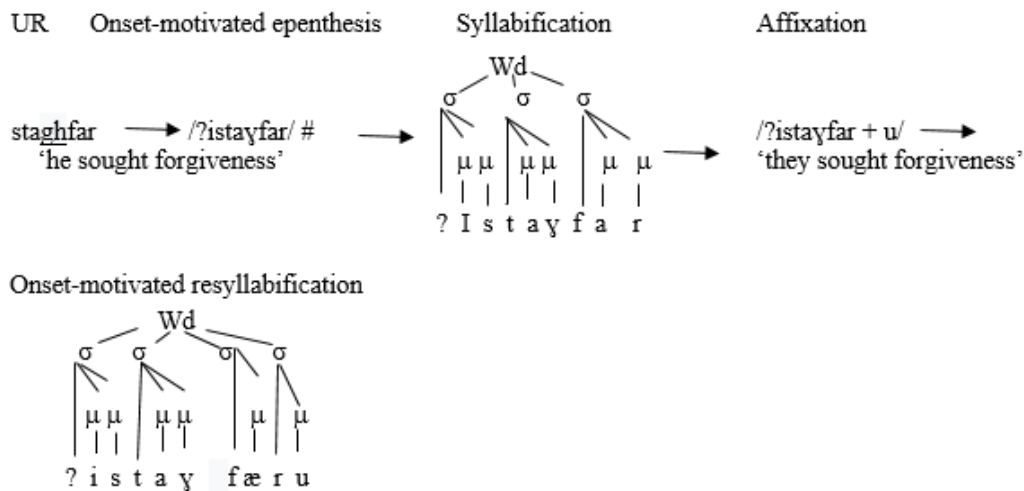
The question that arises here is whether or not it is permissible to break the CC cluster in the underlying form above 'staghfar' by just changing the epenthesis locus. That is, to break such as a cluster by inserting a high short vowel between C_1 and $C_{2,}$ (si.tagh.fa.ru). While this choice seems less costly than epenthesizing

the two other segments (/ʔ/ and /v/), it still remains not the right choice as the resulting output form is ruled out by the M-O-CONTIG constraint.

(15)   Input: /staɣfar+u/ 'they (m.) sought forgiveness'

Table 3: ONSET, *COMPLEX^{ONS}, M-O-CONTIG >> NOCODA, DEP-IO, ALIGN (R)

/staɣfar+u/	ONSET	*COMPLEX ONS	M-O-CONTIG	NOCODA	DEP-IO	ALIGN (R)
a. /is.taɣ.fa.ru/	*!			**	*	*
b.☞/ʔis.taɣ.fa.ru/				*	**	*
c. /staɣ.fa.ru/		*!		**		*
d. /si.taɣ.fa.ru/			*!	***	*	*

The grammar of this variety vehemently rejects onsetless syllables and consistently gives onsets priority over codas. Accordingly, VCCV and VCV are always syllabified as VC.CV and V.CV, respectively. Thus, the first candidate loses the competition as it commits a fatal violation of the top-ranked constraint ONSET. The optimal output (3b) violates the lowest-ranked constraints NOCODA, DEP-IO and ALIGN (R). No doubt, the violation of these constraints is necessary for this candidate in order to avoid the penalty of the higher-ranked constraints ONSET and *COMPLEX^{ONS}. That is, via the resyllabification of the onset [s] as the coda of the epenthetic nucleus [i], candidate (3b) could break up the inadmissible consonant clusters and steer clear of the penalty of the top-ranked constraint *COMPLEX^{ONS}. Also, this candidate could avoid incurring a violation of the undominated ONSET by resyllabifying the coda of the underlying penultimate syllable as the onset of the ultimate syllable. Though it fully satisfies the undominated ONSET, candidate (3c) is excluded by *COMPLEX^{ONS}. In spite of being in full compliance with the two top constraints ONSET and *COMPLEX^{ONS}, the last candidate is ruled out by M-O-CONTIG, which bans morpheme medial epenthesis.

Epenthesis in SA is not restricted to the non-phonemic segments, as shown above. Rather, it is also used to avoid the surfacing of impermissible consonant clusters, which result in trimoriac syllables or a complex onset. For instance, the emphatic form of the imperfect verb *yadrus* 'he studies' is formed by suffixing the syllable –nna to the root morpheme /ja+drus+nna/. For that reason, epenthesizing a low short vowel in the example below seems inescapable to avoid having a trimoraic syllable.

(16)   *3μ
          ('No trimoraic syllables')   (Kager 2004)

(17)   Input: /jadrus+nna/ 'he studies/most certainly will study'

Table 4: *COMPLEX^ONS, *3μ, M-O-CONTIG>> NOCODA, DEP-IO, ALIGN (R)

#jadrus#+#nna#	*COMPLEX^ONS	*3μ	M-O-CONTIG	NOCODA	DEP-IO	ALIGN (R)
a. /jad.rusn.na/		*!		*		*
b. /jad.rus.nna/	*!			*		
c. ☞/jad.ru.san.na/				**	*	*
d. /ja.di.rus.nna	*!		*	***	*	

The first candidate is eliminated by the *3μ constraint as syllables are maximally bimoraic. Candidate (4b) fatally violates *COMPLEX^ONS, which is undominated in this variety. In spite of violating the low ranking constraints NOCODA, DEP -IO and ALIGN (R), the third candidate, emerges as optimal simply by both epenthesizing a low short vowel [a] to break the impermissible consonant clusters [-nna] and resyllabifying the first member of the suffix [-nna] as a coda of the epenthetic nucleus. By so doing, this candidate could avoid the penalty of the high-ranking constraints. A comparison between candidates (4c) and (4d) reveals that M-O-CONTIG bans epenthesis morpheme internally, but not at morpheme juncture. This explains the elimination of the last candidate. Though it is beyond the scope of this paper, the type of the epenthetic vowel (*i*, *a*, or *u*) in Arabic varieties, in general, is primarily governed by the position of the cluster (initial or medial), definite article, and verb pattern.

In addition to epenthesis, SA sometimes resorts to vowel shortening to avoid non-final inadmissible syllable structures. For instance, the long vowel in CVVC syllables in verbs such as *yakha:f* 'he fears' in Table (5) below is reduced when suffixed with the plural feminine morpheme –na. It seems obvious from the constraint interactions below that SA prefers adhering to *3μ over breaching MAX-V-IO. Also, this variety imposes severe restrictions on non-final light syllables.

(18) MAX-V-IO
Input vowels must have output correspondents. ('No vowel deletion.') (Kager 2004)

(19) REDUCE
Minimize the number of non-final light syllables. (Kiparsky 2003)

(20) Input: /jaxa:f+na/ 'they (f.) fear'

Table 5: *COMPLEX^ONS, *3μ, REDUCE >> DEP-IO, MAX-V-IO, NOCODA, ALIGN (R)

#jaxa:f#+#na#	*COMPLEX^ONS	*3μ	REDUCE	DEP-IO	MAX-V-IO	NOCODA	ALIGN (R)
a. /ja.xa:f.na/		*!				**	
b. /ja.xa:.fna/	*!		*			***	*
c. ☞/ja.xaf.na/			*		*	**	
d. /ja.xa:.fa.na/			**!	*		****	*

Candidate (5a) preserves all the input segments. Nevertheless, it is entirely taken out of the race by *3μ, which strictly stipulates that a syllable is maximally bimoraic. Adjoining the last

consonant of the root to the following syllable successfully helps the second candidate avoid the penalty of *3μ. But still, this attempt has rendered candidate (5b) suboptimal as it fatally violates *COMPLEX^{ONS}. Candidate (5c) turns up the optimal output by managing to escape the penalties of both *COMPLEX^{ONS} and *3μ via deleting the input segment [a] and parsing the suffix as a separate syllable. Though it is in harmony with the *3μ, the last candidate is eliminated due to the double violation of the REDUCE constraint, which disfavors non-final light syllables. Although both (5c) and (5d) violate REDUCE, the former is selected since it minimally violates REDUCE. That is, (5c) has only one violation while its contestant (5d) incurs two violations.

3.2 *Resyllabification across word boundaries*

Resyllabification in Arabic varieties, including SA, is applied post-lexically as well. It is worth mentioning that post-lexical resyllabification is common in all colloquial Arabic varieties, which mainly resort to epenthesis to break impermissible consonant clusters. However, the case is substantially different in SA as the last consonant is almost always voweled unless in pre-pause position as a result of the deletion of word-final short vowel utterance-finally. Drawing on cross-sectional analysis of the data below, it clearly emerges that the definite article in Arabic constitutes a major phonological domain for resyllabification.

In the context of the definite article in Arabic, it is essential to mention that this article consists of just one segment /l/, called '*la:m atta^cri:f*' (the definite article). However, since complex onsets cannot surface in Arabic, then the addition of the definite article to a noun (or an adjective) always results in banned consonant clusters (Gadoua 2000). For this reason, a new syllable consisting of a glottal stop followed by a short vowel is inserted before the /l/, as shown in (21).

(21) Indefinite Noun	UR Definite Noun	SR Definite Noun	Gloss
ba:b	l-ba:b	?al.ba:b	'the door'
^calam	l-^calam	?al.^calam	'the flag'

When the definite article */l/* is followed by sun (or solar) letters, coronal consonants, it (*the definite article 'l'*) completely assimilates to the following noun or adjective's initial consonant. As such, the following sun letter becomes a geminate, as shown in (22).

(22) a. ?al-shams-u > /?aʃ. ʃam.su/ 'the sun' (NOM)
      b. ?al-da:r-u > /?ad.da:.ru/ 'the house' (NOM)

It is noticeable that both the definite article [l] and the following sun letter share the same feature, viz., [+cor], consequently violating the Obligatory Contour Principle (OCP).

(23) OCP
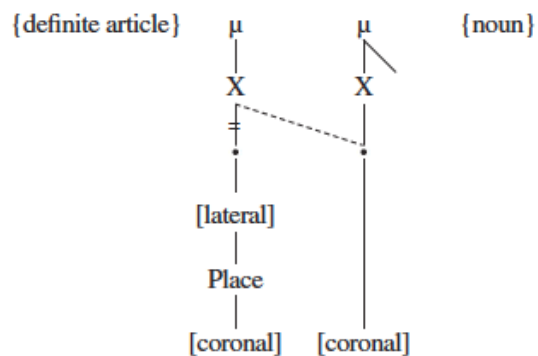   At melodic level, adjacent identical elements are prohibited.   (McCarthy 1986)

Watson (2002: 220) depicts OCP violation, as in (24).

(24)

```
{definite article}    μ              μ          [noun]
                      |              /\
                      x              x
                      |              |
                      •              •
                      |
                  [lateral]
                      |
                   Place
                      |
                  [coronal]      [coronal     OCP violation
```
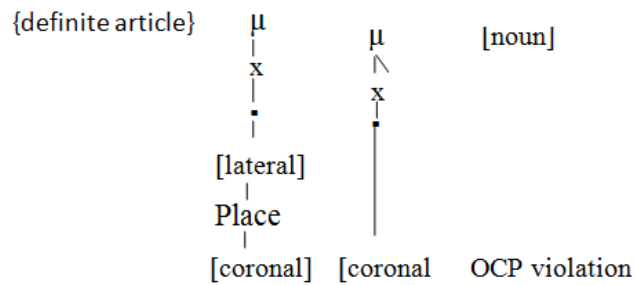
The assimilation of the definite article /l/ to the following coronal triggers the violation of the OCP on the coronal tier. To overcome this dilemma, Watson (2002: 220) notionally represents the assimilation of –l to the following coronal, as shown in (25).

(25)

```
{definite article}    μ              μ          {noun}
                      |              /|
                      X              X
                      ‖- - - - - - - |
                      •              •
                      |
                  [lateral]
                      |
                   Place
                      |
                  [coronal]    [coronal]
```

Based on this representation, the leftmost root node is deleted and the rightmost node spreads to the left node. It is noticeable from the representation in (25) that the two adjacent different segments end up in identical segments (geminates). Irrespective of the violations incurred against IDENT-IO (αF), the newly formed geminate is in absolute harmony with the OCP, [LAT COR] SHARE (F), and MAX-C-IO constraints.

(26) IDENT-IO (αF)
Output correspondents of an input [αF] segment are also [αF]   (McCarthy & Prince 1995: 264).

(27) [LAT COR] SHARE (F)
Across a morpheme boundary, /l/ and the following coronal consonant should be assigned the same token features   (McCarthy 2010)

(28) MAX-C-IO
'Input consonants must have output correspondents. ('No consonant deletion.' (McCarthy & Prince 1995)

(29) Input: /l-da:r/ 'the house'

Table 6: Ocp, [Lat Cor] Share (F), Max-C-IO, *Complex^ONS >> Dep-IO, Ident-IO (αF)

/l-da:r/	Ocp	[Lat Cor] Share (F)	Max-C-IO	*Complex^ONS	Dep-IO	Ident-IO (αF)
a. /?a.da:r/ d | Place | [coronal			*!		**	
b. /?a.dda:r/ d d | | Place Place | | [coronal] [coronal]	*!			*	**	*
c. /?al.da:r/ l d | | [lateral] Place | Place | [coronal] [coronal]	*!	*			**	
d. ☞/?ad.da:r/ l d x x [lateral] | Place | [coronal] [coronal]					**	*

By deleting the definite article *[l]*, candidate (6a) fully satisfies *Ocp, but to no purpose since it has incurred a fatal violation of the Max-C-IO. Avoiding the pitfall of the first candidate has not been sufficient to prevent candidate (6b) from being ruled out by *Complex^ONS, which bans branching onsets. Also, this candidate is less harmonic than candidate (6a) with respect to the top-ranked Ocp constraint. Despite attempting to be quite faithful to the input form, candidate (6c) is excluded as it breaches the top-ranked constraints Ocp and [Lat Cor] Share (F). The optimal candidate (6d) incurs a double violation of Dep-IO and a single violation of Ident-IO (αF). Yet, such violations are least expensive and necessary to escape the penalty of the left-most top-ranked constraints.

Post-lexical resyllabification, which is common in all Arabic varieties, merely creates an environment where word boundaries and syllable boundaries do not coincide. This will unquestionably always lead to the violation of Align (W) constraint, as shown in Table (7).

(30) ALIGN (W)

'Align the right edge of a word with the right edge of a syllable.' (Harris & Gussmann 2003)

(31) Input: /#masaħt-u# #l-ʃa:riʕ-a# 'I wiped the street'

Table 7: OCP, [LAT COR] SHARE (F), ONSET, *3μ, MAX-C-IO >> IDENT-IO (αF), ALIGN (W)

#masaħt-u# #l-ʃa:riʕ-a#	OCP	[LAT COR] SHARE (F)	ONSET	*3μ	MAX-C-IO	IDENT-IO (αF)	ALIGN (W)
a. /ma.saħ.tuʃ.ʃa:r.i.ʕa/			*!	*		*	*
b. /ma.saħ.tul.ʃa:.ri.ʕa/	*!	*					*
c. ☞ /ma.saħ.tuʃ.ʃa:.ri.ʕa/						*	*
d. /ma.saħ.tu.ʃa:.ri.ʕa/					*!		*

In spite of its full compliance with the OCP, the first candidate commits two fatal violations against ONSET and *3μ. It should be made plain that the last consonant in superheavy syllables (CVC, CVCC and CVVC) is extrasyllabic or invisible word-finally since the last consonant in these types

of syllables is always extrametrical. So, having a trimoraic syllable always renders the syllable inadmissible. By parsing the [r] as the onset of the following syllable candidate (7b) successfully avoids the penalty of both ONSET and *3μ. But this attempt has been proven to be of no avail since this candidate is ruled out by the top-ranking constraints OCP and [LAT COR] SHARE (F). Candidate (7c) wins over other candidates (which violate the higher-ranked constraints) as it has successfully satisfied all the leftmost ranked constraints. The last candidate is discarded by MAX-C-IO since consonant deletion is not tolerated in this variety.

Unlike sun letters, where the definite article [1] loses its distinctive features and is assimilated to the initial consonant of the following noun or adjective, the definite article retains all its features when followed by moon (or lunar) letters. As such, by being non-coronals, the definite article [1] does not assimilate to the following moon letters, as shown in (32).

(32)
      a.   l-qamar-u > /?alqamar –u/ 'the moon (NOM)
      b.   l-ba:b-u   > /?alba:b-u/ 'the door'   (NOM)

Drawing on (32), the definite article retains its input features when followed by a non-coronal sound; the input features of the definite article /l/ are preserved by IDENT-IO (PLACE).

(33)   IDENT-IO (PLACE)

      Correspondents in input and output have identical place features. (Kager 2004)

It is evident from Tables (7) and (8) that resyllabification takes place post-lexically irrespective of whether the following noun (or adjective) starts with a sun or moon letter. A comparison between the two examples shows that except for the assimilation of the definite article /l/ to the following sun letters, *parsing* of sun and moon letters is quite symmetrical.

(34)   Input: #fataħ-a# # l-ba:b-a#/ 'he opened the door'

Table 8: OCP, IDENT-IO (PLACE), ONSET, *3μ, MAX-C-IO >> NOCODA, ALIGN (W)

#fataħ-a# #l-ba:b-a#	OCP	IDENT-IO (PLACE)	ONSET	*3μ	MAX-C-IO	NOCODA	ALIGN (W)

a. /fa.ta.ħal.ba:b.a/
 	
  1   b
  | |
[lateral] Place
  | |
Place
  |
[coronal] [labial]
			*!	*		***	*

b. /fa.ta.ħab.ba:.ba/		*!				****	*

l — b x — x [lateral] Place [coronal] [labial]							
c. ☞ /fa.ta.ħal.ba:.ba/ l b [lateral] Place Place [coronal] [labial]						****	*
d. /fa.ta.ħa.ba:.ba/ b Place [labial]					*!	*****	

Candidate (8a) presents the actual output. Yet, it is not the most harmonic, as it incurs fatal violations of ONSET and *3µ. The full assimilation of the definite article to the following non-coronal sound renders the second candidate suboptimal as it excessively violates IDENT-IO (PLACE), which stipulates the preservation of the input place features. Candidate (8c), on the other hand, wins the contest as it has free violations of the topmost set of constraints. The bottommost candidate is ruled out as it incurs a fatal violation of the high-ranking constraint MAX-C-IO.

In addition to the definite article, post-lexical resyllabification occurs when a preposition precedes the elidable glottal stop as in (35). It is rather evident from the example below that the elidable glottal stop deletes and a new syllable is formed after shortening the long vowel of the preposition [fi:] so as to avoid the surfacing of a trimoraic syllable.

(35)

    #fi:# #?al-maktab-i# > /fil.mak.ta.bi/
                        'in the office'

Post-lexical resyllabification runs along similar lines in vocatives with regard not only to the drop of the prothesized syllable, but also to the shortening of the long vowel (Alqaasem 1993). Table (9) illustrates these exceptional cases.

(36) Input: /#ja:#   # bn-a#   #?axi#/
           vocative   son   brother my
               'oh my nephew'

Table 9: *3μ, MAX-C-IO, ONSET, *COMPLEX^ONS >> NO CODA, MAX-V-IO, ALIGN (W)

/#ja:# #bn-a# #ʔaxi#/	*3μ	MAX-C-IO	ONSET	*COMPLEX^ONS	NO CODA	MAX-V-IO	ALIGN (W)
a. /ja:b.na.ʔa.xi/	*!				***		*
b. /ja:.bna.ʔa.xi/				*!	****		*
c. /jab.naa.xi/		*!			**	*	*
d. ☞/jab.na.ʔa.xi/					***	*	*

Being fully faithful to the input form is not sufficient to rescue candidate (9a) as it violates the high-ranked *3μ. By parsing two consonants as an onset of the second syllable, the second candidate could manage to avoid the penalty of *3μ, but at the expense of another fatal constraint *COMPLEX^ONS in addition to the numerous incurred violations of the low-ranking constraints ALIGN (W) and NOCODA. Candidate (9c), on the other hand, is discarded by MAX-C-IO, which prevents the deletion of input consonants. Despite the three incurred violations of the low-ranked constraints, the last candidate emerges as optimal since it has successfully satisfied all the top-ranked constraints.

The organization of segments within a syllable is usually governed by the SSP, where more sonorous segments come close to the syllable peak. However, languages clearly exhibit divergence with regard to the adherence to the SSP. A study conducted on coda clusters in SA shows that clusters that do not adhere to the SSP outweigh the ones that do (Al Tamimi & Al Shboul 2013). It should be reiterated that such clusters result from the deletion of the final short vowel or nunation (upon pausing). The following example shows that the deletion of the word-final short vowel or nunation results in a clear violation of the SSP, as shown in Table (10).

(37) SON
'In a syllable, sonority increases toward the peak and decreases toward the margins.'

(Morelli 2003)

(38)   Input: /ʕasˤr/# 'age'

Table 10: ONSET, M-O-CONTIG>> DEP-IO, SON, NOCODA

/ʕasˤr/#	ONSET	M-O-CONTIG	DEP-IO	SON	NOCODA
a. /ʕa.sir/		*!	*		*
b. ☞ /ʕasˤr/				*	
c. /ʕasˤ.ir/	*!				

Splitting the underlying consonant clusters in coda position via the epenthetic high short vowel [i] has rendered candidate (10a) suboptimal as M-O-CONTIG strictly prohibits morpheme internal epenthesis. Although candidate (10b) violates SON, it is still selected as the optimal candidate owing to being the form with the least violations of the most harmful constraints. Therefore, input-output correspondence in SA is given precedence over the adherence to the SON constraint. The last candidate is eliminated by ONSET, which is undominated in all Arabic varieties.

4 Conclusion

This paper has shed light on some key issues on lexical and post-lexical resyllabification within the OT framework. By concentrating on descriptions, analyses, and fundamental phonological processes, the paper has attempted to explore the phonological and morphological factors that constrained the resyllabification process, in general. Evidence has shown that there are a lot of variations in the processes that SA employs to evade the surfacing of impermissible syllable structures and consonant clusters. Based on the evidence provided, it has become rather obvious that SA adopts several phonological processes such prosthesis, syncope and vowel shortening to avert the surface of impermissible clusters.

References

Abdo, Daud. 1969. *On Stress and Arabic Phonology: A Generative Approach*. Beirut: Khayats.

Abu-Abbas, Khaled. 2003. *Topics in the phonology of Jordanian Arabic: An optimality theory perspective*. Kansas: University of Kansas. (Doctoral dissertation.)

Abu Mansour, Mahasen Hasan. 1995. Optimality and conspiracy in the syllable structure of Arabic. In Beckman, Jill & Dickey, Laura Walsh & Urbanczyk, Suzanne (ed.), *University of Massachusetts Occasional Papers: Papers in Optimality Theory* 18. 1–20.

Abu-Salim, Isam. 1982. *A reanalysis of some aspects of Arabic phonology: A metrical approach*. Urbana: University of Illinois at Urbana-Champaign. (Doctoral dissertation.)

Adra, Mohammad. 1999. *Identity effects and opacity in Syrian Arabic: An optimality theory analysis*. Urbana: University of Illinois at Urbana-Champaign. (Doctoral dissertation.)

Al-Ani, Salman. 1970. *Arabic Phonology: An Acoustical and Physiological Investigation*. The Hague: Mouton.

Algasem, Yahya. 1993. Athar almaqta' almarfood fi buniat alkalimati al'arabiyah [Translated: The Impact of the Rejected Syllable in the Structure of Arabic Word], *Abhaath Al-Yarmouk* 11(2). 149 –178.

Alghazo, Mohammad. 1987. *Syncope and epenthesis in Levantine Arabic: A nonlinear approach*. Urbana: University of Illinois at Urbana-Champaign. (Doctoral dissertation.)

Al Mashaqba, Bassil. 2015. *The phonology and morphology of Wādi Ramm Arabic.* Salford University, UK.

Al Tamimi, Yaser & Al Shboul, Yousef. 2013. Is the phonotactics of the Arabic complex coda sonority-based?, *Journal of King Saud University – Language and Translation* 25(1). 21–33.

Bolinger, Dwight. 1975. *Aspects of Language*, (2nd edition). New York: Harcourt Brace Jovanovich.

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

Btoosh, Mousa A. 2006. Constraint interactions in Jordanian Arabic phonotactics: An optimality theoretic approach. *Journal of Language and Linguistics* 5(2). 192– 221.

Btoosh, Mousa A. 2018. Rules or constraints? A cross methodological comparison of approaches to 'syncope' and 'vowel shortening' in Cairene Arabic. *SKASE Journal of Theoretical Linguistics* 15(1). 59–79.

Chomsky, Noam & Halle, Morris. 1968. *The Sound Pattern of English*. New York: Harper and Row.

Comrie, Bernard. 2006. Transitivity pairs, markedness, and diachronic stability. *Linguistics* 44(2). 303–318.

Davis, Stuart. 2002. Syllable Structure for an Artificial Language Based on Universal Principles. *Journal of Universal Language* 3. 3–15.

Dickins, James. 2007. *Sudanese Arabic: Phonematics and Syllable Structure*. Wiesbaden: Harrassowitz.

Erdmann, Benno & Dodge, Raymond. 1898. *Psychologische Untersuchungen fiber das Lesen auf experimenteller Grundlage*. Halle: Niemeyer.

Fudge, Eric. 1965. The nature of phonological primes. *Journal of Linguistics* 3. 1–36.

Gadoua, Abdulhamid. 2000. '*Consonant clusters in Quranic Arabic'*. *Cahiers Linguistiques d'Ottawa* 28. 59–85.

Goldsmith, John. 2011. The syllable. In John A. Goldsmith, Jason Riggle & Yu, Alan C. L. (eds.). *The Handbook of Phonological Theory*, (2nd edition), 164–196. Chichester and Malden: Wiley-Blackwell.

Goslin, Jeremy & Frauenfelder, Uli. 2001. A comparison of theoretical and human syllabification. *Language and Speech* 44(4). 409–436.

Harris, John & Gussman, Edmund. 2003. Word-Final Onsets. ROA- 575-0203. http://roa.rutgers.edu/view.php3?roa=575 (Retrieved on May 12, 2018).

Haugen, Einar. 1956. The syllable in linguistic description. In: Halle, Morris, Lunt, H. G. & McLean. H. (eds.), For Roman Jacobson, 213–221. Den Haag: Mouton.

Heselwood, Barry & Watson, Janet C. E. 2013. The Arabic definite article does not assimilate. *Leeds Working Papers in Linguistics and Phonetics* 18. 34–53.

Hwaidi, Tamader. 2016. *Syllable structure and syllabification in Al'ain Libyan Arabic*. Newcastle: Newcastle University. (Doctoral dissertation.)

Kager, René. 2004. *Optimality Theory*. Cambridge: Cambridge University Press.

Kenstowicz, Michael. 1994. *Phonology in Generative Grammar*. Cambridge (Mass.) Oxford: Blackwell.

Kiparsky, P. 2003. Syllables and Moras in Arabic. In Caroline Fery & Ruben Van De Vijver (Eds.), *The Syllable in Optimality Theory*, 147–183. Cambridge: Cambridge University Press.

Landman, Meredith. 2003. Morphological contiguity. In Carpenter, Andrea, Coetzee, Andries & de Lacy, Paul (eds.), *University of Massachusetts Amherst Occasional Papers in Linguistics 26: Papers in Optimality Theory II*. Amherst: Graduate Student Linguistic Association Publications.

Liberman, Isabelle & Shankweiler, Donald & Fischer, F. William & Carter, Bonnie. 1974. Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology* 18. 201–212.

McCarthy, John. 1979. *Formal Problems in Semitic Phonology and Morphology*. Cambridge, Mass. (Doctoral dissertation.)

McCarthy, John. 1986. OCP effects: Gemination and antigemination, *Linguistic Inquiry* 17. 207–63.

McCarthy, John. 2010. Autosegmental spreading in Optimality Theory. In John Goldsmith, Elizabeth Hume, & Leo Wetzels, (eds.), *Tones and Features*, 195–222. Berlin and New York: Mouton de Gruyter.

McCarthy, John J. & Alan Prince. 1993. *Prosodic Morphology I: Constraint Interaction and Satisfaction*. Ms., University of Massachusetts, Amherst and Rutgers University, New Brunswick.

McCarthy, John & Alan Prince. 1995. Faithfulness and reduplicative identity. *University of Massachusetts Occasional Papers* 18. 249–384.

McCarthy, John & Joe Parter (eds.). 2016. *Harmonic Grammar and Harmonic Serialism*. London: Equinox.

Mehler, J., Dommergues, J., Frauenfelder, U., & Segui, J. 1981. The syllables role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior* 20. 298–305.

Mobaidin, Hosam Eddin. 1999. Consonant underpaying in Arabic: An optimality theoretic approach. Dirasat, *Proceedings of F.I.C.A.E.C.C.S Special Issue*, University of Jordan.

Morelli, Frida (2003), The Relative Harmony of /s stop/Onsets: Obstruent Clusters and the Sonority Sequencing Principle. In Féry, Carolin & van de Vijver, Ruben (eds.), *The Syllable in Optimality Theory*, 356–371. Cambridge: Cambridge University Press.

Prince, Alan & Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Blackwell, Malden, MA and Oxford, UK.

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

Ryding, Karin. 2005. *A Reference Grammar of Modern Standard Arabic*. Cambridge: Cambridge University Press.

Steriade, Donca. 1999. Alternatives to syllable-based accounts of consonantal phonotactics. In Osamu Fujimura, Brian Joseph & Palek, Bohumil (eds.), *Proceedings of the 1998 Linguistics and Phonetics Conference*, 205–242. Prague: Karolinum Press.

Watson, Janet. 2002. *The Phonology and Morphology of Arabic*. Oxford: Oxford University Press.

Yip, Moira. 1994. *Phonological constraints, optimality, and phonetic realization in Cantonese*. Colloquium, University of Pennsylvania.

Yip, Moira. 2002. *Tone*. Cambridge: Cambridge University Press.

Youssef, Islam. 2013. *Place assimilation in Arabic*: *contrasts*, *features, and constraints*. Tromsø: University of Tromsø. (Doctoral dissertation.)

Zughoul, Muhammad. 1980. Diglossia in Arabic: Investigating solutions. *Anthropological Linguistics* 22(5). 201–217.

Mousa A. Btoosh, Ph.D.,
Dept. of English Language and Translation
Fahad Bin Sultan University
P.O. Box 15700
Tabuk 71454, SA
e-mail: mbtoosh@fbsu.edu.sa
*   mousa.btoosh@gmail.com*

Appendix

Romanized symbols for Arabic consonants and vowels and their corresponding IPA symbols

Romanized Symbols Used for Arabic Consonants and Vowels	International Phonetic Alphabet
?	[ʔ] glottal plosive
b	[b] voiced bilabial plosive
t	[t] voiceless alveolar plosive
th	[θ] voiceless inter-dental fricative
j	[dʒ] voiced post-alveolar affricate
h̲	[ħ] voiceless pharyngeal fricative
kh	[x] voiceless velar fricative
d	[d] voiced alveolar plosive
th̲	[ð] voiced inter-dental fricative
r	[ɾ] alveolar tap
z	[z] voiced alveolar fricative
s	[s] voiceless alveolar fricative
sh	[ʃ] voiceless post-alveolar fricative
s̲	[sˤ] emphatic voiceless alveolar fricative
d̲	[d̪ˤ] emphatic voiced alveolar plosive
t̲	[t̪ˤ]voiceless dental plosive
TH	[ðˤ] emphatic voiced alveolar fricative
c	[ʕ] voiced pharyngeal fricative
gh	[ɣ] voiced velar fricative
f	[f] voiceless labiodental fricative
g̲	[q] voiced uvular plosive
k	[k] voiceless velar plosive
l	[l] alveolar lateral
m	[m] bilabial nasal
n	[n] alveolar nasal
h	[h] voiceless glottal fricative
w	[w] voiced labialized approximant
y	[j] palatal approximant
[a]	[a] low, front, lax, unrounded
[a:]	[a:] low, front, tense, unrounded

[u]	[u] high, back, lax, rounded
[u:]	[u:] high, back, tense, rounded
[i]	[i] high, front, lax, unrounded
[i:]	[i:] high, front, tense, unrounded
[aw]	[aw] diphthong
[ay]	[ay] diphthong

Nominal competition in present-day English affixation: zero-affixation vs. *-ness* with the semantic category STATIVE

Cristina Lara-Clares, University of Granada
Paul Thompson, University of Birmingham

Word-formation processes may compete for the same domain of application in a given paradigm. This paper considers affixal competition (including zero) within the same derivational paradigm, specifically -ness *vs. zero-affixation for the expression of the semantic category* STATIVE*, as in* alert *vs.* alertness. *The results are compared with those obtained previously in Lara-Clares (2017) for the competition between* -ation *and zero-affixation for the expression of the semantic category* ACTION.

The data are a sample extracted from the complete frequency list of the British National Corpus (BNC) further enlarged with data from the Oxford English Dictionary (OED). For descriptive accuracy, the analysis of the sample separates lemmas into senses according to the lexicographic information available in the OED. This is in turn categorised for each affix in terms of the semantic classification of Bagasheva (2017). Potential cases of competition are screened based on a number of morphological constraints and thus researched: i) by entry, ii) by sense, and iii) by specific pairs or groups, both within entry and within sense.

Competing forms of each word class are then ranked by their prevalence, measured in two ways: i) in terms of their C-value (Fernández-Domínguez 2017), and ii) in terms of the distribution and dispersion of each form across the sample categories used by the BNC.

Keywords: *competition*, *nominalisation*, *-ness*, *zero-affixation*, STATIVE

1. Introduction

The borrowing of forms and, thus, affixes, that took place from the 17th century plays a central role in the development of the English language and a consequence of this is the emergence of both native and non-native competing (rival) affixes (Riddle 1985: 452–455; Lieber 2004: 44; Bauer 2009: 189; Kaunisto 2009; Díaz-Negrillo 2017: 119). When affixes acquire new senses, competition also emerges, such that more than one affix may share the same meaning[1] and potentially also the same domain of application; "expansion into new "territory" creates new potential for formation and thus enhances productivity, which leads to further expansion, which leads to further synonymy and potential competitors." (Pounder 2000: 669; see also Lieber 2004: 115). The underlying assumption here is that language tends to an avoidance of synonymy, as has been argued in the previous literature (see §2).

Morphological competition became a topic of research in itself after the publication of Aronoff's (1976) monograph, where he discussed the competition between *-ity* and *-ness*, thus laying the foundations for subsequent studies on the resolution of competition. This topic has received increasing attention ever since, as evidenced by the number of references dealing with it (Kjellmer 1984; Riddle 1985; Kastovsky 1986; Plag 1999; McWhinney et al. 2014, among others), as well as by the fact that the *17th International Morphology Meeting* in Vienna (2016) was dedicated to the discussion of competition in morphology. These studies have benefited

---

[1] "If we can ascribe meanings to affixes themselves, it appears that one affix can bear several meanings, while, conversely, one meaning can be shared by several affixes." (Lloyd 2011: 5).

greatly from research on morphological productivity (see, e.g., Plag 1999; Bauer 2001) and on semantics (see, e.g., Lieber 2004) and serve as a way to understand to what extent both linguistic economy and transparency of expression are at work in situations of morphological competition.

The analysis of competing forms in this paper relies on corpus and dictionary data. Corpora allow for the extraction of frequencies and also context of use, whereas dictionary data provide information regarding the different meanings of the forms under study and the dates of earliest and latest attestation. Specifically, BNC (Davies 2004–), data was used in an initial stage for sample extraction and in a later stage for attestation of forms in present competition (identified using the OED), for a manual sense classification of concordances and for frequencies' analysis. The OED was used both to enlarge the original BNC sample by looking for all potential competitors of the selected competing pattern and for lexicographic data (dates of attestation, senses, etc.).

An overview of the competition between nominalising suffixes is presented here, followed by a study on the competition between *-ness* and zero-affixation[2] for the expression of the semantic category STATIVE. The results are compared depending on whether frequencies are taken by entry, i.e., as given by the BNC, or by sense, i.e., after manual sense classification of concordances in order to put to test the relevance of sense separations. Finally, the results are compared with previous research for identification of patterns in the resolution of competition.

This paper addresses the following questions:

- RQ1: Is there any historical trend in the resolution of the competition between *-ness* and zero-affixation for the expression of the semantic category STATIVE? If so, is this trend confirmed with present-day data?
- RQ2: If a resolution of competition between the affixes is expected, will zero-affixation or *-ness* come out of use or will one or both affixes become semantically specialised?
- RQ3: Does each cluster of competition follow the general trend of resolution that is found for the pattern as a whole for these competing affixes?

This paper is structured as follows: the next section (§2) provides an overview of previous research on competition. The method followed for data extraction and analysis is then described (§3). §4 presents the results, first from a diachronic point of view (§4.1) and then in present-day English (§4.2). §5 draws conclusions from the analysis and compares results with a previous study.

---

[2] *Zero-affixation* is used here for terminological convenience.

2. Morphological competition in English affixation: an overview

Several definitions of competition in derivational morphology have been put forward in the last 40 years. They have key points in common, but there is still a certain level of variation among them. There is a general agreement that competing forms have to share the same base ("based on the same stem" in van Marle 1984: 178; "correlated to the same lexical base" in Fradin 2016) and be derived with different affixes (Plag 2000: 2; Fradin 2016). The third characteristic competitors have to possess is a relation of synonymy, although the degree of synonymy required is not the same for each author. Plag (2000: 2), Pounder (2000: 669) and Koehl (2015: 56), among others, claim that there is a need for competitors to have the *same* meaning, whereas other authors are more flexible in that respect: for Hoesktra & Versloot (2016), competitors should have "*roughly* the same semantic content" (emphasis added). Similarly, van Marle (1984: 178) claims that they should be "semantically similar" and Palacios (2013: 46) holds that they should "overlap in the expression the same or a very similar meaning [sic]." Also, competitors have to have the same distribution (Fradin 2016) or operate in the same domain and, thus, they occupy the same slot in a derivational paradigm (see Pounder 2000: 669; Bauer et al. 2013: 568).

This paper considers that, for competition to take place, forms have to:

i)   share the same base,
ii)   be derived with different affixes,
iii)   take affixes that express the same semantic category(s),
iv)   operate in the same domain, and
v)   no constraints (e.g., phonological, morphological) may apply.

Competition can be between individual words or between word-formation patterns. The latter seems to be dependent on the former: in Bauer's words (2009: 181), "if sufficient doublets are formed and word-formation *x* wins out in the majority of cases, then word-formation *x* will become the dominant process." Competition is expected to be resolved within a period of time, but this resolution is justified in different ways by different authors: some claim that it should be expected due to the principle of linguistic economy, according to which "a linguistic system will avoid having two forms for the same purpose" (Bauer et al. 2010: 1). Other authors claim that avoidance of synonymy is a key driving force behind competition (Lindsay & Aronoff 2013), and some equate competition in natural languages with Gause's (1934) struggle for existence, where "[t]he fate of the less efficient species is local extinction." (Aronoff 2016: 39; see also MacWhinney et al. 2014: 367). Be it as it may, competition is expected to be resolved after an indefinite period of time.

The resolution of competition can result in various outcomes:

i)   one competitor prevails over another, which eventually comes out of use,
ii)   each competitor finds a niche or domain of application, be it phonological (e.g., complementary distribution), stylistic (e.g., formal or informal contexts of use), or of some other kind,
iii)   both forms disappear (if no naming need or if another form fills that slot), or
iv)   both forms coexist for a period of time (e.g., negative prefixes for 700 years, according to Bauer (2009: 193)).

However, Pounder (2000: 322) holds that competition needs *not* be resolved, because "equivalent formations" may coexist and "there does not […] appear to be any competition at the lexical level that would imply a low tolerance for synonymy." We would then expect forms to exist peacefully alongside each other as long as there is not a more frequent or productive operation (Pounder 2000: 322, 669–672).

The relation between competition and productivity is undeniable[3] and the analysis of the productivity of word-formation patterns is essential in any study of competition. However, there is, to the best of our knowledge, only one measure of productivity that is specifically designed for this field or research (see Fernández-Domínguez 2017 for a review of competition and productivity and for details of the measure). The productivity of word-formation patterns is dependent on the number of constraints applying and, thus, the way competition resolves will also be dependent on them.

Constraints can be of different kinds: etymological, morphological, phonological, pragmatic and semantic (for an overview of previous literature on constraints on competition, see Díaz-Negrillo 2017 and Lara-Clares 2017). Etymological constraints are typically related to the origin of the base (e.g., Germanic vs. Latinate, as in *-ity* and *-ness*, see Bauer et al. 2013: 248; Arndt-Lappe 2014: 498, 501), while morphological constraints relate to the affixation already contained in the base of the derivative (e.g., *-ic* vs. *-ical*, where *-ical* is preferred over *-ic* in bases ending in *-olog*, see Lindsay 2012: 193). Phonological constraints are of especial relevance when they lead to the complementary distribution of affixes, as is the case of *-ify* and *-ize* (Plag 1999: 197, 228; see also Plag 2000: 10). Pragmatic or stylistic differences may lead to a specialisation of affixes and, thus, the resolution of competition. Looking at register distribution may prove to be useful in this respect, as in Guz (2009), where *-ness* is shown to prevail in more informal contexts than *-ity*. Finally, semantics can constrain the creation of derivatives, as with deadjectival adverbs, where it is reported that there is a tendency for dynamic adjectives to be more liable to form *-ly* adverbs than stative ones (Kjellmer 1984). The role other factors, such as blocking[4] or analogy[5], play in the resolution of competition is still under debate.

This paper aims at an overview on the competition between nominalising affixes to see how competition may be resolved (or not) in a case study. This considers the constraints that may apply as well as the productivity of the affixes.

3. Method

The sample used for the present study was extracted in two stages: a first sample (henceforth, Sample1) was extracted from the BNC, which was chosen over the *Corpus of Contemporary American English* (COCA; Davies 2008–) for its comparatively more fine-grained sample classification (for a thorough comparison of both corpora and for an analysis of morphological competition, see Fernández-Domínguez 2017). In a second stage, the OED was used to enlarge the sample using lexicographic data (henceforth, Sample2).

---

[3] "[I]t comes as no surprise that competition and productivity cross their paths, as productivity by nature compares two or more derivational rules." (Fernández-Domínguez 2017: 79).

[4] See Aronoff (1976) and Rainer (1988), but type-blocking is dismissed as a morphological mechanism by others (cf. Plag 2000; Giegerich 2001 and Bauer et al. 2013)

[5] See Arndt-Lappe (2014), although other authors tone down the relevance of analogy in word-formation (cf. Bauer 2001: 83).

3.1 *Extraction and analysis of Sample1*

The extraction of Sample1 was based on the complete BNC frequency list, which amounts to approximately 615,000 types. The list was sampled using *Scáthach* (Lara-Clares & Lara-Clares 2016), which allows automatic extraction of non-hyphenated affixed forms of a given word class. The base form of the units extracted was then analysed with OED data (see further below, Table 2). The affixes introduced in the software were taken from two sources: Quirk et al. (1985: 1540–1552, 1557) and Stockwell & Minkova (2001: 194–204). The decision to select only affixed forms stems from the fact that the same base appears more than once in the frequency list (e.g., the base *artificial* would be extracted from *artificialness*, *artificiality* and *artificialism*). Besides, the selection of affixed forms would potentially lead to the extraction of more cases of competition between affixes or between affixes and zero-affixation[6]. The exclusion of forms containing a hyphen served as a way to filter out hyphenated compounds (e.g., *self-awareness*), because compounding as a word-formation process is outside the scope of this paper. Other types of compounds (e.g., *minesweeper*) as well as any unwanted element, such as typos, foreign forms or entries containing symbols or numbers (e.g., *d[sep]amor*, *aimee*), were discarded in a later stage, in which every unit in the sample was analysed manually.

Sample1 was extracted and analysed in three stages for manageability reasons. The size of the sample for each stage was calculated using a tool for the calculation of the sample size of a population (Raosoft Inc. 2014). Both nominal and verbal forms were extracted in each stage, amounting to 1,147 nouns and 1,117 verbs (Table 1). The word class selected for further analysis was 'noun', due to the variety of patterns of competition found in this word class. The initial *nominal* sample was later enlarged as described in §3.2.

Table 1: Size of first sample per stage and total number of forms extracted for each word class

Word class	Stage 1	Stage 2	Stage 3	Total
Noun	384	382	381	1,147
Verb	377	370	370	1,117

Forms extracted from Sample1 (and, later, also Sample2) were analysed with data from the OED and classified using a template (Table 2). Competing forms appear in the first column, followed by the competing sense in the OED and the word-class of their base. The meaning of the competitors is given in the fourth column: first the semantic category, following Bagasheva's (2017) classification, and the definition in the OED. The fifth column provides the senses extracted from the OED, classified into in use, obsolete/dated, dialectal and register/domain and followed by the absolute frequencies of the BNC. Finally, there is the timeline of each form: first, the date of earliest attestation and then, when applicable, the date of latest attestation (†).

---

[6] Forms were considered to be zero-affixed when there is another attested word which is formally identical but of a different word class. Zero-affixed forms did not appear in the frequency list sample but they were later identified using the OED.

Table 2: Example of a competing cluster in the template with the base *representative*

Competing forms	Sense	Base w-class	Meaning		Senses (OED)				Freq.	Timeline	
			Sem. category	Def. (OED)	In use	Obs./ Dat.	Dial.	Reg./ Dom.	BNC	Earliest	†
representativeness	1	Adj	STATIVE	quality, state, or condition of being representative	1				49	1664	
representativeship	1	Adj	STATIVE	state or condition of being representative		1			0	1692	2009
representativity	1	Adj	STATIVE	representativeness	1				1	1901	

Competitors were identified using the list of entries on the right-hand side of the OED web. For example, in the BNC sample the derivative *representativeness* was extracted. It was then searched in the OED, where potential competitors could be identified using the entries list (Figure 1). In this example, *representativeship* and *representativity* as nouns have a sense that is apparently synonymous to that of *representativeness* and, thus, they are a priori identified as competitors.

Figure 1: Caption of the OED web entries list

For forms to be considered competitors and, thus, part of a cluster, they had to comply with the characteristics described above (§2). If we take as an example the cluster from Table 2, competitors had to:

i)   share the same base (*representative*[Adj]),
ii)   be derived with different affixes (*-ness*, *-ship* and *-ity*),
iii)   take affixes that express the same semantic category(s) (STATIVE),
iv)   operate in the same domain (they are not classified as dialectal or pertaining to any particular register or domain in the OED),
v)   and no constraints (e.g., phonological, morphological) may apply (a priori, inasmuch as the forms are attested, it is expected that no constraints apply, but that is verified in a later stage).

The semantic category of the affixes was determined based on the definition of the OED and later tested in a manual classification of BNC concordances (see further below). In the example, both *representativeness* and *representativeship* included the term *state* in the definition, and the definition for *representativity* leads to *representativeness*, which has also been classified as STATIVE.

This paper focuses on present-day competition, so only forms attested in the BNC were selected for further analysis. In the example in Table 2, *representativeness* and *representativity* were thus selected, and *representativeship* was discarded. Note that, in this example, lexicographic and corpus data agree in that *representativeship* is no longer in use. However, that is not always the case: in *warm*[N], the sense corresponding to the semantic category STATIVE is classified as rare in OED2[7] (sense 1) but a manual semantic classification of concordances of the BNC showed that that sense was attested in two out of three concordances.

3.2 *Extraction and analysis of Sample2*

The second sample was extracted after analysing and classifying the initial sample. A specific group of affixes competing for a particular sense was chosen from all the potential patterns[8] for analysis. For this paper, the semantic category chosen was STATIVE, and the affixes that were found in competition for the expression of this semantic category were *-ety*, *-ity*, *-ness*, and zero-affixation. More clusters of these patterns of competition were searched for using the OED's advanced search facility. This allowed for selection of entries from the dictionary according to their language of origin (English), to the affix (*ness, *ety and *ity, in three separate searches) and to the keywords used in their definition (here, *state* or *condition*). Five entries ending in *-ety*, 545 ending in *-ity*, and 1,626 entries ending in *-ness* were analysed in search for additional competing clusters of the said pattern, with the aim of getting as complete an image as possible of this pattern of competition. Besides, using the OED as a source for the extraction of data allows the identification of competing clusters over time, independently of whether forms are attested in contemporary corpora, or not. This resulted in a total of 437 competing clusters (see Table 3).

Table 3: Size of the second sample and number of competing clusters by affix

	-ety	*-ity*	*-ness*	Total
Occurrences	5	545	1,626	2,176
Competing clusters	1	140	296	437

After classification of both samples using the template shown in Table 2, the competing clusters for analysis were selected, namely, in this paper, the competition between zero-affixation and *-ness* for the expression of the semantic category STATIVE.

Both affixes are productive in present-day according to the specialised literature: zero-affixation is an "extremely productive process" (Plag 1999: 219) and *-ness* is reported to be the default resource for derivation of abstract nouns from non-verbal categories (Bauer et al. 2013: 246). Regarding the constraints applying to each process, the suffix *-ness* "is not sensitive to the phonological or prosodic structure of its bases, nor does it affect either the

---

[7] The version of the OED is specified where necessary because some entries have not been updated in more than a century and, thus, their classification as, e.g., in use or obsolete, is to be taken with caution.

[8] See Fernández-Alcaina (2017: 168) for a definition of *cluster* and *pattern*.

segmental phonology or stress patterns of the bases it attaches to." (Bauer et al. 2013: 248). In contrast, restrictions on noun to verb conversion have been found, e.g., monosuffix constraint (Don 2005), but this does not play a role here, because all the bases attested as in present competition are simple. Phonological restrictions on adjective to noun conversion, finally, do not seem "promising […] on cases of conversion involving adjectives." (Lohmann 2016: 229).

Nine clusters of this pattern of competition pattern were found in present competition in our data, amounting to a total of 7,898 concordances. All the concordances of the competing forms were semantically classified because, even though the forms under study are polysemous, the focus is only on one of their senses (STATIVE). The semantic analysis was done manually; in this case, each concordance was classified into two main categories:

i)   STATIVE, and
ii) non-STATIVE, where the keyword of the concordance:
     a.   is of a non-competing semantic category,
     b.   is of a different word class (adjective, verb, adverb), or
     c.   is ambiguous regarding meaning.

The semantic classification proved to be particularly complex either because some concordances can be interpreted in several ways or because slight differences in meaning could be identified within the semantic category under study (see Riddle 1985; Aronoff & Cho 2001; Baeskow 2012 and Díaz-Negrillo 2017 for the difficulty of applying general semantic definitions in order to capture differences in meaning between affixes).

The category STATIVE is described by Bagasheva (2017: 56) as a "particular condition of being, be in a state" (e.g., *sadness*). This category comprises both temporal states, i.e., states that remain for a period of time, and permanent states, which have been classified as quality nouns (*nomina qualitatis*) elsewhere (Rainer 2015; see also Luschützky 2015). The difficulty of applying this distinction in a classification of concordances, however, lies in the fact that many suffixes can be found in both categories, as in the case of the suffix *-ness* (e.g., *cleverness* as a permanent quality, and *drunkenness* as a temporal state) (Rainer 2015: 1269–1271). What is more, in the sample, competing derivatives showed both readings, so a decision was taken to classify as STATIVE in this paper only the concordances which reflect a temporal or non-permanent sense (1), as opposed to a permanent one (2):

(1)   *The only dealers who remain in their jobs are those who can sustain **alertness** during their long working hours* (EUU)
(2)   *Occupation is unquestionably one of the most important factors in preserving mental **alertness** and bodily health* (CKP)

The classification allowed also for an 'ambiguous' tag for those concordances in which no clear meaning could be captured from context (3).

(3)   *Very often, as we've mentioned already in this programme, there's contrast between **alertness**, brightness, whatever you call it, with words in speech and the disability in writing it down* (KRH)

Besides, some concordances were classified as non-STATIVE because they were part of a set phrase (4) or because they were (part of) a proper name (5), e.g.,

(4)   *Or do you think he's got the **hots** for this fabulous brown-eyed career woman he's acquainted with in London?* (GV8)

(5)   *Eliot's reading of Heart of **Darkness** whose Buddha-like clerk, Marlow, saw London as 'one of the dark places of the earth' further blended savage and 'sepulchral city'* (A6B)

The sample thus obtained was further analysed in later stages: the C value of the forms in each cluster was calculated in order to quantify the probability that a form would prevail over another (§4.2.1); the register distribution of forms was explored to look for hints as to a possible specialisation in their usage (§4.2.2) and, finally, the dispersion of the register distribution was computed to test whether competitors are or are not attested evenly across the registers, modes and domains of the BNC (§4.2.3). In sum, BNC data was used in the initial stage for sample extraction and in a later stage for attestation of forms in present competition (identified using the OED) and for frequencies' analysis. The OED was used both to enlarge the original BNC sample by looking for all potential competitors of the selected competing pattern and for lexicographic data (dates of attestation, senses, etc.).

4. Competition of nominalising suffixes

This section first presents an overview of the patterns of competition found in the samples and, more specifically, of the competition between zero-affixation and *-ness* suffixation for the semantic category STATIVE from a diachronic point of view (§4.1) and then looks at the present competition of the selected forms (§4.2).

4.1 *Overview*

After analysis of the samples, 10 senses were found to be in competition for nominal word formation according to the abovementioned conditions (§2 and §3.1). From these senses, STATIVE was chosen for further analysis in this study, and approximately 20 affixes were attested as competing for that sense in the OED. The suffix *-ness*, in competition with nearly 20 affixes, was selected for further analysis. Diachronically, the competition between *-ness* and *-ity* is the most widely attested one in this data (171 clusters), followed by *-ness* and zero-affixation (27 clusters), the latter being the pattern selected for this study (see Table 4).

Table 4: Patterns of competition of *-ness* with the semantic category STATIVE, the number of competing clusters and the total absolute frequencies of the forms attested in the BNC for each pattern (diachronically)

Patterns		N clusters	Abs. freq. BNC
-ness vs.	*-acy*	8	296
	-ance/-ence	3	101
	-(at)ion	22	4,086
	-dom	5	9,736
	-ety	3	378
	-hood/-head	25	2,964
	-ing	2	5,592
	-ism	7	27
	-itude	1	24
	-ity	171	58,389
	-ment	3	84
	-ry	4	153
	-ship	3	138
	-th	3	44
	-ty	3	8,461
	-ure	2	48
	-y	2	46
	zero-affixation	27	10,535

The dates of earliest and latest attestation in the OED of the competitors under study suggest different resolutions (Figure 2). Overall, it seems that the competition between zero-affixation and *-ness* tends to be resolved towards the decay of the former (e.g., *cool*, *glad*, *late*). However, that is not always the case: for some bases (e.g., *altogether*, *low*, *dry*), both affixes seem to remain in competition, because they are classified as in use in the OED; for others, forms do not actually co-exist with the sense STATIVE at any point in time: *laxative*$_N$ decays in 1527 (OED2) and *laxativeness* is attested earliest in 1610 (OED2), *watertight*$_N$ is only attested once in the OED in 1539 (OED3) and *watertightness*'s earliest attestation is in 1826 (OED3). In one case, zero-affixation wins out over *-ness*: *ripe*$_N$ decays in 1500 (OED3) while *ripeness* stays in use until present day (OED3) and, in another, zero-affixation wins shortly but then it also disappears (*warmness* decays in 1681 and *warm*$_N$ in 1839, both in OED2). As simple as this seems to be, the actual image is more intricate (Figure 3).
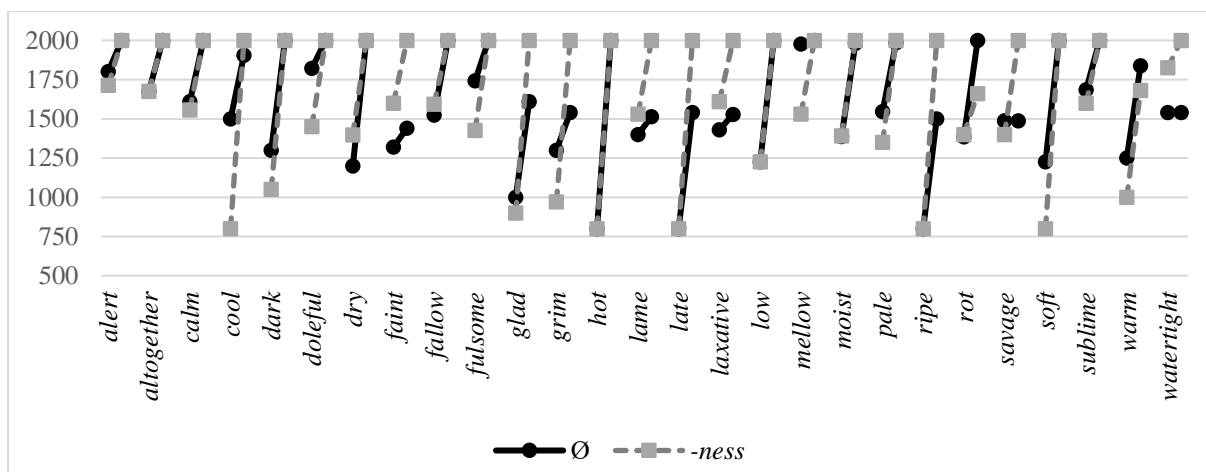
Figure 2: Timeline of the earliest and latest attestation of the competitors zero-affixation (continuous black line) and *-ness* (broken grey line) for the expression of STATIVE (OED)

Figure 3 shows that, quite often, there are more than two members in a cluster of competition. Five bases (out of 27) have at least three competing forms within the cluster. In *cool*~Adj~, the forms derived with *-ness* and *-th* remain in competition because the competing sense is classified as in use in OED3, whereas *cool*~N~ came out of use in 1905 (OED3). In *soft*~Adj~, the *-head* derivative is the one first attested the latest (1350, OED3) but it is also the only form to come out of use (1500), whereas *soft*~N~ and *softness* remain in use and, thus, in competition (both in OED3). The competition for the bases *glad*~Adj~ and *moist*~Adj~ seem to be resolved. In *glad*~Adj~, the *-ship* and zero-affixed forms were attested latest in 1597 and 1609 (both in OED2), respectively, and *gladness* prevails because it remains in use (OED2). Regarding *moist*~Adj~, the three forms are earliest attested within an 11-year period, but *moisture* is the first to disappear (1912, OED3), followed by *moist*~N~ (1981, OED3); the *-ness* form prevails, as it is the only one within the cluster that remains in use in present day according to lexicographic data (OED3). Finally, four competitors were found for the base *savage*~Adj~, where *-ism* and *-ness* remain in competition according to the OED3 data.
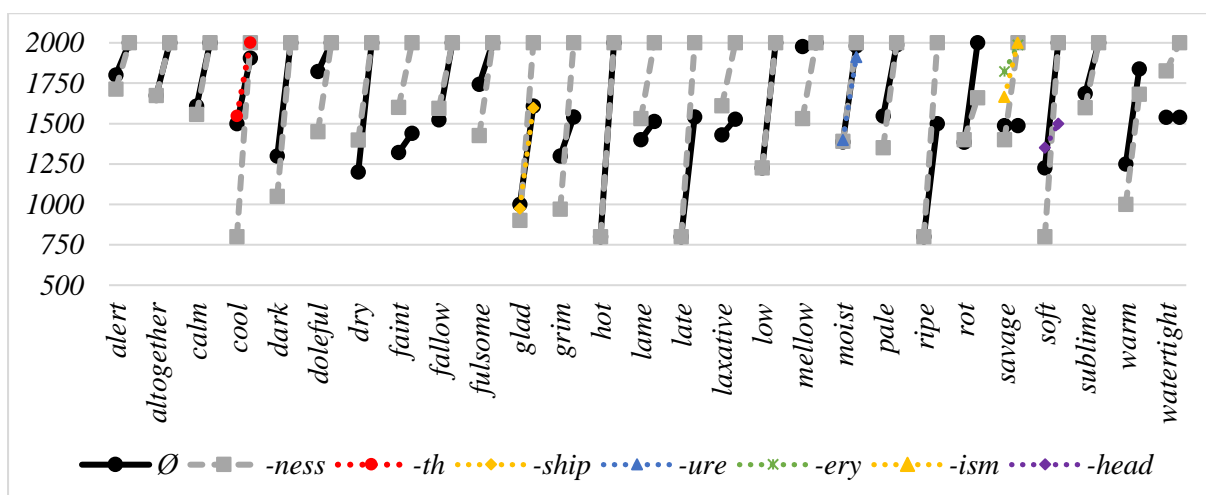

Figure 3: Timeline of the earliest and latest attestation of the competing nominal forms for the expression of STATIVE (OED)

Although lexicographic data of this kind should be taken with caution, it is still a relevant source (Bauer 2001: 156–157; Bauer 2009: 178, 181–182; Kaunisto 2009: 78; Bauer et al. 2010: 3; Arndt-Lappe 2014: 518). These data suggest, in answer to the first research question posed in §1 (RQ1, "is there any historical trend in the resolution of the competition between -*ness* and zero-affixation for the expression of the semantic category STATIVE?"), that there is a historical trend towards the resolution of this pattern of competition that favours the formation of -*ness* nouns over zero-affixed nouns for the expression of the said semantic category.

4.2 *Present-day competition*

Despite the historical trend pointed out in the previous section (§4.1), there still seems to be some cases of present competition of the affixes under study. For this reason, clusters in which at least two forms were attested in the BNC were selected for further analysis. The competition pattern -*ness* vs. zero-affixation for the expression of the semantic category STATIVE in present-day English was selected for further analysis.

Table 5: Manual classification of BNC concordances of competitors (STATIVE and non-STATIVE)

	STATIVE	Non-STATIVE			
		Diff. sem. cat.	Ambiguous	Adjective	Verb
alertness	72	7	6	0	0
alert	34	116	9	0	33
coolness	82	30	21	0	0
cool	39	0	13	44	8
coolth	1	0	0	0	0
darkness	667	2,434	30	0	0
dark	268	1,689	44	1,905	1
dryness	100	16	7	0	0
dry	0	0	0	2	0
faintness	19	0	0	0	0
faint	7	3	1	20	1
hotness	11	1	4	0	0
hot	0	17	1	0	22
lowness	3	1	0	0	0
low	35	123	18	179	0
savageness	1	0	0	0	0
savage	0	242	0	64	0
savagery	82	56	9	3	0
warmness	2	0	1	0	0
warm	0	1	0	2	0

All the concordances of the competing forms within the selected cluster were manually classified, as described in §3.2. Table 5 shows the number of concordances of each competitor

that were classified as being STATIVE or non-STATIVE and, within non-STATIVE, as being of a different semantic category, ambiguous regarding meaning, an adjective or a verb. Overall, the majority of concordances were classified as being of a different semantic category (4,736), followed by adjectives (2,219), and concordances that had a STATIVE reading (1,423). There was semantic ambiguity in 164 concordances and 65 forms were actually verbs.

Competitors suffixed with -*ness* were mainly classified as being of a semantic category different from STATIVE (71%), whereas 27% of concordances did show a STATIVE reading. There was ambiguity in 2% of the concordances. Concordances of zero-affixed forms, in contrast, were classified as being of a different semantic category in a 44% of cases, and as STATIVE in only an 8%. The percentage of ambiguity was again a 2%. Strikingly, in a 45% of concordances the derivative was an adjective and in a 1% it was a verb. That is the reason why, even though zero-affixation was more frequent than -*ness* overall (4,180 zero, 3,562 -*ness*), after-sense-classification absolute frequencies of the semantic category STATIVE are higher for -*ness* suffixation (957 -*ness*, 383 zero).

4.2.1 *Index of Competition (C)*

Frequencies from the BNC (by entry) and after sense classification (by sense) were used for the calculation of the Index of Competition (C) (Fernández-Domínguez 2017), which quantifies the likelihood that a morpheme outlasts its competitors. The index is calculated using the following formula:

$$C = \frac{N/_{Nc}}{Vc},$$

where N is the token frequency of a competing form, Nc is the token frequency of all forms in the cluster and Vc is the number of forms in current competition.

The interpretation of the resulting value is checked against the *Reference C*, a figure which fluctuates from 0 to 1 depending on the number of competitors and which therefore varies across clusters. The measure C posits that the more units are in direct competition, the more challenging their individual survival will be. Under this assumption, the maximum possible result from C is 1, which happens when a cluster is made up of just one unit, i.e., other competitors have disappeared and the unit has succeeded in competition. Likewise, the more competitors there are in a cluster, the lower the Reference C because the mere presence of other units means the existence of rivals, and then each unit gets fewer chances of success. Thus, Reference C is 1 if there is one lexeme in the cluster (i.e. resolved competition); 0.5 if there are two competitors; 0.33 if there are three; 0.25 if there are four; etc. The advantage of this measure is hence that the competitive status of a form is not assessed through an isolated numerical value, but it is set in the context of the cluster in question through several variables. For instance, a C value of 0.24 may indicate complete dominance for a unit if the Reference C is 0.25 (i.e., there are four competitors in its cluster, so 0.25 is the maximum possible result), while 0.24 is a rather poor value for a Reference C of 0.5 (i.e., there are two competitors in its cluster, so 0.5 is the maximum possible result).

According to the resulting value of the computation of C using BNC frequencies, -*ness* is expected to prevail in just one cluster (*dry*), zero-affixation in seven (*alert*, *cool*, *dark*, *faint*, *hot*, *low*, *savage*) and both would remain in present competition in another (*warm*) (Table 6). Opposite patterns are found if frequencies after sense separation are taken: -*ness* suffixation is expected to prevail in 7 clusters (*alert*, *cool*, *dark*, *dry*, *faint*, *hot*, *warm*), zero-affixation in one

(*low*) and a third member of a cluster, *-ery* (*savage*), in another (Table 7). This difference evidences the need to manually classify the concordances of each form for research on competition.

Table 6: Computation of C value for *-ness* vs. zero-affixed derivatives by entry. A plus symbol (+) means that the form is expected to prevail over the others in the cluster

	N	Nc	Vc	C	Exp. prevalence	Reference C
alertness	86	279	2	0.1541		0.5
alert	193	279	2	0.3459	+	
coolness	134	291	3	0.1535		0.33
cool	156	291	3	0.1787	+	
coolth	1	291	3	0.0012		
darkness	3,132	6,271	2	0.2497		0.5
dark	3,139	6,271	2	0.2503	+	
dry	2	125	2	0.008		0.5
dryness	123	125	2	0.492	+	
faint	32	51	2	0.3137	+	0.5
faintness	19	51	2	0.1863		
hotness	16	56	2	0.1429		0.5
hot	40	56	2	0.3571	+	
lowness	5	361	2	0.0069		0.5
low	356	361	2	0.4931	+	
savageness	1	458	3	0.0007		0.33
savage	307	458	3	0.2234	+	
savagery	150	458	3	0.1091		
warmness	3	6	2	0.25	+	0.5
warm	3	6	2	0.25	+	

Results obtained using after-sense-separation frequencies provide an answer to the second question posed in RQ1: if there is a historical trend in the resolution of the competition of these affixes, is it confirmed with present-day data? (§1). The answer is positive, because the aforementioned frequencies are in line with the general trend found according to OED data (§4.1), and they also support the need to semantically classify concordances. This leads to the next research question: if a resolution of competition between the affixes is expected, will zero-affixation or *-ness* come out of use or will one or both affixes become semantically specialised? (RQ2, §1). This question will be answered considering the register distribution of these affixes in the next section.

Table 7: Computation of C value for -*ness* vs. zero-affixed derivatives by sense. A plus symbol (+) means that the form is expected to prevail over the others in the cluster

	N	Nc	Vc	C	Exp. prevalence	Reference C
alertness	72	106	2	0.3396	+	0.5
alert	34	106	2	0.1604		
coolness	82	122	3	0.2240	+	0.33
cool	39	122	3	0.1066		
coolth	1	122	3	0.0027		
darkness	659	927	2	0.3554	+	0.5
dark	268	927	2	0.1446		
dry	0	100	2	0		0.5
dryness	100	100	2	0.5	+	
faint	7	26	2	0.1346		0.5
faintness	19	26	2	0.3654	+	
hotness	11	11	2	0.5	+	0.5
hot	0	11	2	0		
lowness	1	34	2	0.0147		0.5
low	33	34	2	0.4853	+	
savageness	1	83	3	0.00402		0.33
savage	0	83	3	0		
savagery	82	83	3	0.3293	+	
warmness	2	2	2	0.5	+	0.5
warm	0	2	2	0		

4.2.2 *Register distribution*

The register distribution[9] of the competing affixes was looked at using BNC frequencies both by entry and by sense in order to see whether there is any specialisation (register, mode or domain) of either the word as a whole or of concordances classified as STATIVE.

In order for the distribution of forms in the BNC and after sense separation to be comparable, the latter were also classified and normalised using the registers, modes and domains from the BNC. It is important to highlight here that frequencies reflecting the semantic category under study are very low for most of the competitors, but they are evidence of the attestation of the sense under study in each corpus part. Also, low frequencies could either point to an increasing or decreasing use of the form, which could be the cause or the result of competition.

Normalised frequencies from the BNC (Figure 4) show that the use of both affixes is balanced for Spoken (13% -*ness*, 10% zero), so they are expected to remain in competition in

---

[9] Normalised frequencies, both by entry (BNC) and by sense (after sense separation) are presented as percentages in Figures 4–5 for easier comparison, as frequencies are always lower after sense separation. Percentages are calculated with respect to the frequency of the form in the BNC, e.g., *darkness* has a frequency of 23 in Spoken and a total frequency of 659, so it is attested in a 3.49% of cases in Spoken. Tags show the normalised frequency of each item with respect to the size of each corpus section as well as the percentage.

that mode. Regarding the Written mode, zero-affixation is used more frequently overall in all registers but *fiction*, where -*ness* has a normalised frequency of 134.58 and zero-affixation of 122.07. Their frequencies are very similar in *non-academic* and *miscellaneous*.

It should be noted that, in Figure 4, -*ness* seems to be more frequent than zero-affixation in *non-academic* and *miscellaneous* because data is presented in percentages, i.e., with respect to the total frequency of the word. However, it is actually zero-affixation that is slightly more frequent in both registers (23.94 and 22.31 in *non-academic*; 30.58 and 27.07 in *miscellaneous*).
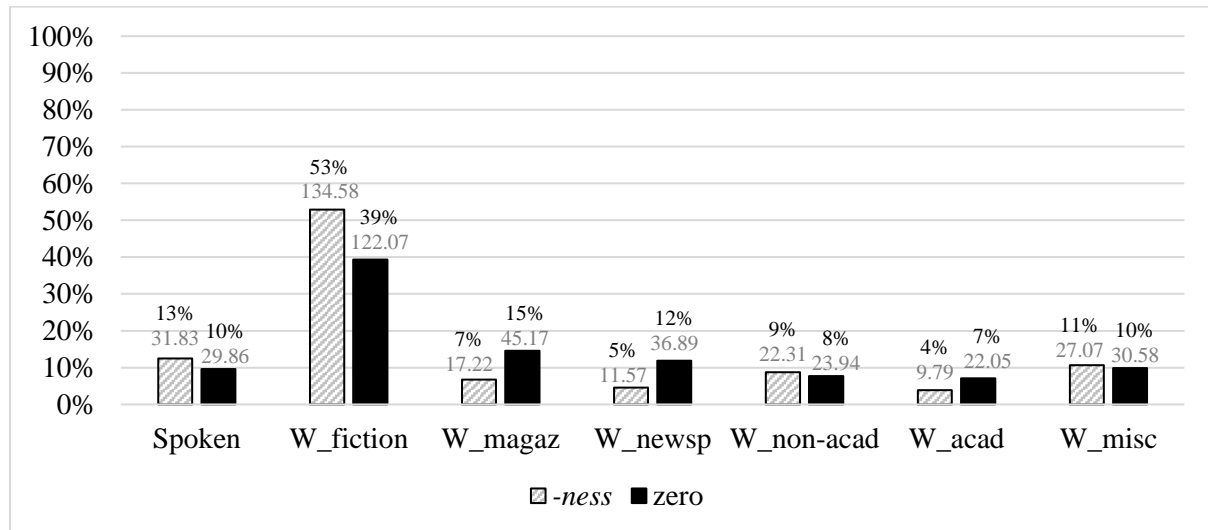

Figure 4: Register distribution of -*ness* (striped grey) and zero-affixed (black) forms with the semantic category STATIVE (normalised frequencies and percentages by entry)

However, if after-sense-separation normalised frequencies are taken (Figure 5), the difference of use between the two affixes becomes wider. Zero-affixation prevails in the Spoken mode (17%), whereas -*ness* suffixation prevails in the Written mode (96% in total) in all registers but *newspaper* (4.29 zero, 3.14 -*ness*). The gap becomes especially relevant for *academic,* where -*ness* suffixation sextuples the frequency of zero-affixation; or *non-academic* and *miscellaneous*, where it is quadrupled.

The data thus lead to the conclusion that there may be a difference between the affixes in question regarding the mode in which they are used. On the whole, -*ness* is expected to prevail over zero-affixation for the expression of the semantic category STATIVE, although there seems to be a process of mode specialisation underway, where zero-affixed forms would prevail in Spoken. If this were confirmed with a larger quantity of data, the competition between these affixes would be considered to be resolved, in that each of them would have a specific domain of application. Still, it might be the case that specific cases of competition do not follow the expected pattern, as claimed elsewhere for other patterns of competition (Lara-Clares 2017) (see §5). In order to test whether each cluster of competition follows the general trend of resolution found for the pattern as a whole for these competing affixes (RQ3, §1), the register distribution after sense separation of each pair of competitors was looked at, and it did show that not every cluster follows the expected trend.
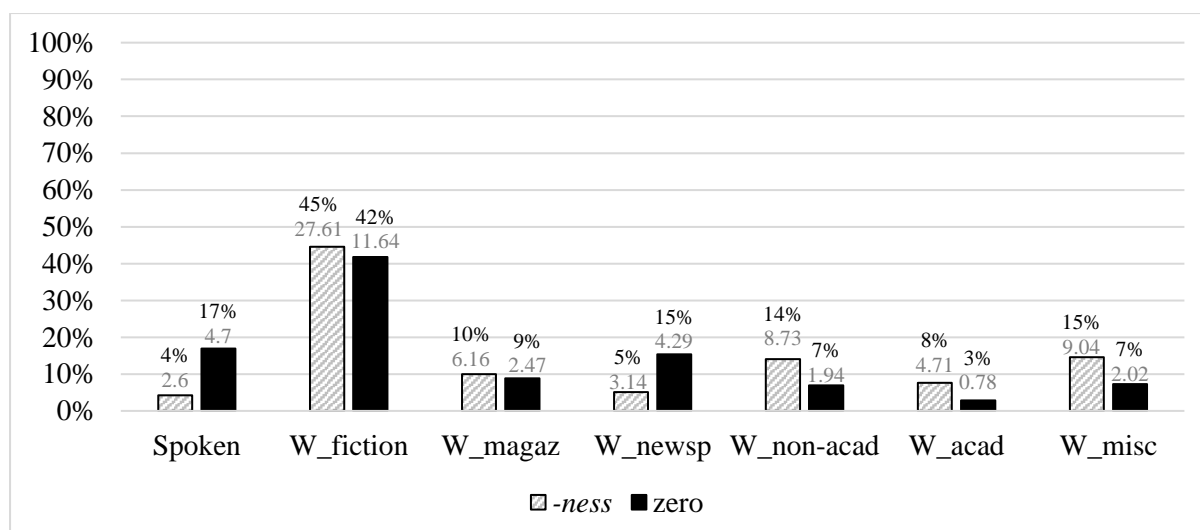
Figure 5: Register distribution of *-ness* (striped grey) and zero-affixed (black) affixed forms with the semantic category STATIVE (normalised frequencies and percentages by sense)

The cluster *darkness* vs. *dark* does follow the pattern described above: *dark* prevails in Spoken (16% zero, 3% *-ness*), and *darkness* does so in the Written mode overall, with the exception of the register *fiction*, where 333 concordances of *darkness* and 155 of *dark* were classified as STATIVE. Examples (6) and (7) are two concordances within the Spoken mode that belong to the same speaker, 'Danny', and which illustrate how similar the context of use of one and the other form with the semantic category STATIVE can be:

(6)   *The light will hit the moon and half the moon will be in **darkness** and half the moon will be in light yes?* (KPA)

(7)   *When you look at the moon half of it will be in light half of it will be in **dark** you will see a moon which looks like half a moon* (KPA)

In *alertness* vs. *alert* (Figure 6[10]), *coolness* vs. *cool* and *faintness* vs. *faint*, the same pattern as in *darkness* vs. *dark* is found, and the zero-affixed form prevails in *news* within the Written mode, as in Figure 5. Again, some instances were found where two forms were used in similar domains (*fiction* for the base *cool*, and *other* for the base *alert*) for the expression of the semantic category STATIVE, even though they were formed with different affixes: *-ness* in (8) and (10) and zero-affixation in (9) and (11):

(8)   *Slowly she chafed her arms with her fingers, aware now of the **coolness** of the late evening (HA6)*

(9)   *The sudden **cool** of evening made it advisable to keep the cooking fire alive (APU)*

---

[10] Absolute frequencies, both by entry (BNC) and by sense (after sense separation) are presented as percentages in Figures 6–8 for easier comparison. Tags show the absolute frequency of each item with respect to the size of each corpus section as well as the percentage. Note that there are two more registers here than in the BNC, as the manual classification of registers allowed for a more fine-grained classification of corpus sections (*W_essays* and *W_letters*).

(10)   *Whilst orthopaedically designed seating and sophisticated ventilation help keep the driver **alert** and in control, even on the most fatiguing of journeys* (CFT)

(11)   *But many find they help to maintain **alertness** and reduce headaches and eye strain from close work* (ED3)

However, other clusters of competition show a different distribution: in *lowness* vs. *low* (Figure 7), zero-affixation prevails in all registers (*lowness* is only attested once with the competing sense (12), in *fiction*, a register in which *low* is attested thrice (13)); in *faintness* vs. *faint* it is -*ness* suffixation that prevails overall (19 -*ness*, 7 zero), and so does in *dryness* vs. *dry*, *warmness* vs. *warm*, *savageness* vs. *savage*, and *hotness* vs. *hot*, where no zero-affixed form was classified as STATIVE. The register distribution of the latter, *hotness* vs. *hot* (Figure 8) is particularly striking in that all the concordances reflecting the semantic category under study are within only two registers of the Written mode: *fiction* and *other*.

(12)   *His depressions were an illness, not merely a feeling of **lowness** or irritability* (FU2)

(13)   *On an emotional and spiritual **'low'**, she had accepted, and been grateful for, his advice, his support, his protection* (JY2)
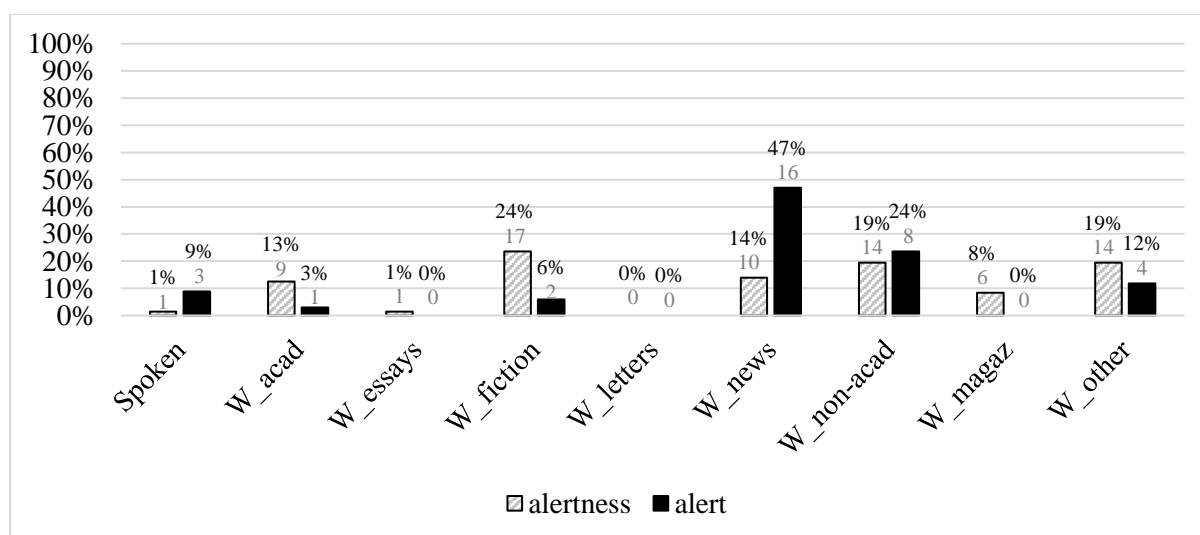

Figure 6: Register distribution of *alertness* and *alert* with the semantic category STATIVE (absolute frequencies and percentages by sense)
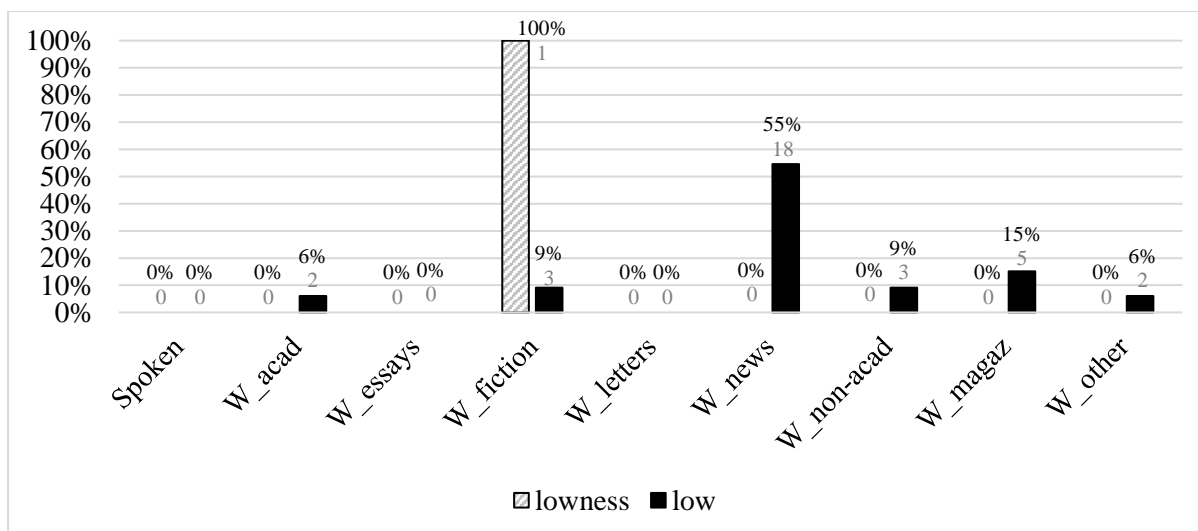
Figure 7: Register distribution of *lowness* and *low* with the semantic category STATIVE (absolute frequencies and percentages by sense)

An example of *hotness* with the semantic category STATIVE is (14):

(14)   *He could feel the **hotness** of inflamed flesh and the steady throbbing of a dozen tooth-sized bruises* (ACW)

The concordances of *hot*, in contrast, where either used as a verb (*to hot up*) (15) or as part of the set phrase *have the hots for* (4), and no instance of *hot* as STATIVE was thus found.

(15)   *And when the sparring **hots up**, the blow can sting* (CGE)

Figure 8: Register distribution of *hotness* and *hot* with the semantic category STATIVE (absolute frequencies and percentages by sense)
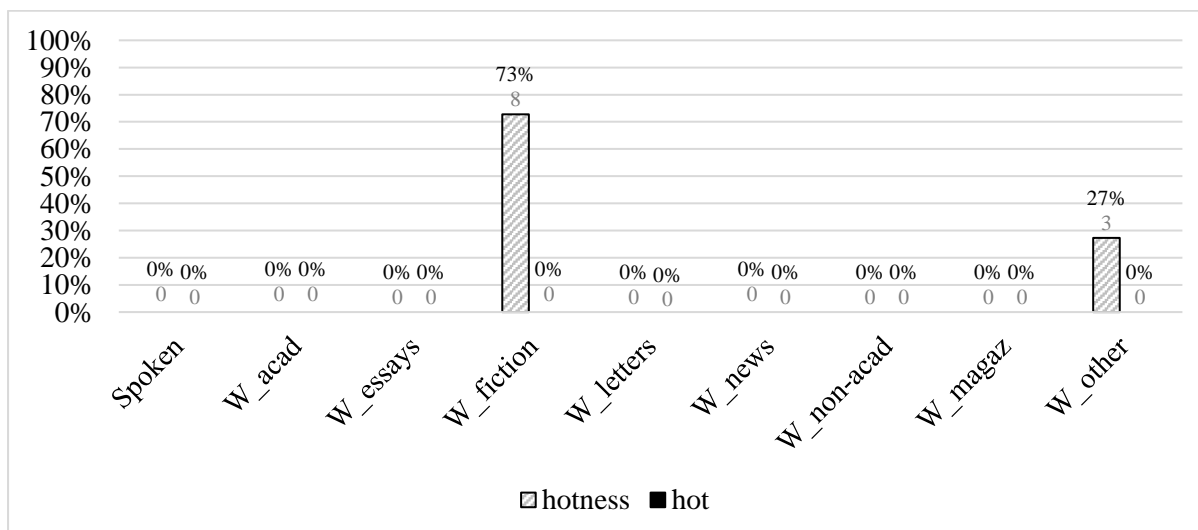
The data presented above suggest that, even though there seems to be a trend for the resolution of this pattern of competition, the analysis of particular clusters of competition yields disparate results.

4.2.3 *Dispersion*

The dispersion of the register distribution was calculated in order to see whether competitors are (or not) attested evenly across registers, modes and domains and, as a consequence, whether the forms may be specialised. The dispersion measure is calculated using the following formula (Gries 2008; Lijffijt & Gries 2012):

$$DPnorm = \frac{DP}{1 - \min(s)},$$

where *DP* is the dispersion and *min(s)* is the size of the smallest corpus part. A value close to zero indicates that the sample is evenly distributed, and if it is close to 1 the sample is unevenly distributed[11], even though results can fall outside that range (see Lijffijt & Gries 2012: 148). The *DPnorm* was calculated for each affix (zero-affixation and *-ness*), using both frequencies by entry and by sense (Table 8). The dispersion of these affixes with respect to the whole corpus shows that zero-affixed forms are more evenly distributed than *-ness* forms. Considering the dispersion of the semantic category STATIVE with respect to the frequency of each affix in the BNC, both affixes are very evenly distributed.

Table 8: *DPnorm* for *-ness* vs. zero-affixed derivatives with respect to the whole corpus and of the semantic category STATIVE with respect to the frequency of the form in the corpus

	-ness	zero-affixation
By entry (whole corpus)	0.445519616	0.292273107
By sense (word)	0.146011956	0.133894084

An even distribution a priori discards a specialisation in the use of competing affixes, which would theoretically lead to a resolution of the competition. Further research is needed, then, to fully test whether the differences in frequency between each of the registers, modes and domains are statistically significant or not.

5. Conclusion

Based on a sample of over 1,000 units from the BNC and over 2,000 from the OED, 437 clusters of competition were found. Diachronically, 27 clusters were identified for the competition between zero-affixation and *-ness* for the expression of the semantic category STATIVE. Of these, only nine clusters have at least two members attested in the BNC and are thus (a priori) in present competition, amounting to a total of 7,898 occurrences in the corpus.

---

[11] "[V]alues close to 0 indicate that *a* is distributed across the *n* corpus parts as one would expect given the sizes of the n corpus parts. By contrast, values close to 1 indicate that *a* is distributed across the *n* corpus parts exactly the opposite way one would expect given the sizes of the n corpus parts." (Gries 2008: 415; Lijffijt & Gries 2012: 147).

Opposite results are found when clusters are analysed by entry and by sense: competition seems to resolve in different ways according to data of each specific cluster of competition. Looking at the Index of Competition (C) by entry, zero-affixation is expected to prevail over *-ness* (zero-affixation would prevail in seven clusters out of nine), but an opposite tendency is found by sense (*-ness* would prevail in seven clusters). The data obtained after sense separation supports the trend found in the OED, in that *-ness* would prevail over zero-affixation for the expression of the semantic category STATIVE (RQ1). The register distribution of these affixes shows that, by entry, both affixes are used equally within the Spoken mode, but zero-affixation prevails overall in the Written mode, with the exception of *fiction*. However, there does seem to be a difference in the domain of use when looking at data by sense: zero-affixation prevails in the Spoken mode, whereas *-ness* prevails in the Written mode (RQ2). It is important to note, however, that the computation of the *DPnorm* shows an even distribution of competitors, both by entry and by sense. However, even though a general trend was found regarding the resolution of the competition between these affixes, results may vary if every cluster is looked at individually, e.g., no attestation of the semantic category STATIVE was found in the BNC for *hot* (RQ3).

A comparison of these results with those obtained in Lara-Clares (2017) for the competition of zero-affixation vs. *-ation* for the expression of the semantic category ACTION shows that one affix (here, zero) may win out when competing with one affix for one sense but it may also be in a process of decay or register specialisation when competing with another affix for a different sense. In the previous study, zero-affixation seemed to prevail overall, even though results by sense were not so conclusive (zero-affixation would prevail over *-ation* in four cluster and *-ation* in three) (Lara-Clares 2017: 224–225, 228–229). The distribution of the competing affixes, both by entry and by sense, showed that zero-affixation prevailed overall, and no hints were found that would point to a specialisation of the affixes (Lara-Clares 2017: 225–226, 230). This contrasts with the results from this paper, where zero-affixation is expected to become specialised (mainly used in the Spoken mode), even though the computation of the *DP* contrasts with that view and should be tested further.

The results presented in this paper focus on two ways of nominalising an adjectival base, that is, by affixation by *-ness* or by zero-affixation. However, the derivational paradigm of each base is much more complex and analysing it for a study of competition could give new insights to the results. Also, competitors are likely to be found in any study of derivational paradigms, as acknowledged by Pounder (2000: 670): "for bases in the common (lexico-semantic) domain […], lexical paradigms may well contain sets of synonymous alternates of equal status (except for the productivity differential)", and should thus not be ignored.

Acknowledgements

References

Arndt-Lappe, Sabine. 2014. Analogy in suffix rivalry: The case of English *-ity* and *-ness*. *English Language and Linguistics* 18(3). 497–548.

Aronoff, Mark. 1976. *Word formation in Generative Grammar*. Cambridge, MA: MIT Press.

Aronoff, Mark. 2016. Competition and the lexicon. In Elia, Annibale & Iacobini, Claudio & Voghera, Miriam (eds.), *Livelli di Analisi e Fenomeni de Interfaccia: Atti del XLVII Congresso Internazionale di studi della Società di Linguistica Italiana (SLI)*, 39–51. Roma: Bulzoni Editore.

Aronoff, Mark & Cho, Sungeun. 2001. The semantics of *-ship* suffixation. *Linguistic Inquiry* 32(1). 167–173.

Baeskow, Heike. 2012. *-Ness* and *-ity*: Phonological exponents of *n* or meaningful nominalizers of different adjectival domains? *Journal of English Linguistics* 40(1). 6–40.

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. 2001. *Morphological productivity*. Cambridge: Cambridge University Press.

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 & 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. 2013. *The Oxford guide to English morphology*. Oxford: Oxford University Press.

Davies, Mark. 2004–. *BYU-BNC* (Based on the *British National Corpus* by Oxford University Press). (http://corpus.byu.edu/bnc/) (Accessed 2018-07-25.)

Davies, Mark. 2008–. *The Corpus of Contemporary American English (COCA): 520 million words, 1990-present*. (http://corpus.byu.edu/coca/) (Accessed 2018-05-06.)

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. 2005. On conversion, relisting and zero-derivation: A comment on Rochelle Lieber: *English word-formation processes*. *SKASE Journal of Theoretical Linguistics* 2(2). 2–16.

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.

Fernández-Domínguez, Jesús. 2017. Methodological and procedural issues in the quantification of morphological competition. In Santana-Lario, Juan & Valera, Salvador (eds.), *Competing patterns in English affixation,* 67–117. 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.)

Gause, George F. 1934. *The Struggle for Existence*. Baltimore, MD: Williams & Wilkins.

Giegerich, Heinz J. 2001. Synonymy, blocking and the Elsewhere Condition: Lexical morphology and the speaker. *Transactions of the Philological Society* 99(1). 65–98.

Gries, Stefan Th. 2008. Dispersions and adjusted frequencies in corpora. *International Journal of Corpus Linguistics* 13(4). 403–437.

Guz, Wojciech. 2009. English affixal nominalizations across language registers. *Poznań Studies in Contemporary Linguistics* 45(4). 447–471.

Hoekstra, Eric & Versloot, Arjen. 2016. Frequency, similarity and the blocking principle. (Paper presented at the *16th International Morphology Meeting*, Vienna, 18–21 February 2016.)

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.

Kastovsky, Dieter. 1986. The problem of productivity in word formation. *Linguistics* 24. 585–600.

Kjellmer, Göran. 1984. Why *great*/*greatly* but not *big*/**bigly*? On the formation of English adverbs in *-ly*. *Studia Linguistica* 38(1). 1–19.

Koehl, Aurore. 2015. Deadjectival nominalizations: Suffix rivalry and synonymy. In Audring, Jenny & Koutsoukos, Nikos & Masini, Francesca & Raffaelli, Ida (eds.), *Proceedings of the 9th Mediterranean Morphology Meeting – Morphology and Semantics*, 54–63. University of Leiden, University of Patras, University of Bologna, University of Zagreb.

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 & Lara-Clares, Alicia. 2016. An online tool for big data sampling in research on competition in English word-formation. (Paper presented at the 4th meeting *Linguistics Beyond and Within* – International Linguistics Conference, Lublin, 20–21 October 2016.)

Lijffijt, Jefrey & Gries, Stefan Th. 2012. Correction to Stefan Th. Gries' "Dispersions and adjusted frequencies in corpora", International Journal of Corpus Linguistics, 13:4 (2008), 403–437. *International Journal of Corpus Linguistics* 17(1). 147–149.

Lieber, Rochelle. 2004. *Morphology and lexical semantics*. Cambridge: Cambridge University Press.

Lindsay, Mark. 2012. Rival suffixes: Synonymy, competition, and the emergence of productivity. In Ralli, Angela & Booij, Geert & Scalise, Sergio & Karasimos, Athanasios (eds.), *Proceedings of the 8th Mediterranean Morphology Meeting – Morphology and the architecture of hrammar*, 192–203. Patras: University of Patras.

Lindsay, Mark & Aronoff, Mark. 2013. Natural selection in self-organizing morphological systems. In Montermini, Fabio & Boyé, Gilles & Tseng, Jesse (eds.), *Morphology in Toulouse: Selected proceedings of Décembrettes 7*, 133–153. Munich: Lincom Europe.

Lloyd, Cynthia. 2011. *Semantics and word formation. The semantic development of five French suffixes in Middle English*. Bern: Peter Lang.

Lohmann, Arne. 2017. Phonological properties of word classes and directionality in conversion. *Word Structure* 10(2). 204–234.

Luschützky, Hans Christian. 2015. Status nouns. In Müller, Peter O. & Ohnheiser, Ingeborg & Olsen, Susan & Rainer, Franz (eds.), *An international handbook of the languages of Europe*, 1268–1284. Berlin: De Gruyter Mouton.

MacWhinney, Brian & Malchukov, Andrej & Moravcsik, Edith (eds.). 2014. *Competing motivations in grammar and usage*. Oxford: Oxford University Press.

Palacios, Amutio. 2013. Suffix competition in Old English word formation. *Revista electrónica de lingüística aplicada* 12. 45–62.

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

Plag, Ingo. 2000. On the mechanisms of morphological rivalry: a new look at competing verb-deriving affixes in English. In Reitz, Bernhard & Rieuwerts, Sigrid (eds.), *Anglistentag 1999 Mainz. Proceedings*, 63–76. Trier: Wissenschaftlicher Verlag Trier.

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

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

Quirk, Randolph & Greenbaum, Sidney & Leech, Geoffrey & Svartvik, Jan. 1985. *A Comprehensive grammar of the English language*. London: Longman.

Rainer, Franz. 1988. Towards a theory of blocking: the case of Italian and German quality nouns. In Booij, Geert & van Marle, Jaap (eds.), *Yearbook of morphology*, 155–185. Dordrecht: Foris.

Rainer, Franz. 2015. Quality nouns. In Müller, Peter O. & Ohnheiser, Ingeborg & Olsen, Susan & Rainer, Franz (eds.), *An international handbook of the languages of Europe*, 1268–1284. Berlin: De Gruyter Mouton.

Raosoft Inc. 2004. *Sample size calculator*. (http://www.raosoft.com/samplesize.html?nosurvey). (Accessed 2018-01-14).

Riddle, Elizabeth M. 1985. A historical perspective on the productivity of the suffixes *-ness* and *-ity*. In Fisiak, Jacek (ed.), *Historical semantics: Historical word-formation*, 435–461. Berlin: Mouton.

Stockwell, Robert & Minkova, Donka. 2001. *English wordsHistory and structure*. Cambridge: Cambridge University Press.

van Marle, Jaap. 1984. *On the paradigmatic dimension of morphological creativity*. Dordrecht: Foris.

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

Paul Thompson
Department of English Language and Linguistics
University of Birmingham
Edgbaston
Birmingham B15 2TT
United Kingdom
p.thompson@bham.ac.uk

Metonymy and meaning construction in Persian nominal compounds: A close-up on one-part metonymical combinations

Masoumeh Diyanati, University of Isfahan, Iran
Alexander Onysko, University of Klagenfurt, Austria

Metonymy has been discussed as one constitutive process among others in the formation of figurative compound meaning in previous research. An explicit focus on the qualitative types of metonymic processes involved in the construction of compound meaning, however, has so far not been pursued. The current study zooms in on metonymy as a constitutive phenomenon of compound meaning taking Persian compounds as a test case. An analysis of 166 Persian noun-noun compounds that contain one metonymical part is geared at answering the question of whether the type of metonymy affects the construction of meaning in the metonymical compounds. The analysis reveals that there are different degrees of metonymic complexity, based on the relation between the metonymical element and the referent. While in some cases, the metonymical constituent does not match with the referent, there are other cases where the metonymical named element represents the referent name in a subordinate relation or is expressed directly in the referent name. In addition, metonymical complexity depends on the prototypical nature of the metonymic association in the frame-based combination of the compound constituents. These two factors imply that a continuum can be postulated that ranges from simple metonymies to complex metonymies.

Keywords: *metonymy, compounds, Persian, meaning construction*

1. Introduction

Since the 1980s, metonymy and metaphor have been considered as key conceptual mechanisms in cognitive semantics as initially laid out in Lakoff and Johnson (1980). They define metonymy as a 'stand for' relationship in which one named entity is used to refer to another (1980: 36). Later, following further development in Cognitive Linguistics, the definition and formulation of metonymy became more diversified (see section 2). Surveys such as that conducted by Drożdż (2014) show that cognitive linguists have still not reached a complete consensus on the definition of metonymy even though there is general agreement on the essential role of metonymy for encoding meaning in language. Among various language forms that highlight the role of conceptual metonymy, the semantics of compound words has been a major interest. Some previous studies have discussed the immanent role of conceptual metonymy (and metaphor) in the meaning of compounds (see section 3). However, research so far has not highlighted the potential qualitative difference of metonymical relations in compound meaning construction. To shed some light on this semantic question, this study is based on a dataset of 166 Persian noun-noun compounds that contain one metonymical part. The analysis of the data aims at clarifying the kind of metonymical relation and the simplicity/complexity of meaning construction in compounds.

The paper first gives a brief overview of how conceptual metonymy is defined in cognitive linguistics (section 2). This is followed by a discussion on the representation of metonymical associations in the meaning of compounds (section 3). Section 4 outlines the methodology used for this study, and section 5 is devoted to the analysis of the selected Persian

metonymical compounds. Central to that will be the discussion of how metonymy relates to the simplicity/complexity of meaning construction in compound words.

2. Metonymy in cognitive linguistics

In the traditional view, metonymy was regarded as a figure of speech in rhetoric that leads to "a shift of a word meaning from the entity it stands for to a 'contiguous' entity" (Ullman 1957: 232; as stated in Croft 1993: 347). Following the advent of cognitive linguistics, metonymy went beyond purely stylistic and rhetorical considerations and turned into a central cognitive semantic process. Lakoff and Johnson (1980: 39) state that metonymy is another conceptual mechanism beyond metaphor that does not function exclusively at the level of words but "structures our thought, attitude, and actions" (Evans & Green 2006: 311). They define metonymy as a mechanism which allows human beings "to conceptualize one thing by means of its relation to something else" (Lakoff & Johnson 1980: 37, 39). Later, Lakoff (1987) revised the definition of metonymy by introducing the concept of Idealized Cognitive Models (ICMs). He states that, metonymy is a stand for relation between two elements of the same ICM or one of the elements of an ICM and the whole ICM (1987: 78–79).

As part of the development in Cognitive Linguistics, the discussion of metonymy has become diversified involving other cognitive concepts. Based on Langacker's (1987) ideas, Croft (1993, 2002) applies the notion of domain matrix and claims that metonymy leads to the foregrounding or highlighting of one domain within a domain matrix (2006: 321). Another widespread definition of metonymy has been proposed by Radden and Kövecses (1999). Using Lakoff's (1987) notion of ICMs, they describe metonymy as follows: "A cognitive process in which one conceptual entity, the vehicle, provides mental access to another conceptual entity, the target, within the same idealized cognitive model (ICM)" (ibid.: 21).

An alternative perspective of metonymy is suggested in Peirsman & Geeraerts (2006). In their prototype-based approach to metonymy, they take spatial part-whole relations as the prototypical core of the category since they show the strongest degree of contiguity. The degree of contiguity weakens for metonymical relations that range from containment and contact to adjacency without contact (2006: 309).

Panther and Thornburg (2007) provide yet another description. They characterize metonymy as a cognitive process where: 1) a source content provides access to a target content within one cognitive domain; 2) there is a contingent relation between the source and the target content; 3) the source content is backgrounded, and the target content is foregrounded; and 4) depending on the conceptual distance between the source and the target and the salience of the source, the metonymic link between the source and the target may be weak or strong (2007: 242).

Another important definition that needs to be mentioned is proposed by Barcelona (2011). Considering metonymy as a reference point phenomenon, he refines the definition of metonymy as "an asymmetric mapping of a conceptual domain, the source, onto another domain, the target" (2011: 52). Barcelona argues that the source and the target are both elements of the same functional domain which are related by a "pragmatic function" (ibid).

The short overview of some important definitions of metonymy shows that researchers mainly rely on the notions of domains and contiguity when defining the process of metonymy. This is also true for some other contributions to metonymy in the literature (cf. Benczes, Barcelona & Ruiz de Mendoza Ibáñez 2011; Kosecki 2007; Panther & Radden 1999). Among

these approaches, the prototype view proposed by Peirsman & Geeraerts (2006) adds a qualitative dimension to the notion of metonymy as, depending on the strength of contiguity, metonymical relations can be more or less prototypical, i.e. central to the phenomenon of metonymy. Our findings on different metonymical relations in the set of Persian compounds discussed in section 5 corroborate a prototypical view of metonymy, which helps to differentiate metonymical compounds on a cline of metonymical complexity. Before the data analysis and discussion, the next section will outline some previous research on metonymy in compounding to show the importance of that process in a semantic analysis of compounds.

3. Metonymy in compounding

Cognitive linguists argue that conceptual processes affect linguistic structures (Evans & Green 2006). Conceptual metonymy is one of the main mechanisms beyond conceptual metaphor that contributes to the semantics of compounds as previous research has shown. Geeraerts (2002) analyzed various interactions of metaphor and metonymy in the semantics of idioms and compounds. He states that on a possible continuum between metaphor and metonymy, at the one end lie "expressions that are fully metonymical" (2002: 449). Among his examples, *hanglip* is mentioned as a completely metonymical compound involving two metonymical extensions (BODY PART FOR PERSON and EFFECT FOR CAUSE) that make up its meaning of 'an unhappy, sulky, pouting person' (2002: 456–457).

Focusing on hyponymic compounds as the most common type of English compounds, Radden (2005) observes that in a hyponymic compound the modifier foregrounds one salient property of the category through metonymy. For example, in *wheelchair* the modifier focuses on the wheels. According to Radden (2005: 19), "compounds thus typically involve a PART FOR WHOLE metonymy". This implies that the morphological process of compounding would in itself be metonymic (also see Janda 2011), which has been cogently criticized in Brdar & Brdar-Szabó (2014).

Another attempt that has elucidated the role of metonymy, specifically in the semantics of selected English and Spanish bahuvrihi compounds is made by Barcelona (2008). He demonstrates that the exocentric nature of all bahuvrihi compounds is motivated by the metonymy of CHARACTERISTIC PROPERTY FOR CATEGORY. He also discusses three ways of conceptualizing the characteristic property mapped onto the target category, including literal (e.g. *humpback*), metonymical (e.g. *hard top*), and metaphorico-metonymical (e.g. *blockhead*) relations. More recently, Barcelona (2012) has also presented examples for the role of metonymy in grammar and discourse to illustrate the function of metonymy on three levels of the lexicon, namely under the lexicon (phonology and morphology), in the lexicon (lexical metonymies), and above the lexicon (grammar). Acknowledging the important role of metonymy in grammar, Barcelona (2012: 261) introduces compounding as one part of grammar that can be motivated by metonymy.

Benczes (2006a) provides a further cognitive semantic analysis of compounds that draws on metonymy and metaphor for explaining the creation of noun-noun compounds. She investigates the American neologism *freedom fries* as an exocentric compound in which both constituents are respectively affected by METHOD OF PRODUCTION FOR PRODUCT (fried in oil for potatoes) and DEFINING PROPERTY FOR CATEGORY (freedom for America) metonymies. In her major study on *Creative Compounding in English*, Benczes (2006b) devotes one chapter to different patterns in which conceptual metonymy can affect English noun-noun combinations.

She discusses five patterns, including metonymy-based modifier, metonymy-based profile determinant, metonymy-based modifier and profile determinant, metonymy-based compound as a whole, and metonymy-based relation between the two constituents of the compound. In addition, Benczes (2011) investigates the semantics of metonymical (and metaphorical) compounds to explore the role of domains in conceptual metonymy. She claims that multiple metonymies (and metaphors) can act upon the meaning of compounds, which chimes in with previous research such as Geeraerts's (2002) discussion on the interaction of metaphor and metonymy mentioned above.

Based on Radden & Kövecses's (1999) definition of metonymy, Kuczok (2007) analyzes a number of English noun-noun compounds which are motivated by one (e.g. *bear jam*) or two (e.g. *white-collar*) metonymies (and/or metaphors).

The effect of conceptual metonymy and metaphor in the semantics of novel English noun-noun compounds has also been shown in Onysko (2014, 2016). Carrying out an empirical study that investigates the process of figurative meaning interpretation, Onysko (2014) confirms that conceptual metonymy (besides conceptual metaphor) underlies various types of meaning associations. As far as metonymy is concerned, his data on meaning interpretations of novel compounds show that metonymy is the most frequent, i.e. basic, figurative process used by the participants (Onysko 2016).

In a detailed examination of the role of metonymy in word-formation, Brdar & Brdar-Szabó (2013) show that metonymy may affect the head or the modifier of compound words, prior or posterior to the compounding process. In other words, they believe that compounding as a concatenative process might be preceded or followed by conceptual metonymy. In a recent major study, Brdar (2017) shows that metonymy can affect the compound constituents and the overall meaning of the compound.

Altogether, previous research on the role of metonymy in compounding has confirmed the important function of metonymy (alongside conceptual metaphor) for the construction of meaning in compounds. The main lines of research so far have investigated different types of metonymies, their relation to the compound constituents and to the overall meaning of the compound, and their interaction with conceptual metaphors. Our study attempts to go a step further and explores the qualitative difference of metonymies and whether that can potentially be related to associative complexity in the meaning of compounds. Before we precede with that question, the data and methods of analysis will be briefly described in the next section.

4. Methodology

As in many languages of the world (see Lieber & Štekauer 2011) compounding is a major process of word formation in Persian. Compounding in Persian stands out from other languages as Persian allows for both right and left-headed compounds (see Kalbassi 1997; Shariat 2005). While the role of headedness in the construction of compound meaning is not the focus of this study, it plays a role when analyzing metonymies according to their location in the modifier or head constituent of the compound (see below).

In general, Persian offers a well-documented and not very frequently studied data set of compounds that serves as a good basis to explore the role of metonymy (or metaphor, cf. Torabian 2013). *Sokhan* (Anvari 2003), an eight volume Persian Monolingual Dictionary, serves as a database of metonymical compounds in this study. All noun-noun nominal compounds were extracted from the dictionary, which amounted to a total of 720 compounds.

As a next step, the possible effect of metonymy and/or metaphor in the meaning of the collected words was examined. This procedure yielded 347 compounds with no metaphor/metonymy, 166 compounds with one metonymical part, 79 items with one metaphorical constituent, 46 terms with one metaphorical and one metonymical constituent, 63 words in which both constituents are metonymical, and 19 compounds in which a metonymical mapping underlies the meaning of the compound as a whole. To zoom in on the qualitative difference of metonymical relations in the meaning construction of compounds, we limited our study to compounds which have just one metonymical component.

Due to the existence of right and left-headed compounds in Persian, the data set of one-part metonymical compounds comprises instances of metonymies in the head or the modifier regardless of the constituent's location in the compound. For example, *xan-ʔamu* (leader + paternal uncle), meaning 'the oldest and respected paternal uncle', is a right-headed compound whose modifier *xan* 'khan' involves the metonymy of PERSON FOR PROPERTY. A good example of a metonymical left-headed compound is *komak-ranande* (help + driver) meaning 'someone who helps the bus driver in intercity travel'. The head element of *komak-ranande*, i.e. *komak* 'help', is motivated by the metonymy of ACTION FOR AGENT. Another example of a metonymical left-headed compound is *ab-andʒir* (water + fig) meaning 'water with the flavor and color of fig', where the modifier, i.e. *andʒir* 'fig', stands metonymically for its properties (ENTITY FOR PROPERTIES). In the analysis below, the head constituent is marked in bold print to provide a visual aid for understanding the compound structure.

4. Data analysis

As mentioned earlier, this study aims to investigate the possible relationship between the quality of the metonymical relations and simplicity/complexity of meaning construction in Persian metonymical nominal compounds. In general, "meaning construction is an inferential process" (Radden, Köpcke, Berg & Siemund 2007: 10). In the case of metonymical units, meaning construction relates to the process of inferring the referent of the metonymical constituent. It seems that this inferential process is not completely the same in all metonymical compounds. Depending on the conceptual relation of the metonymical element and the referent, different degrees of complexity are evident in the inferential process. For the sake of illustration, this difference can be displayed on a scale of complexity. At the upper end of that scale, metonymical compounds show a higher degree of complexity when the metonymical constituent does not match with the referent of the compound. Consider the following examples:

(1)   *xan-ʔamu* (khan + paternal uncle) 'the oldest and respected paternal uncle'

(2)   *xan-daji* (khan + maternal uncle) 'the oldest and respected maternal uncle'

(3)   *xan-dadaʃ* (khan + brother) 'the oldest and respected brother'

(4)   *gol-mix* (flower + nail) 'a decorative nail'

(5)   ʃab-**bu** (night + smell) 'a kind of flower that smells at night'

(6)   **ab**-andʒir (water + fig) 'water with the flavor and color of fig',

(7)   **ab**-alu (water + plum) 'water with the flavor and color of plum'

(8)   **ab**-zereʃk (water + barberry) 'water with flavour and color of barberry'

(9)   **pa**-taɢ (foot + arch) 'the lowest part of the arch'

(10)   **tʃub**-parde (wood + curtain) 'rod where a curtain is hung up'

Examples (1) to (3) comprise three right-headed compounds which follow the same construction of *xan + x*. In Persian, *xan* 'khan' is a title to refer to the most respected tribal elder who decides about all internal and external affairs of a tribe. Thus, the modifier *xan* 'khan' contributes to the meaning of the head element via the same metonymy of PERSON FOR PROPERTY. The referent, however, follows from a literal interpretation of the compound head. Similarly, *gol-**mix*** (flower + nail) in (4) is also a right-headed compound where the modifier, i.e. *gol* 'flower', as a whole thing stands metonymically for its ornamenting function (ENTITY FOR FUNCTION). Another interesting metonymical right-headed compound is ʃab-**bu**. In contrast to examples (1) to (4), in this right-headed compound the head element *bu* 'smell', is metonymical. It refers to the flower via the metonymy of PROPERTY FOR ENTITY while the modifier of the compound specifies the time when the flower exudes its smell.

Examples in (6) to (10) represent five metonymical left-headed compounds. The metonymical modifier is observable in (6), (7), and (8). In these three left-headed compounds the modifier represents the name of a fruit which does not actually exist in the referent as **ab**-andʒir, **ab**-alu, and **ab**-zereʃk are produced by soaking the respective dried fruit in water and removing it before consumption. Thus, the fruit stands for its color and flavor as expressed in the metonymy of ENTITY FOR PROPERTY. By contrast, **pa**-taɢ and **tʃub**-parde in examples (9) and (10) are two left-headed compounds, in which the head element is motivated by metonymy. In **pa**-taɢ, *pa* 'foot' stands metonymically for its low position in the body (BODY PART FOR POSITION IN SPACE). In **tʃub**-parde, the head *tʃub* 'wood' stands for the object (rod, even though it is usually made of metal nowadays) and creates a metonymic link to the referent of MATERIAL FOR ENTITY.

A close look at the metonymical constituent and the referent in (5), (9), and (10) shows that they share a common feature. In all these examples, the conceptual relation between the head constituent and the referent of the compound is non-literal but can be construed via metonymy. Thus, in ʃab-**bu**, 'smell' is the named element but 'flower' is the actual referent, and, in **pa**-taɢ, 'foot' is named but the meaning only relates to the low position of the entity. The metonymic links between the head element and the referent are based on the typical metonymies of PROPERTY FOR ENTITY (as a specification of a PART FOR WHOLE metonymy) in (5), ENTITY FOR PROPERTY (as a specification of a WHOLE FOR PART metonymy) in (9), and on a MATERIAL FOR ENTITY metonymy in (10).

Alternatively, there are some compounds where the metonymical element is expressed in the referent, typically in a subordinate/superordinate relation. Let us take a closer look at compounds where the metonymical component represents such a relationship with the referent.

(11)   ***dʒaʔbe**-pargar* (box + compass) 'box including compass, ruler, protractor, etc.'

(12)   ***dʒaʔbe**-rang* (box + color tube) 'painting box set; box including paint tubes, brush, canvas board, etc.'

(13)   ***maʃin**-hesab* (machine + counting) 'calculator'

(14)   *dam-**pezeʃk*** (cattle + doctor) 'veterinarian'

(15)   *qalpaq-**dozd*** (hubcap + thief) 'thief who steals exterior accessories of cars'

Examples in (11) and (12) are two left-headed compounds which are based on the construction of *dʒaʔbe*-x. In (11), the modifier, *pargar* 'compass', stands for the whole tools in the box through the metonymy of MEMBER FOR CATEGORY. In the same way, the modifier in (12), i.e. *rang* 'color tube', refers to all tools included in a painting box, evoking the same metonymy of MEMBER FOR CATEGORY. *Rang* 'color tube' and *pargar* 'compass' exist as part of the whole referent. Another example of a left-headed compound with a metonymical modifier is provided in (13). In ***maʃin**-hesab*, the modifier, i.e. *hesab* 'counting', refers to calculating through the metonymy of PART OF ACTION FOR WHOLE ACTION.

The examples in (14) and (15), *dam-**pezeʃk*** and *qalpaq-**dozd***, are two right-headed compounds that contain a metonymical modifier. They also evoke a MEMBER FOR CATEGORY metonymy as *dam* 'cattle' in (14) stands for the category of animals, and *qalpaq* 'hubcap' in (15) refers to all exterior accessories of a car.

As these examples show, in some cases, the metonymical constituent is conceptually an inherent part of the referent, because the named element represents the referent more directly via a subordinate/superordinate relation. Based on this kind of relationship with the referent, the claim can be made that inferring the referent in these compounds is not as associatively complex as in compounds in which the metonymical link between the metonymical constituent and the referent is not based on purely taxonomic relations but emerges from a contiguous relation within one conceptual domain. Another example of a superordinate – subordinate relation can be seen in (16).

(16)   ***kolah**-pust* (hat + skin) 'lambskin hat'

***kolah**-pust* is a left-headed compound whose modifier, *pust* 'skin', refers to lambskin through the metonymy of CATEGORY FOR MEMBER. Although the metonymical constituent in this compound is also present in the referent, it connects to the referent through a 'type of' relation, which is different from the 'part of' relation in the previous examples (11–15).

Keeping the case of ***kolah**-pust* in mind, consider the following compounds which show a lower level of complexity:

(17)   *dʒudʒe-**kabab*** (chicken + kebab) 'chicken meat kebab'

(18)   ***halim**-buqalamun* (porridge + turkey) 'porridge containing turkey meat'

Example (17) and (18) share the same metonymical relation in the modifier. In *ʤuʤe-**kabab***, the modifier, *ʤuʤe* 'chicken', stands for chicken meat via the metonymy of ANIMAL FOR EDIBLE SUBSTANCE OF ANIMAL. The same metonymy holds in *buqalamun* 'turkey', as the modifier in (18). In (17) and (18), the metonymies are prompted by the other component of the compound. For instance, in *ʤuʤe-**kabab*** (chicken + kebab) 'chicken meat kebab', the head element evokes the frame of food and prepares the ground for the metonymic reference to the meat of the animal. In the other compounds discussed earlier, the metonymical components do not relate so closely to the meaning of the constituent. In other words, in (17) and (18) the frame of the non-metonymical compound constituent creates a context that prompts the specific metonymic relation, which becomes associatively salient, i.e. expectable.

Example (16) ***kolah*-*pust*** (hat+ skin) can illustrate that difference. Even though *kolah-pust* shares a general WHOLE FOR PART metonymy with the examples (17) and (18), its specific 'type of' relation is less predictable than in the latter group of compounds. Thus, it appears that examples (17) and (18) in the data at hand constitute the lowest degree of metonymic complexity.

To sum up, the close analysis of Persian compounds that involve one metonymical constituent shows that different degrees of metonymic complexity can be distinguished in the data. In that type of analysis, complexity can be gauged by two interpretative criteria. On the one hand, it is based on the conceptual relation between the intended referent of the metonymical constituent and its lexical expression. On the other hand, complexity can be related to how likely the metonymical target emerges from the frame-based interaction of the compound constituents. Figure 1 illustrates a continuum of metonymical complexity from low to high based on these criteria.

Frame-based selection of metonymical relation is highly restricted/expectable

Frame-based selection of metonymical relation is less restricted/expectable

| Low | | High |

Compounds whose metonymical element is directly expressed (named) in the referent

Compounds whose metonymical element is part of the referent category (subordinate relation)

Compounds whose metonymical element is not expressed in the referent

Figure 1. Continuum of metonymical complexity in one-part metonymic compounds

The two dimensions of complexity in Figure 1 describe different aspects that are indicative of the associative relation and its linguistic expression on the surface. In a view of modifiers filling a role in the head frame of the compound (cf. Wisniewski 1997; Onysko 2010), this frame-based specification can follow from more or less prototypical associations. The assumption

here is that more prototypical (i.e. likely) associations underlie less complex meaning construction in conceptual combination. The second dimension takes the similarity/difference between the intended referent and its lexical expression as a qualitative criterion of metonymical associations. If, as shown in the data analysis, the metonymical target is (partly) expressed in the linguistic form, the construal of the referent follows from an associatively less complex metonymy.

One of the questions that arise from the continuum of metonymical complexity is whether it makes a difference if the metonymy occurs in the modifier or the head constituent of the compound. This question follows from the observation that the link between the compound constituents and the overall referent of the compound is frequently not symmetrical, but the head constituent tends to be prominent in building the referent category (just consider the simple, non-metonymical compound *table cloth* whose referent category is a type of cloth). Empirical testing of this issue and of the other predictions emerging from the model in Figure 1 would be a next necessary step to consolidate or falsify the proposal inferred from the data.

6. Concluding remarks

Considering the important role of metonymy in the meaning of compounds, this study set out to investigate differences in the metonymic construals in one-part metonymic Persian compounds. As the analysis showed, the relation between the metonymical constituent and the referent of the compound is not the same in all cases of metonymic compounds. While in some examples, the metonymical element does not match with the referent (e.g. *pɑ-tɑɢ* foot + arch), there are some instances where the metonymical element represents the referent name in a subordinate relation, for example through the metonymy of MEMBER FOR CATEGORY (e.g. *dʒaʔbe-pargɑr* box + compass). The 'type of' relation in *kolah-pust* (hat + skin), on the other hand is similar to compounds in which the metonymical element is expressed in the referent name. This implies that metonymic strength varies in different metonymical compounds and there are different degrees of metonymic complexity based on the relation between the metonymical element and the referent. A second dimension of metonymical complexity follows from how prototypical the metonymic association is in the frame-based combination of the compound constituents. Based on these factors, a continuum can be postulated that ranges from simple metonymies to complex metonymies. The analyzed compounds in this study lie on different points of this continuum. The current proposal calls for future studies that will allow consolidating the continuum of metonymical complexity empirically with different datasets and experimental studies.

References

Anvari, Hasan. 2003. *Farhang-e Bozorg-e Sokhan* [Sokhan Comprehensive Dictionary]. Tehran: Sokhan.

Barcelona, Antonio. 2008. The interaction of metonymy and metaphor in the meaning and form of "bahuvrihi" compounds. *Annual Review of Cognitive Linguistics* 6. 208–281.

Barcelona, Antonio. 2011. Reviewing the properties and prototype structure of metonymy. In Benczes, Reka, Antonio Barcelona & Francesco José Ruiz de Mendoza Ibáñez (eds.), *Defining metonymy in cognitive linguistics: Towards a consensus view*, 7–57. Amsterdam: John Benjamins.

Barcelona, Antonio. 2012. Metonymy in, under, and above the lexicon. In Alegre, Sará Martín, Melissa Moyer, Elisabeth Pladevall & Susagna Tubau (eds.), *At a time of crisis: English and American studies in Spain*, 254–271. Barcelona: Departament de Filologia Anglesa i de Germanistíca, Universitat Autònoma de Barcelona.

Benczes, Reka. 2006a. Analyzing metonymical noun–noun compounds: The case of freedom fries. In Benczes, Reka & Szilvia Csábi (eds.), *The Metaphors of Sixty: Papers Presented on the Occasion of the 60th birthday of Zoltán Kövecses*, 46–54. Budapest: Eötvös Loránd University.

Benczes, Reka. 2006b. *Creative compounding in English: The Semantics of Metaphorical and Metonymical Noun–Noun Combinations*. Amsterdam: John Benjamins.

Benczes, Reka. 2011. Putting the notion of "domain" back into metonymy: Evidence from compounds. In R. Benczes, Reka, Antonio Barcelona & Francesco José Ruiz de Mendoza Ibáñez (eds.), *Defining Metonymy in Cognitive Linguistics: Towards a Consensus View*, 197–215. Amsterdam: John Benjamins.

Benczes, Reka., Barcelona, Antonio. & Ruiz de Mendoza Ibáñez, Francesco José (eds.). 2011. *Defining Metonymy in Cognitive Linguistics: Towards a Consensus View*. Amsterdam: John Benjamins.

Brdar, Mario. 2017. *Metonymy and Word-Formation: Their Interactions and Complementation*. Newcastle upon Tyne: Cambridge Scholar Publishing.

Brdar, Mario & Brdar-Szabó, Rita. 2013. Some reflections on metonymy and word-formation. *Explorations in English Language and Linguistics* 1.1. 40–62.

Brdar, Mario & Brdar-Szabó, Rita. 2014. Where does metonymy begin? Some comments on Janda (2011). *Cognitive Linguistics* 25 (2). 313-340.

Croft, William. 1993. The role of domains in the interpretation of metaphors and metonymies. *Cognitive Linguistics* 4. 335–370.

Croft, William. 2002. The role of domains in the interpretation of metaphors and metonymies. In Dirven, René & Ralf Pörings (eds.), *Metaphor and Metonymy in Comparison and Contrast*, 161–205). Berlin: De Gruyter.

Croft, William. 2006. On explaining metonymy: Comment on Peirsman and Geeraerts, "Metonymy as a prototypical category". *Cognitive Linguistics* 17(3). 317–326.

Drożdż, Grzegorz. 2014. The development of the theory of metonymy in cognitive linguistics. *Linguistica Silesiana* 35. 119–151.

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

Geeraerts, Dirk. 2002. The interaction of metaphor and metonymy in composite expressions. In Dirven, René & Ralf Pörings (eds.), *Metaphor and Metonymy in Comparison and Contrast*, 435–465. Berlin: De Gruyter.

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

Kalbassi, I. 1997. *Sakhte eshteghaghiye vage dar zabane Farsi emrooz* [The derivational structure of word in modern Persian]. Tehran: Institute for Humanities and Cultural studies.

Kosecki, Krzysztof (ed.). 2007. *Perspectives on Metonymy*. Frankfurt am Main: Peter Lang.

Kuczok, Marcin. 2016. The interplay of metaphor and metonymy in English noun+noun compounds. In Drożdż, Grzegorz (ed.), *Studies in Lexicogrammar: Theory and Applications*, 193–210. Amsterdam: John Benjamins.

Lakoff, George. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: The University of Chicago Press.

Lakoff, George & Johnson, Mark. 1980. *Metaphors We Live by*. Chicago: The University of Chicago Press.

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

Lieber, Rochelle & Štekauer, Pavol (eds.). 2011. *The Oxford Handbook of Compounding*. Oxford: Oxford University Press

Onysko, Alexander. 2010. Casting the conceptual spotlight: Hybrid compounding in German as an example of head-frame internal specifier selection. In Onysko, Alexander & Sascha Michel (eds.), *Cognitive Perspectives on Word Formation*, (243–300). Berlin & New York: De Gruyter.

Onysko, Alexander. 2014. Figurative processes in meaning interpretation: A case study of novel English compounds. *Yearbook of the German Cognitive Linguistics Association* 2. 69–88.

Onysko, Alexander. 2016. Enhanced creativity in bilinguals? Evidence from meaning interpretations of novel compounds. *International Journal of Bilingualism* 20(3). 315–334.

Panther, Klaus-Uwe & Radden, Günter (eds.). 1999. *Metonymy in Language and Thought*. Amsterdam: John Benjamins.

Panther, Klaus-Uwe & Thornburg, Linda. 2007. Metonymy. In Geeraerts, Dirk & Hubert Cuyckens (eds.). *The Oxford Handbook of Cognitive* Linguistics, 236-262. Oxford: Oxford University Press.

Peirsman, Yves & Geeraerts, Dirk. 2006. Metonymy as a prototypical category, *Cognitive Linguistics* 17(3). 269–316.

Radden, Günter. 2005. The ubiquity of metonymy. In Otal Campo, Jose Luis, Ignazio Navarro I Ferrando & Begoña Bellés Fortuño (eds.), *Cognitive and Discourse Approaches to Metaphor and Metonymy*, 11–28. Castelló de la Plana: Publicacions de la Universitat Jaume.

Radden, Günter. & Kövecses, Zoltán. 1999. Towards a theory of metonymy. In Panther, Klaus-Uwe & Günter Radden (Eds.), *Metonymy in Language and Thought*, 17–59. Amsterdam: John Benjamins.

Radden, Günter., Köpcke, Klaus-Michael, Berg, Thomas & Siemund, Peter. (2007). The construction of meaning in language. In Radden, Günter, Klaus-Michael Köpcke, Thomas Berg, Peter Siemund (eds.), *Aspects of Meaning Construction*, 1-15. Amsterdam: John Benjamins.

Shariat, M. J. 2005. *Dastoor-e zaban-e Farsi* [Persian grammar 1]. Tehran: Asatir.

Torabian, A. (2013). *Barresi-e este'are dar esm-ha-je morakkab-e esm-esm-e zaban-e Farsi dar čarčub-e zabanšenasi-e šenaxti* [Metaphor in Persian noun-noun nominal compounds in cognitive linguistics]. Master's thesis. Shiraz: University of Shiraz.

Ullmann, Stephen. 1957. *The Principles of Semantics* (2nd edition). New York: Barnes & Noble.

Wisniewski, Edward J. 1997. When concepts combine. *Psychonomic Bulletin & Review* 4. 167–183.

Masoumeh Diyanati
Department of Linguistics
Faculty of Foreign Languages
University of Isfahan
Azadi Square, Isfahan, Iran
m.dianati@fgn.ui.ac.ir

Alexander Onysko
Department of English
University of Klagenfurt
Universitaetsstrasse 65-67
9020 Klagenfurt, Austria
alexander.onysko@aau.at

Modern German Prison Discourse: Mental Resource

Iryna Osovska, Yuriy Fedkovych Chernivtsi National University
Liudmyla Tomniuk, Yuriy Fedkovych Chernivtsi National University

The modern studies on the background of anthropocentric methodological matrix with the fundamental idea of learning the language in its functioning as an integrator of thinking and mediator in the mechanisms of consciousness of a speaker that is immersed in life, undergoes transformation, acquires a certain socio-cultural context and becomes a discourse. Such an approach substantiates the interest in the study of the mental basis of various discourse practices, for which a prison discourse covering the communication of prisoners and penitentiaries in places of imprisonment and being a sociocultural center of the prison language is rather specific.

The results of the study of the mental resource of modern German prison discourse confirm that there is a significant difference between the literary and real-prison forms that are presented by the authors of literary works on the corresponding subjects and real prisoners. Thus, at the forefront of the literary prison discourse, there is a general psychological knowledge about a person as a biological being, a representative of society, the features of his external and internal body structure, his spiritual space and typical activities, among which the central place in the German author's picture of the world is given to communicative activities. Instead, in real prison discourse, there is a fixed knowledge of prisoners about the prison as a reflection of German society, the prison as an epicentre of a specific subculture and its impact on the emotional and psychological state of the prisoner; cognitive changes taking place in the convicted person's ideas during the stay of the prisoners; the nature and peculiarities of the German penitentiary system, the classification of crimes, the types of punishment under criminal law; processes and physical activities that imbue the prison everyday life.

Keywords: *prison discourse, concept system, domain, concept, conceptual analysis.*

1. Introduction

The growing interest in the study of human consciousness through the concepts as its operational units, the access to which for a linguist is provided by a word, led to the establishment of anthropological paradigm in the field of linguistics, the supporters of which consider a comprehensive study of the entire spectrum of discourse practices as varieties of social interaction between parts in the communication process to be one of the priority tasks.

Of all the diversity denoted by a very popular word 'discourse', we choose the aspects relevant of our study: a) "the forms of discourse are as diverse as the forms of human life" (Kibrik 2009: 4); b) "every kind of discourse appears to be a special speech practice, a specific sublanguage that creates its own picture of the world and attends to a certain functional area of the linguistic community" (Naumenko 2003: 122); c) "every discourse also has power because whenever language is used there is an effect in the context surrounding discourse" (Lirola & Rubio 2012: 60).

The above mentioned gives grounds to consider the chosen object of research – *contemporary German prison discourse* (further referred to as PD) as one of the discursive practices that function in the space-time continuum of the German ethno-community.

The importance of selecting such an object is substantiated by both the socio-cultural factors (establishment of a criminal subculture through the media, penetration of its leitmotif into all spheres of public life, the non-obvious nature of the answer to the question of the corrective function of a prison, etc.), and purely linguistic ones, being determined by the constant growth of scientific interest to the mutual influence of social relations, thinking of man and language, general orientation of modern linguistic studies on the study of language through the prism of cognitive processes in consciousness. Due to the social significance of the prison as a social institution located within the society (and not beyond its borders, as previously thought), there is an objective need for a scientific justification of the modern PD as a space of coexistence of the convicts and prison staff, during which both their cognitive worlds and linguistic personalities, manifestation of cognitive mechanisms and forms of their verbal representation interact.

Based of a clear understanding of the direct connection between language and thinking, human consciousness, categorization, memory, and other cognitive functions, the task of linguistic analysis of discursive practice within the cognitive approach is regarded as the reconstruction of its mental foundation, because in the process of speech it is the language means, that help man to explain his own mental models, in which his knowledge of the surrounding world is preserved (Viehöver & Keller 2013: 14).

2. Discussion

PD now appears to be a research object that has hardly ever been interesting for linguists. Still there are some works by scientists in the field of criminology and jurisprudence, in which an attempt was made to distinguish between the legal (O. Dzioban), criminological (A. Zhalinsky), psychological and criminological (Yu. Irkhin), historical and legal discourse (V. Timashov). The existence of a discourse relating to the criminal world is mentioned by D. Altaide, K. Gregoriu and L. Fillippini, but there is no substantial linguistic research in which the PD would be the object of research. Instead, the felonious / criminal jargon as a component of the criminal subculture has long been of interest to linguists (E. Golin, A. Gurov, R. Günther, V. Driomin, K. Laubental, L. Masenko, V. Pirozhkov, Y. Tsimer).

3. Prison discourse as a discursive practice

Prisons, as well as orphanages, psychiatric hospitals, barracks and monasteries (Schumacher 2013: 28) belong to the so-called 'total institutes' (Goffman 1980: 314), the common features of which are: the concentration of all members within the institution, where the management is carried out by the central authorities; compelled interaction with other members of the community; availability of regulations; constant monitoring and supervision of the social community members, their work and life.

The penitentiary system of Germany is one of the oldest and most advanced ones in Europe. There are open (*offene Gefängnisse*) and closed prisons (*geschlossene Gefängnisse*). Today in the country about 70,000 prisoners are serving sentences in nearly 200 prisons (de.statista.com). Each prison is subordinate to the Ministry of Justice of the federal state, where it is situated, thus the prison conditions are different. Recently, two new types of prisons appeared in Germany: the private one (*das Privatgefängnis*) and a corrective institution for the elderly (*das Seniorengefängnis*).

In modern Germany, prison is regarded as a separate legal micro-society, which is much more permeated with power, control, regulation and punishment. The total institute differs from other forms of social institutes by the restriction of social connection with the outside world.

The significance of linguistic culture in the prison was substantiated by sociologists in 1970s. Thus, Clemmer came to the conclusion that only due to the intermediary role of the language the interpersonal contacts are established, relations developed and the culture of both an individual and the social community are reflected (Clemmer 1968: 88). For separation from the outside world and safe communication, prisoners have created their own language – a criminal jargon. In the process of prisoners' communication symbols, gestures, signs, nicknames, and tattoos that function in an organized subcultural hierarchy as a distinctive sign, are of certain meaning.

The above mentioned proves that the prison subculture is characterized by a special picture of the world, because its values and norms are not only within the limits of the criminal environment, but also are spread in the whole society, influencing its culture and the spiritual world of each person.

Despite the permanent interest the term 'prison discourse' was not used in the field of science, but its meaning was explained at the intersection of the legal ("activity aimed at the control and regulation of social relations by formulating standards and legitimization (legal form) as well as regulation of social relations between persons with powers, freedoms and responsibilities of the subjects of discourse" (Kozhemyakin 2011: 63)), and the criminal (all thinking and speech activity related to cases of violation of human law).

The sphere of PD's implementation is a prison and its participants are the convicted and the representatives of other social groups in the institutions of confinement, in particular the prison authorities, representatives of the criminal-executive system and prison staff, including special services such as psychologists, social workers, priests who carry out their professional activities in close contact with prisoners, that definitely affects their verbal behavior (Klocke 2004: 22). The sociolects of each of these groups are in constant close interaction (Günther 2005: 14); this testifies to the relevance of introducing the notion of 'prison discourse'.

To summarize we define the modern German PD as a complex, interdisciplinary, socially determined phenomenon, an area for the realization of the thinking and speaking activities of representatives of German ethnic community that violated the law by committing an unlawful socially dangerous act whose guilt had been proved during the trial, on the basis of which they were convicted and sentenced to imprisonment. While serving their sentence, the mechanism of consciousness of all representatives of the criminal world, whose focus is the prison, is reflected in the concepts as units of the mental resource of individual consciousness.

Thus, the *aim* of the research is to reconstruct the mental resource of contemporary German PD – the conceptual space reflected in the literary and spoken forms of representation (hereinafter – LPD and SPD), which reproduces the literary (author's) and everyday prison pictures of the world.

Since the form of PD representation is an important factor for analysis and, as the study has shown, for conclusions, we will concentrate on each of the analyzed forms separately.

4. Literary discourse as fictional reality modeled by the author

Cognitive-discursive approach to the text (Babenko 2005: 15), considering the content of the text as a "functional field of meaning, represented by the linguistic units, fixed in literary text

concepts" (Bechta 2011: 246) determines identifying the correlation between elements of the triad *author – text – non-textual activity*, the reconstruction of the concept of literary discourse, its mental models and conceptual space as the main research task, since the representation of the basic concepts that are evaluated in the text and undergo rethinking from the position of the author and / or the character is the purpose of literary communication.

Literary discourse (LD) is of particular interest to linguists, because it has features that are very different from those of other discourse types. One of these key features of the LD is its purpose, which is the intention of the author to directly influence 'the spiritual space' (Leibniz 2010: 126) of the recipient (the reader) in order to further influence his personality, worldview and values. Another divergence is reduced to secondary nature of LD in relation to primary spoken genres and to its property to form a multilevel structure of meaning (Burdyugova 2012: 16).

Consequently, LD is a sociocultural interaction between the author and the reader, which involves aesthetic, spiritual, cultural, social values, acquired knowledge of the world, peculiarities of the worldview and perception of reality, a system of feelings, beliefs and ideas, which is aimed at making corrections to the addressee's *spiritual space* and causing his emotional reaction through influencing him. Unlike other texts, the text of a literary work is a work of art of the literary word, the result of the creative speech and intellectual activity of the author that reflects the peculiarities of the individual author's perception of reality, embodies the picture of a fictional reality in the unity of form and content (Koltsova 2007: 3) and through conceptual and linguistic means reproduces the author's vision of the world.

On the basis of the above, the literary prison discourse (LPD) is interpreted as a kind of socio-cultural creative, mental and speech activity of the author, reflecting the specificity of his national linguistic culture and the peculiarities of the individual author's perception of the prison as the subject of reality, the result of which is the text as fictional reality modeled by the author's imagination. In it, in the dialogic communication of characters, represented by the author as the status-role relations between prisoners and employees of the penitentiary system, the author's knowledge and understanding of the prison as a social institution, the focus of the criminal subculture and the environment of social interaction prisoners and employees of the penitentiary system that plays a fragment of author's vision of the world are represented with the help of conceptual and linguistic means.

Another form of representation of the PD is a spoken discourse.

5. Spoken discourse as a real discursive activity

Spoken discourse (SD) is the initial form of expression and the primary discursive form that serves our daily lives, "exploring the traces of latent processes of deploying thought in a verbal statement that occurs during internal speech" (Sedov 1999: 10).

The process of sampling and graphic fixation of oral empirical material is complex and problematic. In the context of our study, this is due, first of all, to the closedness of the system, as well as laboriousness of transcribing videos, justified only by the need "to preserve the semantic and medial quality of communicative practice and to record in writing what has been said for a long period of time" (Diaz-Bone 2015: 180).

However, the difference between the oral and written modi of discourse is much deeper. Many researchers (I. Boldonova, A. Budnyk, A. Depermann, A. Kibrik, V. Ong, V. Podlesskaya, N. Turanina, V. Kharchenko etc.) believe that SD and LD are at different poles, since the literary

word belongs to the fictional, modeled by the author reality, while the oral word is "a part of the ontological reality" (Ong 2016: 94) taking into account the fact that the participants are existing living people.

The differences between the two modi can be observed in their structure. Thus, the important categories of written discourse are completeness, connectivity, compositional structure, consistency and integrity (Budnik 2016: 41). The basis of SD is formed by discursive markers, pauses, meaningful changes in the frequency of the main tone; the division of the sound stream into quanta, characterized by prosodic, semantic and syntactic unity; the determination of the main types of illocutions and the connection between elementary discursive units (Kibrik & Podlesskaya 2003: 47).

Thus, SD is a live speech, characterized by emotionality, instantaneity, ease, informality, spontaneity, low level of planning, and it regulates the provision of socialization and individualization of the person.

Taking into account the above, we define the spoken prison discourse (SPD) as social, everyday thinking and spoken activity, which is reproduced in the communication of prisoners and employees of the penitentiary system in institutions of confinement by verbal and nonverbal means through natural channels of communication.

6. Material and methods

Selection of the research material – the corpus of LPD and SPD texts as forms of the PD representation, met two key criteria: 1) participation of at least one convicted in the process of communication; 2) a modern German penitentiary institution as the place of the deployment of a communicative situation.

The selection of the LPD excerpts, which represent the communication acts unfolding in institutions of confinement between the characters, given the roles of convicts or penitentiaries by the author, is performed by the continuous sampling method. The excerpts from 1,000 dialogues, 140,227 words in total, from 186 works of modern German literature (2003-2017) have been selected.

Live dialogues between prisoners and employees of the German penitentiary institutions (1,000 dialogues from audio- and video-interviews of 100 hours total duration and 138,403 words used), representing the SPD obtained from radio broadcasts from institutions of confinement, videos from prison cells, and documentary videos recorded there, access to which to a greater extent was obtained from the participants of '*Podknast*' project that included penitentiary institutions of 10 federal states.

Written fixation of the passages of the convicts' *live* speech was carried out through *minimal discursive transcription*, the main task of which, unlike the phonetic and phonological transcription, is not in taking into account the sound composition of individual word forms, but in fixing phenomena associated with the organization of a local discursive structure, so in most cases it is done by writing words in the standard spelling form, which corresponds to the norm of the literary language, provides simplicity and speed of text perception and processing (Makarov, 2003: 109). The method of minimal discursive transcription involves preserving the key characteristics of *live* speech (phonetic morphological and syntactic ones) and omitting accents, pauses, intonation and timbre peculiarities, length and spontaneous interruptions typical for oral speech (Adamzik 1994: 370).

The theoretical and methodological basis of the research consists in the idea that mental resource of a certain discursive practice is a specific conceptual system that can be reconstructed by applying a cognitive mapping technique. Related to the PD its application may result in cognitive maps of its varieties – LPD and SPD, which present the process of preferring use of mental units and are specific 'information portraits'.

The attempts of scientists to represent complex mental processes are reflected in the suggestion of various types of structures, such as interpretation frames and cognitive maps. The latter, though differently understood by the researchers (N. Kravchenko, O. Selivanova, R. Akselrod, E. C. Tolman), but being aimed mainly at representing the global picture of the communicants' programs, represent the mental frames of discursive practices of the linguistic-cultural community and reveal the established way of comprehending a certain segment of its members' activities.

The creation of complete presentation of the PD concept system, reflected in the cognitive map, requires the determination of its 'skeleton' – the autochthon concepts and the partial outline of the corpus of PD allochthons as discursive variables. At that, the adequate methodological approach was to defining the synthesis of *conceptual analysis*, corpus linguistics techniques and *linguoquantitative methods*. This enables to observe the complete mental representation of the PD as a conceptual system in a statistically verifiable conceptual structure.

Modeling the LPD and SPD cognitive maps includes several stages: 1) defining basic situational formers that outline the communicative frame as a LPD and SPD situational 'skeleton'; 2) determining the LPD and SPD allochthons as a general complex of possible information elements; 3) determining autochthons as regular elements of LPD and SPD through the procedures of statistical verification of actuals.

Stage 1 is substantiated by the fact that the actualization of the PD semantic space is provided by an adequate situational framework. Due to anthropocentricity, the deictic characteristics of the dialogue discourse and understanding of the action transformations as a continuum of "causal chains", from which the consciousness of the interpreter "pulls out separate links" (Croft, 1991: 159), this situational framework can be outlined by the main formers presenting the participants – the prisoners and employees of the prison, – the global strategic objectives and the spatiotemporal localization of the PD communicative act.

Stage 2 requires additional explanation. The lexical-semantic space can be regarded as a combination of means for verbal interpretation of a certain conceptual space whose nuclear component is compared with a verbal unit capable of activating the given concept in the consciousness of the language native speaker. Considering such units the objectivators of concepts, one can detect the concepts that are regularly or occasionally represented in LPD and SPD.

Conceptual analysis of the PD, despite closedness of the prison "world", enables to define it as an open discourse, which has a certain ability to borrow concepts that are inherently characteristic of other discourses, which, obviously, is explained by the specification of different types of knowledge of various perspectives of human and linguistic existence of personality in it.

The holistic theory as a methodological basis of modern cognitive linguistics combines the formats of the presentation of the semantic and the conceptual, considering language as the main means of categorization and conceptualization of the world, an open cognitive system that interacts with general mental capabilities. Cognitive structures are deep thought-based entities, the discovery of which can only be done through the study of meanings of the language units (Tukaeva 2009: 862), because "the lexical meaning is a concept activated by the word in thought" (Zhabotinskaja 2013: 76). The word used by a person is the "flash" of information, – both the nominator of a certain notion or meaning, and an element of knowledge of the language native speaker in the

context of his national mentality, social, psychological, gender, and worldview characteristics. Any mental formation is constantly in the relations and dependences, and therefore the determination of the algorithm for the exteriorization of concepts in the space of the text enables not only to implement their profile, but also to construct the concept system of LPD and SPD.

The verbal implementation of concepts leads to the need of identifying their actualizers (known as objectivators, explicators, nominants, verbalizers, conceptual qualifiers, etc.) – verbal units that provide the conceptual basis of the PD types, despite the fact that communicative and pragmatic parameters for defining qualifying senses of the lingual units are quite subjective both from the side of the communicants, and from the perspective of the researcher. In order to isolate the constants of the concept system of PD types, an inventory of the PD texts was made, the objectivators of concepts were detected, the classification of which into the semantic domains allowed defining the matrix of concepts as the basic beams of meaning within the PD.

Since only the repetition of particular verbals can indicate a certain relevance and regularity, it seems appropriate to use the elements of statistical analysis enabling to determine the statistical significance of a) domain presentation in the LPD and SPD texts, and b) certain concepts in the selected domains at <u>Stage 3</u> of the study to give the status of LPD and SPD autochthons to certain concepts from other fields. Selected lexemes as linguistic signs, fixed to the subjects of cognition and by relations between them, represent elementary meanings in the consciousness; the names of domains formed – generalized concepts that semantically combine all the elements of a group that, through their presentation in the family interaction, indicate the points of meanings condensation.

The need to distinguish between natural and random phenomena causes applying *quantitative analysis*. Verification of data using the techniques of *calculating the χ^2-criterion* and *contingency coefficient K* allows to select the main meaning dominants, categorial units of the basis of its concept system from the set of fixed allochthons – conceptual variables presenting atypical knowledge quanta, regular autochthons of LPD and SPD concept system – by determining the correspondences between the frequency distributions of concept actualizers of a certain domain in the LPD and SPD texts and specific concepts in statistically significant domains.

The traditional study of the fullness of conceptual structures is based, as a rule, on the definition of the frequency of lexemes that designate certain elements or nominal characteristics of the concept. However, quantitative techniques, in particular, one of the basic methods for verifying hypotheses in linguistics – the chi-square (χ^2) criterion, enable to determine the existence of correspondences or discrepancies between distributions of frequencies of the quantities under observation (Levickij 2012: 156), actually verifying their regularity of this discursive environment. The most widely used formula for calculating the χ^2 criterion is $\chi^2 = \sum \frac{(O-E)^2}{E}$, where O – actually observed values, E – theoretically estimated ones, Σ – the total amount.

The relation between the features is confirmed by the χ^2 index, which is larger than the critical value, and its degree is determined by Chuprov mutual contingency coefficient K (Levickij 2012: 160) by the formula $K = \sqrt{\frac{x^2}{N\sqrt{(r-1)(c-1)}}}$, where N is the total number of observations, r is the number of lines in the table, c is the number of columns.

The values of the mutual contingency coefficient can be from 0 to +1, while the significance is determined in correspondence with the value of χ^2.

Exceeding the value of χ^2 testifies to the prevalence of the empirical use of the domain over the theoretical expectations, confirming its selective character and, consequently, its importance for the LPD and SPD.

A similar procedure determines the value of χ^2 for all allochthons actualizers within each of the distinguished domains. As a result of such calculations, the most significant concepts for each of the predefined statistically significant domains are determined in the LPD and SPD text. The detected concepts are autochthons of LPD and SPD.

Having explained in the above described way the social intragroup structure of thinking of the PD participants on the basis of texts produced by them, one can identify the common elements in the structure of thinking that integrate LPD and SPD within the network – the concepts as cementing elements in the national consciousness. All the information obtained results in the cognitive map – schematically presented information archive as "the way of representing a person's mental image of the surrounding world; a schematic image of a fragment of the world picture" (Yesipovich 2013: 254–255).

7. Concept system of LPD

In the LPD, seven statistically significant domains (SPRECHTÄTIGKEIT, MENSCHLICHER ORGANISMUS, RELIGION / PARANORMALE PHÄNOMENE, BEWEGUNG, GEFÜHLE UND EMOTIONEN, LAUT-UND LICHTERSCHEINUNGEN, PHYSIKALISCHE EIGENSCHAFTEN) and 34 concepts-autochthons (Table 1) are identified. We will analyze the conceptual content of each of the domains selected in more detail.

Table 1: *Significance of domains in LPD and SPD*

Domain	LPD		SPD	
	χ^2	К	χ^2	К
Medicine	27,68			
Physical characteristics of an object	93,2	0,4		
Sound / light phenomena	113,68	0,4		
Assessment / quality of an object			19,6	
Emotional and psychic state			88,24	0,32
Time			190,88	0,56
Features of character	2,72			
Money / property	14,56			
Objects/devices/ weapons / mechanisms	14,08			
Nature	30,08			
Emotions and feelings	309,13	0,62		
Law and legislation			607,46	0,82
Human body	712,37	0,88		
Religion / supernatural	579,12	0,81		
Process			185,84	0,48
Physical act			574,91	0,81
Physiological processes	15,12			
Space			34,48	
Colour / tint	23,52			
Speech activity	588,4	0,81		
Thinking activity			170,32	0,48

Substance / material		0,24	
Occupation / work		0,4	
Appearance	0,96		
Food and beverages	2		
Family and relations between people		24,8	
Clothes		4,2	
Motion	464,56	0,8	
Building / structure / institution	0,72		
Prison subculture		947,59	0,95

In the conceptual space of the LPD, the most important part is given to the domain MENSCHLICHER ORGANISMUS, which combines the concepts with somatic component, which objectify the biological knowledge of the convicted person about the body and its structure, peculiarities of the worldview, typical features of the character and appearance of the German community representatives *(KOPF, AUGEN, FUSS, HAND, MUND, HALS)*. The recorded autochthons are presented according to their share from the most significant to the least relevant ones:

(1)   *Einen KOPF kürzer machen!*
'Cut the HEAD off!' (Chaplet 2003: 101; examples hereinafter are in the authors' translation)

(2)   *Nachhe drehte ma die Daumen in die Faust. – Nachhe kugelt me die AUGEN raus.*
'Nachhe made a fist with his fingers. – Nachhe will put my EYES out.' (Meyrink 2007: 301)

(3)   *Wenn du wiederkommen und auch nur versuchen solltest, einen FUSS durch das Tor zu setzen, wird man dich töten.*
'If you come back or just even try to step your FOOT in the gate, you will be killed.' (Benedikt 2012: 126)

(4)   *Anderen Leuten das Essen aus der HAND schlagen.*
'Knock the food out of other people's HANDS.' (Chaplet 2003: 206)

(5)   *Haltet also bitte euren MUND, auch im eigenen Interesse.*
'Shut your MOUTHS. It's in your best interest.' (Ackermann 2015: 97)

(6)   *Das erklärt, warum sie Holoch totgeschlagen hat, aber was hatte es mit dem Beton in Jossecks HALS auf sich?*
'This explains why Holoch killed himself, but how can this be related to the concrete in his THROAT, Josseck?' (Hartung 2014: 143)

The SPRECHTÄTIGKEIT domain contains knowledge of the communicative activities of representatives of the prison world. Analysis of the empirical material proves that in the LPD the convicts start to communicate in order to satisfy their daily needs, receive and transmit information, maintain relations. The speech activity of the prisoners in the LPD is short-term, incomplete, expressive, marked with specific prison vocabulary and abusive language that mostly

have a negative connotation. The domain is represented by active concepts *(KNASTJARGON, REDE, SCHIMPF, DROHUNG, SCHREI, FRAGE, ERZÄHLUNG, ANTWORT, STREIT)*:

(7)   *Ist das ein SITTICH?*
         'Is he a PAEDOPHILE?' (Bausch 2012: 152)

(8)   *Da ihr Tölpel AUSGEPLAUDERT habt, was keiner wissen soll.*
         'That's because you, fools, BLATHERED about something that nobody should know.' (Ebert 2008: 99)

(9)   *Du SCHLAMPE. Du HURE. Du NUTTE.*
         'You're a DRAB, you're a SLUT. You're a WHORE.' (Chaplet 2003: 292)

(10)   *Eine falsche Bewegung, und du BIST TOT.*
         'One wrong move and you will DIE.' (Ebert 2008: 449)

(11)   *Ihr könnt SCHREIEN, so viel Ihr wollt, Weib.*
         'SCREAM as much as you wish, woman.' (Benedikt 2015: 145)

(12)   *Du sprichst nur, wenn du GEFRAGT wirst!*
         'You will talk only when you are ASKED!' (Ebert 2006: 173)

(13)   *Was hast du ihr ERZÄHLT?*
         'What have you TOLD her?' (Funke 2005: 513)

(14)   *Oh, darauf möchte ich jetzt nicht ANTWORTEN.*
         'Oh, I don't want to ANSWER that now.' (Spaniel 2015: 98)

(15)   *Das war unser großer STREITpunkt.*
         'That was our big CONTROVERSIAL point.' (Link 2012: 618)

The religious beliefs and knowledge about the world, its structure and origin conventional for the German ethnic community are represented by the domain RELIGION / PARANORMALE PHÄNOMENE, which is a religious component of the conceptual picture of the world, based on belief. Active domain concepts are *GOTT, JENSEITS, HÖLLE, KIRCHE:*

(16)   *Und es gibt das GOTTESGERICHT, vor dem einen keine Macht und kein Schmiergeld schützt.*
         'There is the GOD'S JUDGEMENT, before which nobody can be protected by power or bribery.' (Milstein 2012: 30)

(17)   *Mit so einem Delikt ist man normalerweise fast so angesehen wie jemand, der seine untreue Frau ins JENSEITS geschickt hat.*
         'Such crime usually makes you have the same authority as the one, who sent his cheating wife into ETERNITY.' (Chaplet 2003: 240)

(18)   *Wie ist sie aus dieser HÖLLE entkommen?*

'How did she escape this HELL?' (Hartung 2014: 143)

(19)   *DIE ANDACHT ist für alle unsere Abteilungen offen.*
         'The WORSHIP is open to all our departments.' (Chaplet 2003: 181)

The domain BEWEGUNG is the cognitive fragment of the conceptual portrait of the German LPD. In the conceptual content of the BEWEGUNG domain, we observe confining the general knowledge of the convict about motion as the basis of life and the key process that occurs in nature, to specific information about their limitations in the place of confinement, where the main kind of motion is walking, which is recorded in autochthons *GEHEN* and *KOMMEN:*

(20)   *Wann GEHEN wir los und knallen Patrick ab?*
         'When do we GO and shoot Patrick?' (Hartung 2014: 122)

(21)   *Morgen KOMMT der Kerl, der den kleinen Felix umgebracht hat.*
         'The guy, who killed little Felix will COME tomorrow.' (Chaplet 2003: 194)

Feelings and emotions are an integral part of the inner world of man, even in places of confinement. The domain GEFÜHLE UND EMOTIONEN in the LPD is represented by a number of concepts for denoting both negative (much more often) and positive emotions and feelings. The scope of negative emotions is reflected in the concepts *ANGST, HASS, TRAUER, VERDACHT, ABSCHEULICHKEIT:*

(22)   *Aber ich habe ANGST davor, dass ich dabei ausraste oder dass ich das nicht überlebe.*
         But I am AFRAID that I will lose control or won't be able to stand this.' (Bausch 2012: 212)

(23)   *Ich HASSE dich!*
         'I HATE you.' (Anonymus 2009: 263)

(24)   *Sie TRAUERT um eine zerstörte Unschuld.*
         'She MOURNS the broken innocence.' (Hartung 2014: 146)

(25)   *Wir hatten den VERDACHT, dass Jenkins ein doppeltes Spiel spielte.*
         'We SUSPECTED Jenkins to play a double game.' (Neuhaus 2009: 996)

(26)   *Ich sag ich Ihnän; es war Ihnän SCHAISLICH.*
         'I am telling you, you felt DISGUSTED.' (Meyrink 2007: 304)

The concentrators of positive feelings include *HOFFNUNG* and *FREUDE*:

(27)   *Man kämpft in der HOFFNUNG, dass der Andere aufgibt, denn man weiß, man wird ein Leben nehmen.*
         'You fight with HOPE that the enemy gives up, because you know that your life will be taken.' (Aukett 2015: 655)

(28) *Meine Schwester war bis zu dem Tag, da sie mit Gewalt verschleppt wurde, glücklich und FROH.*
'Before the day my sister was hijacked by means of violence, when she was happy and JOYFUL.' (Benedikt 2015: 144)

Actualization of knowledge of the convicts about sound / light phenomena and physical characteristics of the objects of reality are sufficiently rarely observed, which is a consequence of the specificity of the place of criminal communication. From the research material, we receive confirmation that the prisoners are able to observe only a narrow spectrum of light and sound phenomena such as the sunset and the illumination of the prison territory at night from behind the bars, the light of the lamp in the cell, the knock on the steel door of the cell, the rustle, whispers, human sounds. At the same time, there is a restriction of the prisoner in the objects of everyday life that are restricted to minimum in order to prevent the offense in the places of imprisonment. Therefore, there is no need to discuss the size, shape, length, height, and other physical properties of objects. The obtained statistical values of the domains of LAUT-UND LICHTERSCHEINUNGEN and PHYSIKALISCHE EIGENSCHAFTEN indicate that they are most passive part in the formation of the LPD conceptual space. The reserve for the LAUT-UND LICHTERSCHEINUNGEN domain is represented by the concepts that contain audio information (*GERÄUSCH* and *TON*) and that of light phenomena *(LICHT):*

(29) *Ich hatte ein GERÄUSCH gehört, einen dumpfen Schlag von oben.*
'I heard some NOISE, the thud from above.' (Adlon 2014: 89)

(30) *Das Einzige, was sie als Antwort bekam, war EIN NERVTÖTENDER DAUERTON.*
'The only thing she got as an answer was an EXHAUSTING ENDLESS DIAL TONE.' (Adlon 2012: 75)

(31) *DAS HELLE LICHT beleuchtete das in Gram verzogene Gesicht.*
'The BRIGHT LIGHT was illuminating the face, distorted by sorrow.' (Hartung 2014: 146)

The domain PHYSIKALISCHE EIGENSCHAFTEN combines the concepts *LÄNGE* and *GRÖSSE*:

(32) *Und außerdem sind es höchstens FÜNFZIG KILOMETER von dort.*
'Besides, it's more than FIFTY KILOMETERS away.' (Chaplet 2003: 92)

(33) *Er wog DAS SCHMALE MESSER in seinen Händen.*
'He juggled with A LITTLE KNIFE.' (Aukett 2015: 723)

Consequently, the statistically selected concepts as elements of the content of meaningful domains are autochthons, discourse-forming units of criminal communication in the German literary space that characterize it as a separate form of discursive practice through the prism of German authors' knowledge and ideas about the prison and the prisoners. Having selected the most active in the literary form of representation PD concepts that obviously represent the author's picture of the world, we will represent the mental portrait of the LPD in the cognitive map (Figure 1).

Figure 1. *Concept system of LPD*
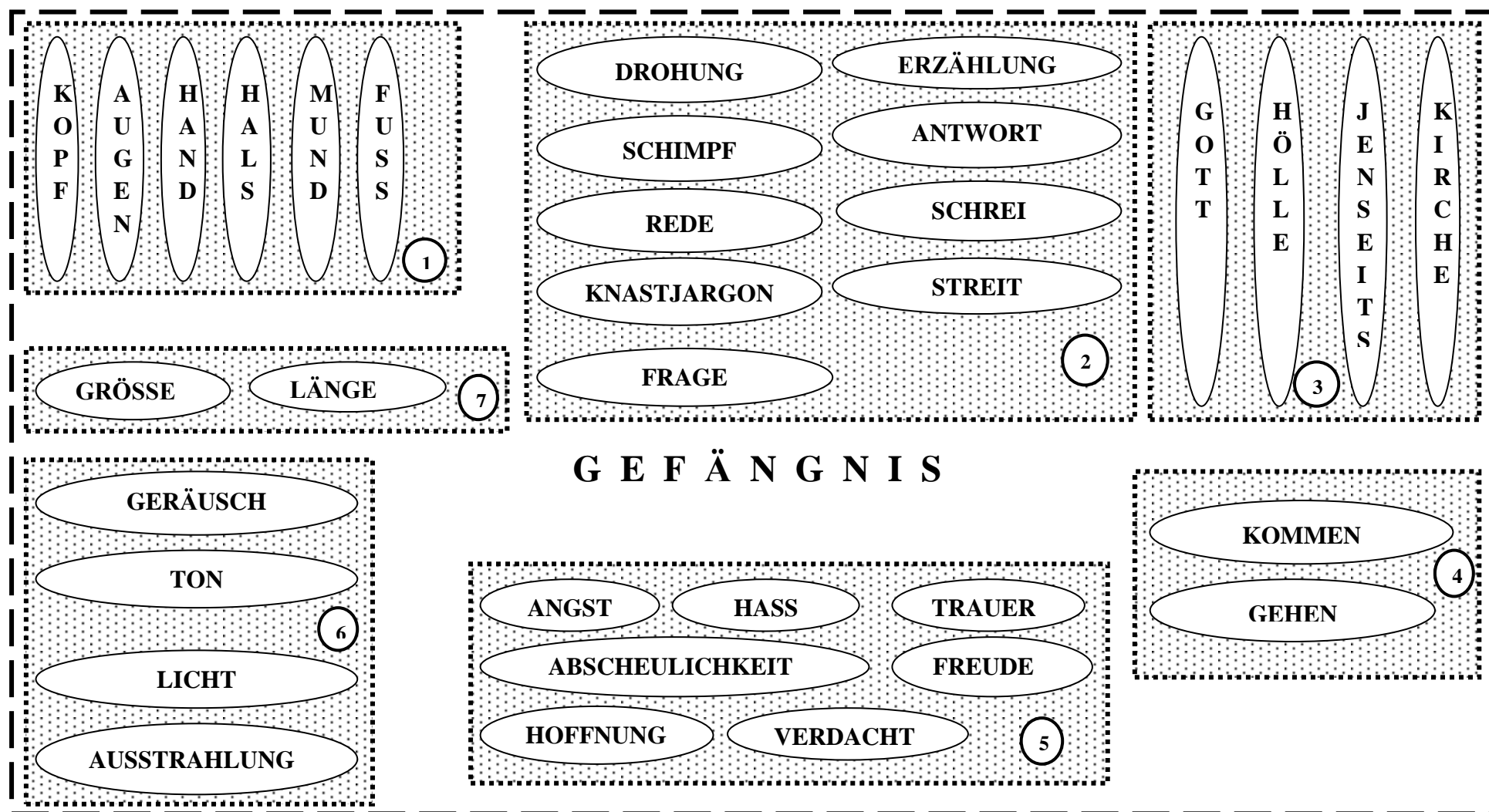
- - – metachthon   1 – domain 'human body'   4 – domain 'motion'   7 – domain 'physical characteristics of an object'

… – domain   2 – domain 'speech activity'   5 – domain 'emotions and feelings'

__ – autochthon   3 – domain 'Religion / supernatural'   6 – domain 'sound / light phenomena'

The cognitive map shows the relevance for a contemporary German author of a work of literature of information about the prison, a man as a biosocial being, whose prior activity is interpersonal communication, religion in the form of the accumulated knowledge of God and paranormal phenomena, human feelings and emotions, motion, light and sound phenomena, as well as the physical characteristics of the objects of the reality.

8. Concept system of SPD

The use of linguostatistical methods to analyze empirical material enabled to identify 7 significant domains (GEFÄNGNISSUBKULTUR, GESETZ UND RECHT, PHYSISCHE HANDLUNG, ZEIT, DENKTÄTIGKEIT, PSYCHISCHER ZUSTAND and PROZESS) and 42 concepts-autochthons within them that are actualized in the process of *live* criminal communication and form mental basis of the SPD in the conceptual sphere of the German ethnospace (see Table 1).

    The GEFÄNGNISSUBKULTUR domain includes a set of concepts that present the knowledge of prisoners about the prison subculture, which is "a system of distorted human values, rules and norms of behavior, traditions and customs adapted for themselves by the representatives of the criminal world" (Khisamutdinov 2015: 46). Domain in the SPD is represented by the nomenclature of concepts *KNASTTATOOS, KNASTJARGON, KNASTGEWALT, VERBRECHERHIERARCHIE, KNASTALLTAG, KNASTABTEILUNGEN and ANTI-GEWALT-PROGRAMM:*

(34)   *Der Klassiker – 5 PUNKTE: ALLEINE IN 4 WÄNDEN.*
        'That's classic – 5 POINTS: ALONE WITHIN FOUR WALLS.' (Würth & Born 2017)

(35)   *Die haben darauf geachtet, dass nur AUF DIE LEBER, AUF DIE MILZ, AUF DIE INNEREN ORGANE WURDE RICHTIG ABGEZIELT.*
        'They were paying attention to the place of hit, in particular the LIVER, SPLEEN, INTERNAL ORGANS.' (Panorama 2015)

(36)   *Die Männer, um die sie sich kümmert, gehören zu den gefährlichsten in ganz Deutschland: rabiate GEWALTVERBRECHER, GEISELNEHMER, VERGEWALTIGER, MÄNNER, DIE SICH AN KINDERN VERGANGEN HABEN.*
        'Men she takes care of are ones of the most dangerous in the whole Germany: fierce CRIMINALS, HIJACKERS, VIOLATORS, PAEDOPHILES.' (Würth 2016)

(37)   *Komm mal reine. HAFTRAUMKONTROLLE.*
        'Come in. CELL CHECK.' (Schenk & Frauke 2016)

(38)   *Gewalttäter, die nicht anders ruhig zu stellen sind, kommen in den so genannten BGH – den BESONDERS GESICHERTEN HAFTRAUM.*
        'Violators, which can't be calmed down, get to so called PLACE UNDER SPECIAL PROTECTION.' (Leunig 2016)

(39)   *Er hat ihn durch dessen ANTI-AGRESSION-THERAPIE begleitet.*

'He convoyed him during the ANTI-AGRESSION-THERAPY.' (Helm, 2017)

A significant area of knowledge relevant to persons in the institutions of confinement is the GESETZ UND RECHT domain, which has such conceptual content in the SPD (*STRAFTAT, STRAFPROZESSORDNUNG, FREIHEITSSTRAFE* and *WIEDEREINGLIEDERUNG:*

(40) *Dieses Mal wurde er wegen EINBRUCHS und DIEBSTAHLS, einer HANDFEUERWAFFE verurteilt.*
      'This time he was convicted for HOME BURGLARY with FIREARMS.' (Dinsenbacher 2016)

(41) *Doch die Grenzen zwischen den erlaubten freien BEWEISWÜRDIGUNG und der STRAFBAREN VORSÄTZLICHEN RECHTSBEUGUNG sind fließend.*
      'But the limits between the permitted free EVALUATION OF EVIDENCE and CRIMINAL INTENTIONAL VIOLATION OF LAW are conditional.' (Süddeutsche TV 2017)

(42) *Frank Keller wurde nach einem Geständnis zur lebenslangen HAFT verurteilt.*
      'After the confession in the commission of offence, Frank Keller was convicted to life IMPRISONMENT.' (Wendelmann 2016)

(43) *Dazu gehören kürzere Haftstrafen und mehr Gewicht auf die WIEDEREINGLIEDERUNG.*
      'This includes such punishment as an arrest for short term, and more attention is paid to the process of RESOCIALIZATION.' (Leunig 2016)

The PHYSISCHE HANDLUNG domain in the SPD covers information on the physical activity of prisoners in German institutions of confinement, where, being an integral part of prison life, it should facilitate the correctional reeducation of the criminal and increase his chances of successful reintegration into society after serving his sentence. Domain content is formed by concepts *SCHREIBEN, ANZIEHEN, TRAGEN, AUFMACHEN, NEHMEN, KÄMPFEN:*

(44) *Mach es gut, Digga. SCHREIB Briefe.*
      'Good luck, man. WRITE me letters.' (Leunig 2016)

(45) *Ich geh schnell Zähne putzen, ZIEH MICH AN, dann gehen wir los.*
      'I'll quickly brush my teeth, DRESS, and then we will go.' (Leunig 2016)

(46) *Die sind Jugendlichen verpflichtet, diese Sachen verdeckt zu TRAGEN.*
      'The youth must WEAR such things in secret.' (Leunig 2016)

(47) *MACH die Lücke AUF!*
      'OPEN the slot!' (Naber 2015)

(48) *Du hast doch den letzten GENOMMEN.*
      'It was you, who TOOK the last one.' (ZDFinfo Doku 2016)

(49)   *KÄMPFEN – das bin einfach ich.*
       'FIGHTING – it's all about me.' (SRF 2019)

The conceptual content of the ZEIT domain also appears in the German SPD. Time is a universal fundamental category of cultural anthropology, which in the national picture of the world of the Germans belongs to the main cultural values. The results of numerous cognitive studies show that in German society, where everything is based on the agreement concerning dates and observance of terms, the speakers try to measure and cover the temporal continuum with the help of technical means as accurately as possible, since the time in their consciousness is the magnitude that is in continuous communication with the dynamics of the processes taking place in society and the life of each of them, as well as one of the most valuable entities, because it is impossible to compensate for the time wasted.

As a result of the analysis of the empirical material it was found that the German prison is a special temporal dimension, where the only reference point in time is not the clock, but the key, since every morning the prisoner's day begins with the guard's opening the door to his cell and ends with its locking at night. It is the sound of getting a heavy key in the lock and a loud sound of double rotation of the key before the cell door opens that signals to the prisoner that the time of his imprisonment decreases. Here the time is measured not in the way usual for man – not in minutes, hours, etc., but in periods from letter to letter, from visit to visit, from the moment of imprisonment and to release from custody.

The ZEIT domain is formed by nuclear concepts to denote both metric (unit of measurement) and existential (indefinite duration) time. METRIC TIME explodes time concepts *JAHR, MONAT, STUNDE, WOCHE* and *TAG:*

(50)   *Der 18-Jährige wurde zu 65 JAHRE Haft verurteilt.*
       'The 18 years old was convicted to 65 YEARS of imprisonment.' (Leunig 2016)

(51)   *Wir haben einen MONAT lang in den Mädchen- und Jungengefängnissen gehemmt.*
       'We spent a MONTH in prisons for girls and boys.' (Schenk & Frauke 2016)

(52)   *Über Nacht werden die Häftlinge 8 STUNDEN lang eingeschlossen.*
       'During the night time prisoners are locked for 8 HOURS.' (Podknast 2015)

(53)   *Jede WOCHE kommen frisch verurteilte Straftäter.*
       'New criminals come every WEEK.' (Riedel 2016)

(54)   *Jeden TAG kommen die 48 Mädchen zum Mittagessen in die Kantine.*
       'Every DAY 48 girls come to the canteen for lunch.' (SWR 2017)

The EXISTENTIAL TIME is objectified in the SPD with concepts *MOMENT* and *WEILE:*

(55)   *Aber seine Wut kann sich jeden MOMENT in Gewalt entladen.*
       'But his fury may burst as violence at any MOMENT.' (Bremer 2014)

(56)   *Eine WEILE lief es richtig gut und sie dachte, jetzt kommt sie raus.*

'One INSTANT everything was really good and she thought she was going out of prison.' (Helm, 2017)

One of the key tasks of punishment in the form of imprisonment is the creation of preconditions for a criminal to get aware of his deed and for his repentance. Therefore, prisoners spend a lot of time in cells alone in order to rethink their existence. We include information about the mental processes of the prisoners into the domain DENKTÄTIGKEIT, the prevailing concepts of which are *WISSEN, BEGREIFEN, GEDANKE, VERSTEHEN, EINSTELLUNG and MERKEN*:

(57) *Ich WEIß nur, dass ich mit etwa 18 herauskomme. Mehr WEISS ich nicht.*
      'I only KNOW that I will go out of prison approximately at 18. I KNOW nothing more.' (Leunig 2016)

(58) *Mir ist BEWUSST, dass das hier ein Gefängnis ist.*
      'I REALIZE that here is a prison.' (ZDFinfo Doku 2016)

(59) *GLAUBST du wirklich, du kannst mir hier diese Scheiße erzählen?*
      'Do you really THINK that you can tell me this shit?' (ZDFinfo Doku 2016)

(60) *Aber das scheint keiner zu VERSTEHEN!*
      'But it seems like nobody UNDERSTANDS this.' (Leunig 2016)

(61) *Für wie bescheuert HÄLTST du mich eigentlich?*
      'How crazy you THINK I am?' (Panorama 2015)

(62) *Manchmal MERKT man in Gesprächen, dass sie die Sache nicht durchdacht haben.*
      'Sometimes during conversations it is NOTICED that they haven't thought this over.' (Riedel 2016)

In the process of analyzing the empirical material it was found that convicts do not tend to express their feelings during *live* communication. The proof of this is the low index of χ^2 criterion in the SPD of the PSYCHISCHER ZUSTAND domain, which contains knowledge of the prisoner's internal conditions in places of confinement. Concepts point to the state and changes in the psyche of the criminal while staying behind bars *(WUT, SCHOCK, NERVOSITÄT, JÄMMERLICHKEIT, GLAUBE, AUFREGUNG and REUE)*:

(63) *Die meisten hier sind voller WUT, WUT auf die Welt.*
      'The majority here is full of FURY, FURY to the world.' (ZDFinfo Doku 2016)

(64) *Ich bin immer noch ziemlich GESCHOCKT.*
      'I am still SHOCKED.' (Greine 2015)

(65) *Es NERVT total rum, dass ich mich verantwortlich und reumütig zeigen soll.*
      'It makes me so NERVOUS that I have to pretend as if I am responsible and I regret.' (NDR 2017)

(66)   *Es KOTZT MICH ECHT AN.*
        'This IRRITATES me.' (Kimmel 2014)

(67)   *Ich schätze, sie GLAUBEN daran, dass es aus mir ein besserer Mensch werden kann.*
        'It seems like they BELIEVE that I can be a better person.' (Schram 2014)

(68)   *Ich bin AUFGEREGT.*
        'I am ANXIOUS.' (Würth & Born 2017)

(69)   *Wir haben mit den Psychotherapeuten daran gearbeitet, dass sie ihre Tat BEREUT und begreift, was sie getan hat.*
        'Psychotherapeutists and I were working on them to REGRET about their conduct and realize what they have done.' (Greine 2015)

The most significant element of the conceptual portrait of the SPD was the PROZESS domain, which is the reservoir of knowledge of the convicts about the processes in their daily lives. Domain concepts are *EINFLUSS, STATTFINDEN, BESCHÄFTIGUNG, WERDEN and VERNICHTUNG:*

(70)   *Das kannst du BEEINFLUSSEN!*
        'You can INFLUENCE this!' (Leunig 2016)

(71)   *Wenn negative Dinge STATTGEFUNDEN haben, dann können aus diesen 3 Monaten durchaus Jahre werden.*
        'If the negative events HAPPENED, then 3 months might become years.' (Würth 2016)

(72)   *Mit den folgenden Erziehungsmethoden will er SICH nicht mehr BESCHÄFTIGEN.*
        'He doesn't want to be ENGAGED in the following educational measures.' (Spiegel TV 2014)

(73)   *Ich übernehme inzwischen die volle Verantwortung, aber es ist echt schwer, jemand zu WERDEN, der man nie gelernt hat zu sein.*
        'I take full responsibility for this, but it's really hard to BECOME the one you never learned to be.' (Leunig 2016)

(74)   *Wir werden ihn VERNICHTEN.*
        'We will DESTROY him.' (Brachwitz 2016)

Thus, the recorded autochthons are a discourse-making framework in the conceptual system of the German SPD, revealing the originality and uniqueness of the *live* interpersonal communication in the German institutions of confinement, which are the social center of the formation and functioning of prison discourse practice. On the basis of the selected domains and the nuclear concepts that are part of them, we model the conceptual portrait of the SPD in the form of a cognitive map (Figure 2).
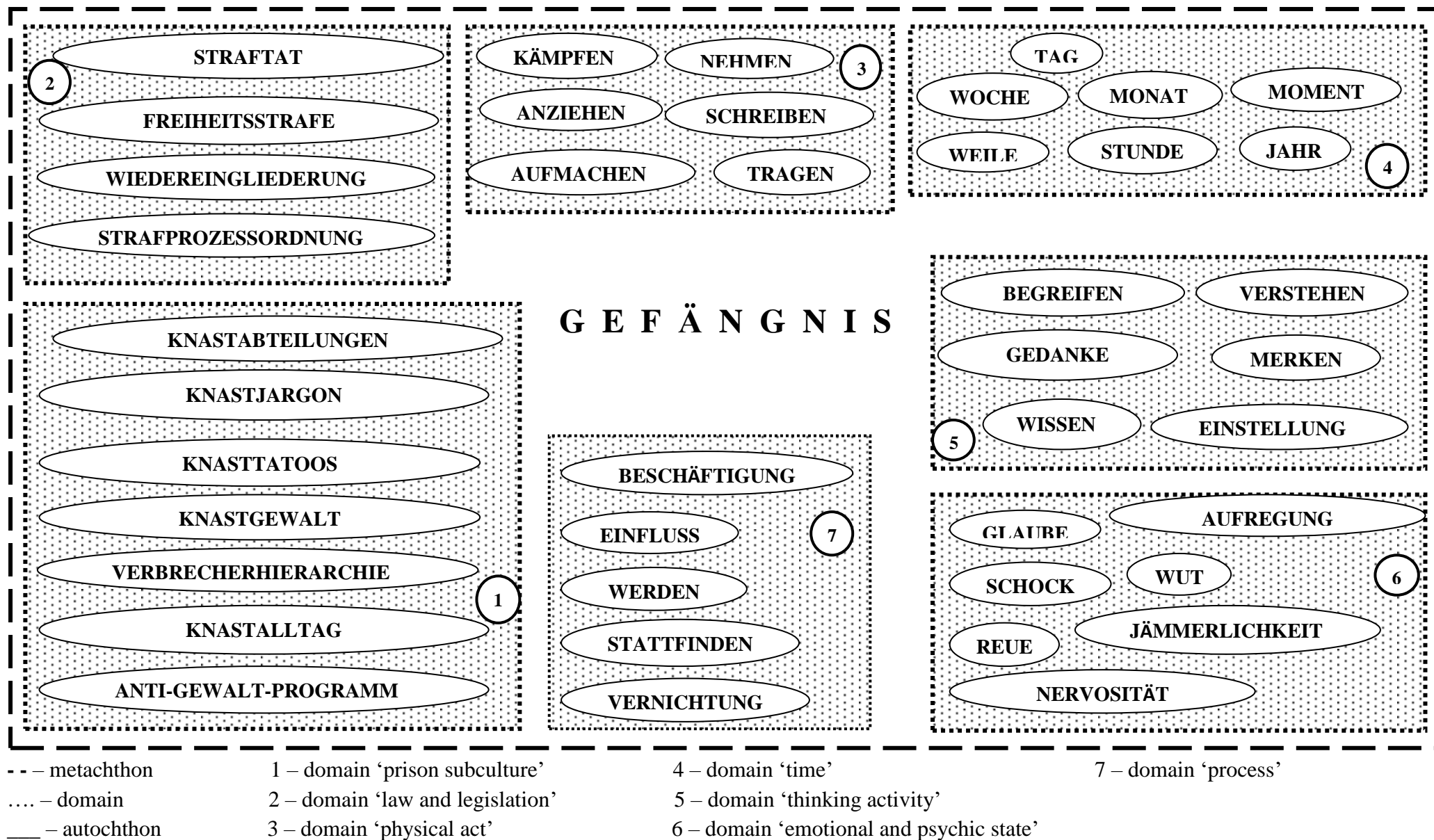
Figure 2. *Concept system of SPD*

The cognitive map shows relevance for a contemporary participant in a real German PD – prisoner or penitentiary system of prison – information on subculture, the main attributes of which are prison jargon, the German Code of Criminal Procedure, the legal basis for the activities of prison as a social institution, the law and the duties of prisoners in institutions of confinement, the everyday processes and physical activity of the prisoners, occurring in prison, the time, the cognitive processes of the convicts during his reeducation in institutions of confinement, as well as the mental state of the prisoners and prison staff during their life and activities in prison.

Comparing conceptual systems reconstructed for LPD and SPD and the author's and real prison-everyday pictures of the world represented in them, it is possible to assume that the literary form differs considerably from its real dimension, the conceptual spaces of LPD and SPD are fundamentally different, and the ideas reflected in the author's and the real prison-crime picture of the world, have almost nothing in common.

Thus, on the basis of the obtained statistical data (Table 1) it can be stated that the most significant domains in the LPD are MENSCHLICHER ORGANISMUS (χ^2=712,37; K=0,88); SPRECHTÄTIGKEIT (χ^2=588,4; K=0,81); RELIGION / PARANORMALE PHÄNOMENE (χ^2=579,12; K=0,81); BEWEGUNG (χ^2=464,56; K=0,8); EMOTIONEN UND GEFÜHLE (χ^2=309,13; K=0,62); LAUT- UND LICHTERSCHEINUNGEN (χ^2=113,68; K=0,4); PHYSIKALISCHE EIGENSCHAFTEN (χ^2=93,2; K=0,4). However, in SPD the following are significant: GEFÄNGNISSUBKULTUR (χ^2=947,59; K=0,95); GESETZ UND RECHT (χ^2=607,46; K=0,82); PHYSISCHE HANDLUNG (χ^2=574,91; K=0,81); ZEIT (χ^2=190,88; K=0,56); DENKTÄTIGKEIT (χ^2=170,32; K=0,48); PSYCHISCHER ZUSTAND (χ^2=88,24; K=0,32) and PROZESS (χ^2=85,84; K=0,48). It is obvious that the relevant domains do not demonstrate any correspondence: in the collective consciousness of German authors there is a perception that in the PD the general psychological knowledge of a person, peculiarities of his external and internal constitution and characteristic types of activities are prevailing, among which the central place is undeniably occupied by communicative activities.

Instead, the verbal space reflecting the mental processes of convicts in the process of *live* communication testifies to the thematic arrangement of knowledge about the prison as a hermetically closed world, the mirror of the whole society, which reflects its main problems (the consequences of the person's improper or lacking education, lack of his socialization and failed attempts to integrate, his growth within a certain subculture, as well as the consequences of living in a modern society of fierce competition, where satisfaction of the needs of one individual becomes more important than the welfare of the whole community. This results in social isolation, poverty, lack of a clear civic stand). Here the knowledge of the prison is presented as a form of punishment for the crime, as the epicenter of the 'prison subculture' and its impact on the emotional and mental state of the prisoner, as well as on the process of re-education of the convicted person in harsh conditions, and the cognitive changes that occur while the prisoners are staying behind the bars. The prison everyday fragment of the picture of the world is filled with the ideas of the SPD participants about the essence and features of the German penitentiary system, the concept of crimes, their classification and types of criminal punishment in the framework of current legislation.

Significant differences are also observed in the representation of the emotional and volitional spheres of the convicted person in the author's and real everyday prison picture of the world. Thus, in the views of German authors, the generalized knowledge about feelings and emotions as an integral part of the inner world of a person is fixed. The expressiveness of

the convicts' speech is represented by the emotions and feelings typical of a person as a representative of certain community, with no regard to the peculiarity of the prison as an institution of the criminal communication development, the impact of a specific microclimate on the organization and the relations between the staff and the imprisoned.

However, in the real everyday prison picture of the world, the fragment of the emotional and volitional spheres of communicants is presented in a different way. Here we state the low level of relevance of knowledge and ideas about the emotional and volitional spheres, only the domain EMOTIONAL-PSYCHOLOGICAL STATE gains communicative relevance, serving as a reservoir of knowledge of the convicts and prison staff about the state of being "newcomers", who find themselves in a criminal environment, the convicted "with experience", for whom prison is a kind of "advanced training courses", as well as legal agents. Empirical evidence confirms that prisoners and staff are always in the state of tension within the precincts of the German prison.

9. Conclusions

The research performed enables to state that the main conceptual constants of the literary and spoken norms of the PD representation are the segments of the specific concept systems. Having explained the structure of thought of the author of the work of literature and that of the real prisoners on the basis of the texts produced by them, we tried to determine the elements and concepts-autochthons in the structures of thought, which make up the matrix and frame of LPD and SPD.

The conceptual space of the LPD is represented by the domains SPRECHTÄTIGKEIT, MENSCHLICHER ORGANISMUS, RELIGION / PARANORMALE PHÄNOMENE, BEWEGUNG, GEFÜHLE UND EMOTIONEN, LAUT- UND LICHTERSCHEINUNGEN, PHYSIKALISCHE EIGENSCHAFTEN, the conceptual content of which proves the views of the German authors of the works of literature that the knowledge of man as a biosocial being, his inner world and spiritual space, in which the faith in God is essential, as well as of the types of human activity, the main of which is communicative activity, is relevant for the PD participants. At the same time, according to the authors, the least significant areas of knowledge used by the prisoners and employees of the penitentiary system are physical characteristics of objects of reality, sound and light phenomena.

The conceptual space of the SPD in the conceptual sphere of the German ethnospace is represented by the domains GEFÄNGNISSUBKULTUR, GESETZ UND RECHT, PHYSISCHE HANDLUNG, ZEIT, DENKTÄTIGKEIT, PSYCHISCHER ZUSTAND and PROZESS, the conceptual content of which shows that in the dimension of reality prisoners and prison staff, the knowledge about the prison subculture, the German Code of Criminal Procedure, time, as well as the emotional and psychological state of the convicted person and the cognitive operations occurring during the reeducation of the criminal, is sufficient.

Comparative analysis of the conceptual spaces of LPD and SPD allows to state that in the author's and the real everyday prison world pictures the prison communication fragment is represented by fundamentally different discourse-making concepts, the ensemble of which is a system of ideas and knowledge about the world of crime and prison.

The significant differences between the author's and the real everyday prison pictures of the world in the German language space is apparently due to the following: a) the author's

intentions to influence the spiritual space of the reader at superficial nature of his ideas and knowledge about the prison, because the researched area belongs to the passive zone of collective consciousness of the representatives of German ethnospace; b) the closedness of the institute of prison, its isolation from society, the careful concealment of what is happening behind bars by the leadership of the penitentiary institutions; c) increased orderliness of the representatives of the German-speaking community and their subconscious fear of being locked up among those whom the society considers outlaws; d) the low social status of the prison in the German-speaking space, which equates to the shame for life, of the stigma, which permanently deprives the convicted the right for successful reintegration into society after serving sentence in institutions of confinement; e) misrepresentations of argot as a language and crime as a subject of PD.

Supplementing cognitive maps with statistically verified interconceptual correlations that fix certain mental connections, as well as the use of the proposed method to other types of discourse, including the comparative linguocultural sphere can be mentioned as the prospects for further research in this perspective.

Abbreviations

LD for "literary discourse"
LPD for "literary prison discourse"
PD for "prison discourse"
SD for "spoken discourse"
SPD for "spoken prison discourse"

References

Ackermann, Rolf. 2015. *Das Geheimnis der Karawane*. München: dotbooks Verlag.

Adamzik, Kirsten. 1994. Beziehungsgestaltung in Dialogen. *Handbuch der Dialoganalyse*. 357–374.

Adlon, Andreas. 2012. *Ausradiert: Nicht ohne meine Tochter.* Köln: CreateSpace Independent Publishing Platform.

Adlon, Andreas. 2014. *Tod im Netz.* Köln: Kindle Edition.

Anonymus. 2009. *Das verbotene Buch*. München: riva Verlag.

Anzahl der Gefangenen und Verwahrten in den Justizvollzugsanstalten in Deutschland nach Bundesländern am 30. November 2017. (https://de.statista.com/statistik/daten/studie/-72216/umfrage/gefangene-und-verwahrte-in-justizvollzugsanstalten-nach-bundes-laendern/) (Accessed 2018-03-29.)

Aukett, Ewa. 2015. *Stunde der Drachen: Zwischen den Welten*. München: BookRix Verlag.

Babenko, Lyudmila & Kazarin, Yuriy. 2003. Lingvisticheskij analiz hudozhestvennogo teksta. *Teorija i praktika.* Moscow: Flinta-Nauka.

Bausch, Joe. 2012. *Knast*. Berlin: Ullstein Buchverlag.

Bekhta, Ivan. 2011. Suchasni napriamy vyvchennia khudozhnoho dyskursu. *Naukovi zapysky Luhanskoho natsionalnoho universytetu* 1(33). 244–255.

Benedikt, Caren. 2012. *Die Feinde der Tuchhändlerin*. München: Piper Taschenbuch Verlag.

Benedikt, Caren. 2015. *Die Rache der Duftnäherin*. München: Knaur Taschenbuch Verlag.

Brachwitz, Tom. 2016. *Eingesperrt: Leben hinter Gittern*. (https://www.youtube.com/watch?v=W9n4SODnz9A) (Accessed 2018-09-11.)

Bremer, Ulrike. 2014. *Die großen Kriminalfälle. Der Elternmord von Morschen*. (https://www.youtube.com/watch?v=Fwl1UmY2J6Y) (Accessed 2018-11-23.)

Budnik, Angela. 2016. Osoblyvosti vyvchennia dyskursu v pysemnii ta usnii komunikatsii. *Naukovyi visnyk Drohobytskoho derzhavnoho pedahohichnoho universytetu imeni Ivana Franka* 5(1). 39–42.

Burdiuhova, Nataliya. 2012. Prahmatychni osoblyvosti inversii v anhlomovnomu khudozhnomu dyskursi. Sotsialno-humanitarni aspekty rozvytku suchasnoho suspilstva 1. 16–17.

Chaplet, Anne. 2003. *Die Fotografin*. München: Antje Kunstmann Verlag.

Chaplet, Anne. 2003. *Schneesterben*. München: Antje Kunstmann Verlag.

Clemmer, Donald. 1968. *The Prison Community*. New York: Holt, Rinehart & Winston.

Croft, William. 1991. Syntactic categories and grammatical relations. Chicago: University of Chicago Press.

Diaz-Bone, Rainer & Krell Gertraude. (eds.). 2015. Diskurs und Ökonomie: Diskursanalytische Perspektiven auf Märkte und Organisationen. Wiesbaden: Springer-Verlag.

Dinsenbacher, Gabriele. 2016. *Vater Mörder: Das Erbe des Niklas Frank*. (https://www.youtube.com/watch?v=lF1gxoeKOZQ) (Accessed 2018-10-11.)

Ebert, Sabine. 2006. *Das Geheimnis der Hebamme*. München: Knaur Taschenbuch Verlag.

Ebert, Sabine. 2008. *Die Entscheidung der Hebamme*. München: Knaur Taschenbuch Verlag.

Ebert, Sabine. 2008. *Die Spur der Hebamme*. München: Knaur Taschenbuch Verlag.

Funke, Cornelia. 2005. *Tintenblut*. Hamburg: Cecilie Dressler Verlag.

Goffman, Erving. 1980. The characteristics of total institutions. A sociological reader on complex organizations. Canada: Holt, Rinehart & Winston.

Greine, Meike. 2015. *Experiment Jugendknast: Das Leben im Gefängnis*. (https://www.youtube.com/watch?v=fpUf_KSX_no) (Accessed 2018-11-02.)

Günther, Ralf. 2005. Seelsorge auf der Schwelle: Eine linguistische Analyse von Seelsorgegesprächen im Gefängnis. Göttingen: Vandenhoeck & Ruprecht.

Hartung, Alexander. 2014. *Bis alle Schuld beglichen*. Luxembourg: Amazon Publishing Verlag.

Helm, Ingo. 2017. *Gefängnis Doku 2017: Psychopathen im Knast!* (https://www.youtube.com/watch?v=LwKYEMVKDMk) (Accessed 2018-09-24.)

Hisamutdinov, Farid & Shalagin Anton. 2015. Kriminal'naja subkul'tura i ee preduprezhdenie. *Vestnik Kazanskogo Juridicheskogo instituta MVD Rossii* 2(20). 46–52.

Iesypovych, Kateryna. 2013. Fenomen kohnityvnoho kartuvannia v suchasnii linhvistychnii paradyhmi. *Studia linguistic* 7. 254–257.

Kibrik, Andrey. 2009. Modus, zhanr i drugie parametry klassifikacii diskursov. Moscow: Institut jazykoznanija RAN.

Kibrik, Andrey & Podlesskaja Vera. 2003. K sozdaniju korpusov ustnoj russkoj rechi: Principy transkribirovanija. Nauchno-tehnicheskaja informacija 6. 5–11.

Kimmel, Erika & Isecke Bernd. 2014. *Die großen Kriminalfälle. Monika Weimar und der Kindermord*. (https://www.youtube.com/watch?v=ERRSWfByxOA) (Accessed 2018-11-02.)

Klocke, Gabriele. 2004. *Über die Gleichheit vor dem Wort: Sprachkultur im geschlossenen Strafvollzug (Sozialtheorie)*. Bielefeld: Transcript Verlag.

Kol'cova, Lyudmila. 2007. Hudozhestvennyj tekst cherez prizmu avtorskoj punktuacii. Voronezh.

Kozhemjakin, Evgeniy. 2011. Lingvisticheskie strategii institucional'nyh diskursov. Sovremennyj diskurs-analiz. Lingvisticheskie izmerenija diskursa 3. 62–69.

Kriminalgericht Luzern: Frau muss wegen Tötung ihrer Zwillingsbabys ins Gefängnis am 30. Januar 2019. (https://www.srf.ch/) (Accessed 2019-01-30.)

Lejbnic, Gottfried. 2010. *Trudy po filosofii nauki*. Saint Petersburg: Librokom.

Levickij, Victor. 2012. *Semasiologija*. Vinnica: Nova Kniga.

Leunig, Christian. 2016. *Jung im Knast (2): Das Recht des Stärkeren*. (https://www.dailymotion.com/video/x6jk5f5) (Accessed 2018-08-26.)

Leunig, Christian. 2016. *Jung im Knast (3): Blinde Wut*. (https://www.youtube.com/watch?v=dIlrL897rLM) (Accessed 2018-09-19.)

Leunig, Christian. 2016. *Jung im Knast (4). Höchststrafe.* (https://www.dailymotion.com/video/x6jlwcl) (Accessed 2018-11-01.)

Link, Charlotte. 2012. *Im Tal des Fuchses.* München: Blanvalet Verlag.

Lirola, María Martínez & Rubio Fernando D. 2012. Foreign language as a social instrument in the European Higher Education Area. *Skase Journal of Theoretical Linguistics*, vol. 9, no. 1, pp. 59–74, 2012.

Makarov, Mihail. 2003. *Osnovy teorii diskursa.* Moscow: ITDGK Gnozis.

Meyrink, Gustav. 2007. *Der Golem.* Bremen: Carl Schünemann Verlag.

Milstein, Ilia. 2012. *Julia Timoschenko: Die autorisierte Biografie*. München: Redline Verlag.

Naber, Nils. 2015. *Härtester Knast in Deutschland.* (https://www.youtube.com/watch?v=n7PzR3ZzXAU) (Accessed 2018-07-20.)

Naumenko, Lyudmyla. 2003. Dilovyi dyskurs ta yoho funktsionalno-stylovi kharakterystyky. Kultura narodov Prychernomoria 37. 122–124.

Naumovych, Tetyana. 2016. Verbalizatsiia imeni kontseptu ZEIT / CHAS u nimetskomovnykh leksykohrafichnykh dzherelakh. *Naukovyi visnyk Mizhnarodnoho humanitarnoho universytetu* 21(2). 86–90.

NDR. 2017. *Gefängnis: Die größte Justizvollzugsanstalt Niedersachsens.* (https://www.youtube.com/watch?v=kGafJOTjMVY) (Accessed 2018-10-11.)

Neuhaus, Nele. 2009. *Unter Haien.* Münster: Prospero Verlag.

Ong, Walter. 2016. *Oralität und Literalität: Die Technologisierung des Wortes.* Wiesbaden: Springer-Verlag.

Panorama. 2015. *Gewalt hinter Gittern.* (https://thewikihow.com/video_A3mJghHvarM) (Accessed 2018-04-11.)

Riedel, Anke. 2016. *Justizirrtum: Unschuldig hinter Gittern.* (https://www.youtube.com/watch?v=-iRPHu0jUfQ) (Accessed 2018-10-22.)

Schenk, Ulrike & Frauke Siebold. 2016. *Frauen im Knast. Zwischen Hoffnung und Koller.* (https://www.youtube.com/watch?v=SQkW1aFs7Ak) (Accessed 2018-08-17.)

Schram, Peter. 2014. *Gewalt hinter Gittern: JVA Herford.* (https://www.youtube.com/watch?v=sCvg_RFgOjc) (Accessed 2018-02-21.)

Schumacher, Monika. 2013. Sprachgebrauch in der Deutschschweizer Justizvollzugsanstalt Pöschwies: Zum sprachlichen Umgang zwischen fremdsprachigen Gefangenen und den Anstaltsmitarbeitern im Spannungsfeld von sprachlicher und institutioneller Asymmetrie. Winterthur: ZHAW.

Sedov, Konstantin. 1999. Stanovlenie struktury ustnogo diskursa kak vyrazhenie jevoljucii jazykovoj lichnosti. Saratov.

Spaniel, Ellen. 2015. *Die Goldprinzessin*. München: dotbooks Verlag.

Spiegel TV. 2014. *Gesichter des Bösen: Teil 2.* (https://www.youtube.com/watch?v=ESalNJUkgSw) (Accessed 2017-02-16.)

Starkov, Oleg. 2010. *Kriminal'naja subkul'tura: Speckurs.* Moscow: Volters Klive.

Süddeutsche TV. 2017. *In den Fängen der Justiz.* (https://www.youtube.com/watch?v=gIrkyWzIpBM) (Accessed 2018-06-16.)

SWR. 2017. *Hinter verschlossenen Gittern.* (https://www.youtube.com/watch?v=_KTg9jmfji0) (Accessed 2018-04-12.)

Tipps für Erstinhaftierte am 29. August 2015 (http://www.podknast.de/index.php) (Accessed 2018-11-14.)

Tukaeva, Hulshat. 2009. Issledovatel'skij potencial kognitivnyx struktur. *Vestnik Bashkir. un-ta* 14(3). 861–865.

Viehöver, Willy & Keller, Reiner & Schneider, Werner. 2013. *Diskurs - Sprache - Wissen: Interdisziplinäre Beiträge zum Verhältnis von Sprache und Wissen in der Diskursforschung (Interdisziplinäre Diskursforschung).* Wiesbaden: Springer VS.

Wendelmann, Ulli. 2016. *Knast auf Ewig? Der Streit um die Sicherungsverwahrung.* (https://www.youtube.com/watch?v=iGlwblLKDBQ) (Accessed 2018-10-24.)

Würth, Nicole & Born Sigrid. 2017. *Knast in Deutschland (2). Herz, Schmerz, Sühne.* (https://www.zdf.de/dokumentation/zdfinfo-doku/knast-in-deutschland-herz-schmerz-suehne-102.html) (Accessed 2018-06-27.)

Würth, Nicole. 2016. *Knast in Deutschland (1). Strafe, Liebe, Hoffnung.* (https://www.zdf.de/dokumentation/zdfinfo-doku/knast-in-deutschland-strafe-liebe-hoffnung-102.html) (Accessed 2018-06-19.)

ZDFinfo Doku. 2016. *Jung im Knast: Geburtstag hinter Gittern.* (https://www.youtube.com/watch?v=pyaOywlAXS8) (Accessed 2017-08-30.)

Zhabotinskaja, Svetlana. 2013. Imja kak tekst: Konceptual'naja set' leksicheskogo znachenija (analiz imeni jemocii). *Kognicija, kommunikacija, diskurs* 6. 47–76.

Iryna Osovska
Department of Germanic, General and Comparative Linguistics
Faculty of Foreign Languages
Yuriy Fedkovych Chernivtsi National University
Sadova Street 5a, 58001, Chernivtsi
Ukraine
germmova@rambler.ru

Liudmyla Tomniuk
Department of Germanic, General and Comparative Linguistics
Faculty of Foreign Languages
Yuriy Fedkovych Chernivtsi National University
Sadova Street 5a, 58001, Chernivtsi
Ukraine
liudmyla_tomniuk@meta.ua

Meta-Intertext in German Parliamentary Discourse

Oleksandr Stasiuk, Taras Shevchenko National University of Kyiv, Ukraine
Galyna Stroganova, Kyiv National University of Trade and Economics, Ukraine

The paper aims to apply the theory of intertextuality to explore German parliamentary discourse. It investigates the usage of meta-intertext in parliamentary texts. Meta-intertext is considered as fragments of a text that accompany intertext and describe it. It is determined that meta-intertext contains markers which provide information about the author of hypotext, the hypotext type, intertext evaluation, time and place of hypotext production, its title and its recipient. The use of these markers in German parliamentary texts is analyzed and described. The syntagmatic structure of meta-intertext is revealed. It is established that meta-intertextual markers describing the author of hypotext are usually placed at the beginning, whereas evaluative markers – at the end of the fragment summarizing it.

Keywords: *intertext, intertextuality, meta-intertext, parliamentary discourse, quotation.*

1. Introduction

In recent years, there has been an increasing interest of linguists in institutional communication. One of its most important branches in the countries with parliamentary democracies is communication in parliaments. The parliamentary discourse is considered to be the main part of the political discourse in these countries. Communication is essential for parliamentarians. Cornelia Ilie (2015: 1114) states that "distinguishing characteristics of parliaments as institutions is that parliamentary work essentially consists of speaking (monologic communication) and debating (dialogic communication)".

The communication processes and the language of the German Bundestag is described in a number of papers (e.g. Burkhardt (2003); Dörner & Vogt (1995)). Researchers claim that German parliamentary communication is double addressed. Explicitly, parliamentarians talk to their colleagues in the chamber, trying to convince them. But the true addressee of the message is the electorate. Thus, the primary communication goal of parliamentarians is to present themselves and to discredit their political opponents.

Linguists try to ascertain in what way the language of parliament influences the audience, and "to explore the recurring linguistic patterns and rhetorical strategies used by MPs that help to reveal their ideological commitments, hidden agendas, and argumentation tactics" (Ilie 2015: 1113).

One of the linguistic devices extensively used in the parliamentary discourse due to its pragmatic value is intertext. A number of studies were conducted on the use of intertext in various parliamentary texts (Berrocal 2016; Constantinescu 2012; Gruber 2012).

Numerous quotations can be found in texts of parliamentary debates. Parliamentarians quote their colleagues, scientists, journalists, citizens talking about political issues etc. As Kolesarova (2008: 33) points out, quotations are used by parliamentarians to support their opinions. It can be argued that the role of quotations in the parliamentary discourse is much wider, and parliamentarians can, for example, quote their political opponents, trying to discredit them.

Despite the great importance of intertextuality in German parliamentary discourse, the phenomenon still remains relatively unexplored. In our study, we are trying to fill this gap by addressing the concept of meta-intertextuality.

2. Terminology issue

2.1 *Parliamentary discourse*

To discuss the use of meta-intertext in the parliamentary discourse, it is necessary to introduce and define some basic terms that will be used further in this paper.

Trying to explore the structures of human communication that are larger than a text, we inevitably come to the notion of discourse. Being an extremely polysemic term, one of the main meanings of *discourse* is a dialogue of texts. Heinemann M. & Heinemann W. (2002: 61) consider discourse as "more or less fast set of texts". Jorgensen & Phillips (2002: 1) remark that "underlying the word 'discourse' is the general idea that language is structured according to different patterns that people's utterances follow when they take part in different domains of social life". Discourse is also considered to be "a form of social practice" (Fairclough & Wodak 1997: 258). Thus, parliamentary discourse, which we discuss in this paper, is considered as a set of texts reflecting social practice in parliament. Parliamentary texts are interconnected by quotations, paraphrases and allusions and, in this way, intertextuality contributes to constituting the parliamentary discourse.

The conducted research does not aim to cover all the text types which can be attributed to the parliamentary discourse. It concentrates on debate texts in plenary meetings of the German Bundestag. These texts can be considered as the central and the most important part of all communication in the parliament. Debates in plenary meetings include different text types, such as speeches of members of the parliament, interjections, speeches of ministers in the parliament, government declarations, questions of parliamentarians to ministers and answers of ministers to the questions of parliamentarians etc.

2.2. *Text*

After determining the scope of parliamentary texts the research will be based on, it is important to define the term *text* in such a way that will give the possibility to clearly distinguish one text from another.

Text is the main term of text linguistics, but it does not have any generally accepted definition (Adamzik 2004: 31). In this paper text is understood as a set of utterances that belong to the same author and are expressed in the same communicative situation. In accordance with this definition, parliamentary debates refer not to a single text, but to a set of texts having intertextual relations with each other and with texts from outside the parliamentary discourse. Parliamentary speech is considered to be a single text, but it can be interrupted by interjections of other parliamentarians. Replies to these interjections can be defined as separate texts distinct from the parliamentary speech.

2.3 *Intertext*

The next term that requires a closer look is intertextuality. The term *intertextuality* was

introduced by Julia Kristeva (1980) in 1960s, but the idea of dialogue of texts is to find by Bakhtin (1981). Over the past few decades, intertextuality has attracted the attention of scholars from various disciplines (Hermann & Hübenthal 2007: 7). Intertextuality has been actively studied by researchers in the field of literature and culture, but it has remained relatively unexplored by linguists (Panagiotidou 2011: 173). Despite its broad use the term intertextuality is neither transparent nor commonly understood (Allen 2000: 2).

Text linguistics aims to describe the linguistic nature of intertextuality. From the point of view of linguistics "intertextuality refers to the phenomenon that other texts are overtly drawn upon within a text, which is typically expressed through explicit surface textual features such as quotations and citations" (Wu 2011: 97). Gérard Genette defines intertextualty in a more restrictive way compared to Julia Kristeva as "the actual presence of one text within another" (Genette 1997: 2). In this paper, we use this more restrictive understanding of intertextuality, because it offers the possibility to detect the connection between texts using linguistic methods.

In text linguistics referential and typological intertextuality are distinguished (Klein & Fix 1997: 34; Antos & Tietz 1997: 97). Referential intertextuality implies the use of fragments of certain texts, while typological intertextuality refers to the use of patterns and structures typical of the text type. This research concentrates on the referential intertextuality.

While discussing the referential intertextuality, we always deal with two different texts: a preceding text and a following text. In this case, the concepts of hypertext and hypotext introduced by Gérard Genette are used. Hypertextuality means "any relationship uniting a text B (hypertext) to an earlier text A (hypotext), upon which it is grafted in a manner that is not that of commentary" (Genette 1997: 5). Thus, in this paper we try to describe in what way hypotext is introduced in the hypertext.

2.4 *Meta-intertext*

Referential intertextuality can be marked and not marked (Broich & Pfister 1985: 31). When not marked, or implicit, intertext is used, the author may not even know that she or he is employing intertext, or may not wish to show that a piece of her or his text is taken from another source.

Marked, or explicit, intertextual fragments contain linguistic markers that describe intertext. The pieces of text with such markers are called *meta-intertext* (Gavenko 2011: 38–39; Stasiuk 2016: 336–337). The term *meta-intertext* is a fusion of two terms used in text linguistics – intertext and metatext.

The term *metatext*, introduced in 1971 by Anna Wierzbicka, is considered as a statement, the subject of which is the text itself (see e.g. Witosz 2017: 108). Meta-intertext is similar to metatext in that it also describes a text. But unlike the metatext, it describes the quoted text, or hypotext. Thus, we define meta-intertext as a statement that describes the intertext. Meta-intertext identifies the hypotext and provides various information about it, and not about the current text.

The purpose of the presented research is to describe meta-intertext as the way of introducing intertextual fragments in German parliamentary discourse.

3. Material and methodology

The conducted research was based on the analysis of fragments of parliamentary texts of the German Bundestag. These fragments contain intertext in form of quotations, and meta-intertext. As noted by Mac & Szwed (2016: 269), it is often difficult to distinguish between quotes and similar types of intertextuality, such as allusion and paraphrase. In written texts quotations are marked with quotation marks or as block quotations. In this paper, all quotations marked in the stenographic protocols of plenary meetings were selected for the analysis. A total of 590 text fragments were analyzed.

To extract meta-intertextual markers from the meta-intertext, a semantic analysis was carried out. Extracted markers were classified in accordance with their meaning.

The principle of meta-intertextual markers differentiation in German parliamentary discourse can be illustrated by the following text fragment:

(1)   *Christine Buchholz (Die Linke): In ihrem Koalitionsvertrag sprechen SPD und Union davon, die – Zitat – "globale Ordnung aktiv mitgestalten" zu wollen.*
        'Christine Buchholz (The Left): In their coalition agreement SPD and Union speak of – quotation – their wish "actively participate in shaping the global order".'
        (Deutscher Bundestag 2013–2014: 84)

Fragment (1) contains the intertext (*globale Ordnung aktiv mitgestalten* 'participate in shaping the global order') and the meta-intertext (*in ihrem Koalitionsvertrag sprechen SPD und Union davon, […] – Zitat* 'in their coalition agreement SPD and Union speak of, […] – quotation') that describes the intertext. The meta-intertext contains markers of following types:

     1. Quotation marker (*Zitat* 'quotation').
     2. Author of hypotext (*SPD und Union* 'SPD and Union').
     3. Hypotext type (*Koalitionsvertrag* 'coalition agreement').

Different types of intertextual markers and the frequency of their use in German parliamentary discourse are shown in Table 1.

Table 1: Frequency of meta-intertextual markers use

Type of intertextual marker	Number of examples	Frequency
Author of hypotext	356	60.3%
Type of hypotext	303	51.4%
Time of hypotext production	92	15.6%
Quotation marker	90	15.3%
Positive evaluation of intertext	77	13.1%
Negative evaluation of intertext	57	9.7%
Place of hypotext production	12	2.0%
Recipient of hypotext	7	1.2%
Text title	6	1.0%

Total	590	100%

In order to determine the sequential structure of meta-intertextual markers, syntagmatic analysis was performed. We have tried to define which markers are usually placed before intertext and which markers are usually placed after it. The sequential structures demonstrate a great variety (more than 100 different sequences of meta-intertextual markers have been detected). That is why only the most frequent sequential structures of meta-intertextual markers are shown in Table 2:

Table 2: Sequential structures of meta-intertextual markers

Sequential structures of meta-intertextual markers	Number of examples	Frequency
Type of hypotext – Intertext	110	18.6%
Type of hypotext – Author of hypotext – Intertext	69	11.7%
Author of hypotext – Intertext	68	11.5%
Intertext (quotations without meta-intertextual markers)	62	10.5%
Author of hypotext – Intertext – Evaluation of intertext	16	2.7%
Author of hypotext – Time of hypotext production – Intertext	14	2.4%
Intertext – Evaluation of intertext	10	1.7%
Author of hypotext – Type of hypotext – Intertext	9	1.5%
Author of hypotext – Type of hypotext – Intertext – Evaluation of intertext	6	1.0%
Other sequential structures	226	38.4%
Total	590	100%

In fragment (1), we can observe the sequential structure of type "Quotation marker – Type of hypotext – Author of hypotext – Intertext". This structure is not widely used, but it is similar to the structure "Type of hypotext – Author of hypotext – Intertext" which occurs much more frequently as it is shown in Table 2.

4. Results and discussion

4.1 *Author of hypotext*

Among all markers of intertextuality used in German parliamentary discourse, the most frequent one is the marker of author of hypotext. This marker serves to name the person or the

organization that has produced the quoted text, and gives additional information about the author.

The author of hypotext is frequently referred to by her or his given name and surname: *Heinz Riesenhuber; Nelson Mandela; Wolfgang Thierse*. As a rule, this way of presenting the author of hypotext is used in the cases when she or he is a politician (usually a member of the parliament). Other well-known people are also referred to in this way: *Christa Wolf* (German novelist, literary critic and essayist)*; Carl Schmitt* (German jurist and political theorist)*.

Similar is the case when only surname of the author of hypotext is given, accompanied by the address form *Frau* 'Mrs' or *Herr* 'Mr': *Frau Nahles; Herr Pofalla; Herr de Mazière*. These people are also usually well known to parliamentarians.

Another way to present the author of hypotext is to name her or his position: *Präsident des Bundesverfassungsgerichts* 'the President of the Federal Constitutional Court'; *Herr Bundespräsident* 'Mr Federal President'. The social position can be combined with a name: *Staatsminister Michael Link* 'Minister of State Michael Link'.

In German parliamentary discourse numerous quotations from mass media are used. In this case, the most important thing is the newspaper or magazine that has published the article, and not the journalist who has written it. That is why in meta-intertext we encounter markers that are names of newspapers or magazines the intertext is taken from: *Spiegel; Spiegel Online*.

In meta-intertext we can also find organization names, e.g. *Ratingagentur Standard & Poor's* 'rating agency Standard & Poor's'; *der Bundesverfassungsgericht* 'the Federal Constitutional Court'; *der Familienbund der Katholiken* 'the Family Union of the Catholics'; *die Deutsche Bank* 'the Bank of Germany'; *das Bundesfamilienministerium* 'The Federal Ministry for Families'.

To refer to the author of hypotext, personal pronouns *Sie* 'you' or *wir* 'we' can be used. Personal pronoun *Sie* is used when the parliamentarian addresses a politician who is present in the chamber and quotes her or him at the same time. Another case is when the speaker addresses a group of politicians, e.g. a political fraction in the parliament or governing coalition.

(2)   *Paul Lehrieder (CDU/CSU): Im dritten Absatz Ihrer Begründung schreiben Sie: "Besonders betroffen von Stundenlöhnen unter 10 Euro sind in Deutschland Frauen".*
'Paul Lehrieder (CDU/CSU): In the third paragraph of your statement you write: "Especially affected by salaries lower than 10 euro hourly are in Germany women".' (Deutscher Bundestag 2013–2014: 174)

In (2), the meaning of the personal pronoun *Sie* is *the Member of the Left Party* and the bill of this political party is discussed.

The personal pronoun *wir* 'we' is used in a similar way:

(3)   *Kerstin Grise (SPD): Wir haben in unserer Koalitionsvereinbarung, so sie denn von unseren Mitgliedern unterstützt werden wird, dazu eine schöne Aussage, die ich Ihnen zitieren will: "Die Herausbildung einer europäischen Zivilgesellschaft ist eine essentielle Voraussetzung für eine lebendige europäische Demokratie. Besonders wichtig ist es, dafür auch die Jugendpolitik weiterzuentwickeln. Das sind zwei der guten Sätze in diesem Koalitionsvertrag.*
'Kerstin Grise (SPD): We have in our coalition agreement, so as it will be supported by our members, a nice expression about it, which I want to cite to you: "The creation of a European civil society is an essential prerequisite for a

living European democracy. Especially important is it to also continue developing the youth policy". These are the two good sentences in this coalition agreement.' (Deutscher Bundestag 2013–2014: 189)

In (3), the first-person plural pronoun *wir* 'we' is used. *Wir* is a pronoun that changes its meaning depending on the context. In this case it refers to CDU/CSU and SPD – the two governing parties in the German Bundestag that have signed the coalition agreement.

In addition to the above-mentioned cases, meta-intertext can provide additional information about the author of hypotext, as illustrated in the following example:

(4) *Petra Sitte (Die Linke): Niemand Geringerer als der Bundestagspräsident selbst hat es in seiner Antrittsrede klargestellt – ich zitiere –: "Und selbstverständlich bedarf eine geschäftsführend amtierende Bundesregierung nicht weniger parlamentarischer Kontrolle als eine neu gewählte". Und weiter in seiner Rede: "Niemand wird deshalb ernsthaft erwarten dürfen, dass der Bundestag seine Arbeit erst nach Abschluss der Koalitionsverhandlungen aufnehmen wird". Die Linke sieht das genauso. Deshalb haben wir den Antrag gestellt, und deshalb ist uns Ihr Verhalten völlig unverständlich.*
'Petra Sitte (The Left): No less a person than the President of the Bundestag has declared it in his introduction speech – I quote –: "And of course the Federal Government requires not less parliamentary control than the newly elected one". And further on in his speech: "Nobody will seriously expect that the Bundestag will start its work only after the end of coalition negotiations." The left party sees it in the same way. That is why we have proposed this motion and that is way your behavior is completely incomprehensible.' (Deutscher Bundestag 2013–2014: 70)

Fragment (4) begins with meta-intertext that is used to show the importance of the author of hypotext in political life of Germany. Pointing out the significance of the author of the quoted text gives more value to the words expressed in the parliamentary speech.

Markers pointing to the author of the hypotext are in most cases placed before the intertext at the beginning of the meta-intertextual fragment. Parliamentarians first present the person, and then quote her or him.

The importance of choosing the right person for quoting and presenting him or her in an appropriate way in the meta-intertext can be observed in the following example:

(5) *Sahra Wagenknecht (Die Linke): Für den Fall, dass Sie mir nicht glauben, zitiere ich den Inhaber des Lehrstuhls für Finanzierung und Kreditwirtschaft an der Ruhr-Universität Bochum, Professor Stephan Paul: "Der jetzige Richtlinienentwurf macht die Tür auf und ermöglicht es, für – so heißt es dort – "gesunde Banken mit tragfähigem Geschäftsmodell" auch vorher schon – also vor der Haftung von Eigentümern und Gläubigern – Hilfszahlungen von staatlicher Seite zu gewähren. Ich frage mich an der Stelle aber: Wenn eine Bank gesund ist, wenn sie ein tragfähiges Geschäftsmodell hat, wozu braucht sie dann staatliche Hilfszahlungen"? So weit Professor Paul.*
'Sahra Wagenknecht (The Left): For the case that you don't trust me I quote the Head of the Chair of Banking and Finance at the Ruhr University Bochum, professor Stephan Paul: "The current motion opens the door and makes it possible for – as it is called there – "healthy banks with reliable business model" also before – that means

before the responsibility of owners and creditors – to be supported by the state. But I ask me at this point: When a bank is healthy, when it has a reliable business model, why does it need to be supported by the state"? So far professor Paul.' (Deutscher Bundestag 2013–2014: 244)

In (5), the parliamentarian explicitly points out that he is probably not trusted. That is why he has to quote a person that will be trusted. He presents the person quoted as *Inhaber des Lehrstuhls für Finanzierung und Kreditwirtschaft an der Ruhr-Universität Bochum, Professor Stephan Paul* 'the Head of the Chair of Banking and Finance at the Ruhr University Bochum, professor Stephan Paul'. The author of hypotext is presented as a scientist and as an independent person who can be considered as impartial, on the one side, and as a person that is of a high qualification and can make judgments about economic issues, on the other.

4.2 *Hypotext type*

Another marker type that is used in German parliamentary discourse is marker of hypotext type. This marker points to the text being quoted rather than to its author. But instead of giving specific text title, the fragment of meta-intertext mostly contains a name that can be applied to a range of similar texts.

The following names represent text types according to their belonging to a particular discourse:
– political (*Tagesordnung* 'agenda'; *Stimmkarte* 'voting card'; *Zwischenruf* 'interjection'; *Antrittsrede* 'inaugural address'; *Rede* 'speech'; *Frage* 'question' (of the parliamentarians to the ministers); *schriftliche Antwort* 'written answer' (of the ministers to the parliamentarians); *Verordnung* 'regulation'; *Ratsverordnung* 'council regulation'; *Minderheitsvotum* 'minority vote'; *Schreiben* 'letter'; *Eid* 'oath');
– juridical (*Urteil* 'judgement'*; Gesetzentwurf* 'bill');
– mass media discourse (*Interview* 'interview');
– literary discourse (*Roman* 'novel'; *Lied* 'song').

If meta-intertext contains the marker of text type, it is in most cases still possible to understand what specific text is meant. For example, the name *Grundgesetz* 'constitution' means the constitution of Germany. *Tagesordnung* 'agenda' means the document of the actual day and *Koalitionsvertrag* 'coalition agreement' means the document of the current legislation period authored by CDU/CSU and SPD.

In many cases, we get some additional information that shows what part of the text is being quoted: *Tagesordnungspunkt 1* 'agenda item 1', *Art. 38 unseres Grundgesetzes* 'Article 38 of our Constitution'.

Markers indicating the hypotext type are often combined with other markers contained in the meta-intertextual fragment. They are often used together with markers describing the author of hypotext. Meta-intertextual markers of text type are usually placed before the intertext.

4.3 *Evaluation of intertext*

Meta-intertext contains expressions that are used to evaluate the hypotext in a positive or negative way. They do not identify the hypotext, but express the author's attitude to it. The obtained results show that approximately equal number of cases with positive and negative

evaluation is used, as shown in Table 1.

Positive evaluation is expressed through adjectives with positive evaluative meaning: *schön* 'beautiful'; *richtig* 'right'; *klar* 'clear'; *gut* 'good'; expressions *zu Recht* 'rightly'; *völlig zu Recht* 'completely rightly'; *Recht haben* 'to be right'; *Recht geben* 'to agree with'.

Besides the positive evaluation, parliamentarians often use the negative evaluation of intertext. It can be expressed through combination of adjectives with positive meaning like *richtig* 'right' with the negation *nicht* 'not'. Adjectives with negative meaning are also used, e.g. *falsch* 'false'; *zynisch* 'cynical'. Other expressions used for negative evaluation are *wackelige Formulierungen* 'wobbly expressions'; *erstaunlich* 'striking'; *Unsinn* 'nonsense'; *völlig weltfremd* 'totally unworldly'.

Negative evaluation can be emphasized by using adverbs and repeating adjectives with negative evaluative meaning: *Das ist definitiv falsch. Das ist eine falsche Behauptung*. 'This is definitively wrong. This is a wrong statement.'

Negative evaluation can be expressed through pointing out the euphemistic nature of the expression used in the intertext. Parliamentarians reveal the words used by their political opponents.

(6)   *Eva Bulling-Schröter (Die Linke): Wie steht es so schön im Koalitionsvertrag: Begrenzung der „Kostendynamik beim Ausbau der erneuerbaren Energien", und man will „der Entwicklung der konventionellen Energiewirtschaft einen stabilen Rahmen" geben. Was hinter dieser Verklausulierung steht, ist ja wohl klar: Der Ausbau der Erzeugung von Energie aus Wind und Sonne wird gebremst; schwarz wie Kohle ist die Zukunft.*
'Eva Bulling-Schröter (The Left): As it is written so beautifully in the coalition agreement: limitation of "cost dynamics by the development of renewable energies", and it is desired to give "a stable framework for the development of economy of conventional energy". It is clear what is hidden behind these formulations: The development of energy production from wind and sun will be hampered; black as coal is the future.' (Deutscher Bundestag 2013–2014: 113)

In fragment (6), we observe positive evaluation of the text written in the coalition agreement that is followed by pointing to the euphemistic nature of these words.

There is a close relation between the evaluation of the intertext by the author of text and evaluation that is made inside the intertextual fragment.

(7)   *Katja Dörner (Bündnis 90/Die Grünen): Da kritisiert der Familienbund der Katholiken völlig zu Recht, dass die neue Bundesregierung „mit einem Wortbruch in die neue Legislaturperiode" startet.*
'Katja Dörner (Alliance 90/The Green Party): Here the Family Union of the Catholics criticizes completely rightly that the new Federal Government starts "the new legislative period breaking its oath".' (Deutscher Bundestag 2013–2014: 96).

In fragment (7), a positive evaluation of intertext by the author of text is observed. At the same time the author points out that the intertext contains critical attitude to his political opponents.

4.4 *Time of hypotext production*

Meta-intertext contains markers that point to the time when the quoted text was produced. This type of markers is usually placed after the markers of author of hypotext, but before the intertext, as in the fragment below.

(8)   *Kerstin Andreae (Büdnis 90/Die Grünen): Die Kanzlerin hat gestern gesagt, „dass wir uns unvermindert anstrengen müssen", um in die Zukunft zu investieren.*
        'Kerstin Andreae (Alliance 90/The Green Party): The Chancellor said yesterday "that our efforts must continue unabated" in order to invest in the future.' (Deutscher Bundestag 2013-2014: 704)

In (8) we have marker of author of hypotext *Die Kanzlerin* 'The Chancellor' that is directly followed by the marker of time when the intertext was produced *gestern* 'yesterday'. After that comes the intertext.

The linguistic ways to represent markers of time in the meta-intertext are quite different. In some cases, markers of time are expressed through dates or expressions like *im Juli* 'in July' (that means in July this year); *vor vier Jahren* 'four years ago' (a legislative period ago)*; gestern* 'yesterday'; *vor einem Jahr* 'a year ago'; *Antwort vom 14. November* 'answer from the 14th of November'; *vor elf Monaten* 'eleven months ago'; *2012; im Mai 1996* 'in May 1996'; *Schreiben vom heutigen Tage* 'letter from today'; *vor zehn Jahren* 'ten years ago'; *einmal* 'once'; *in diesem Jahr* 'this year'; *seither* 'from that time'; *in den letzten Tagen* 'last days'.

In other cases, time markers are expressed related to some important events, as shown in the following example.

(9)   *Hans-Christian Ströbele (Bündnis 90/Die Grünen): Sie sind aus den USA zurückgekommen und haben gesagt – soll ich es Ihnen vorlesen? –: Alle Vorwürfe haben sich „in Luft aufgelöst". – Ich habe immer geguckt, weil das schon damals nicht richtig war.*
        'Hans-Christian Ströbele (Alliance 90/The Green Party): You came back from the U.S. and said – shall I read it to you? –: All accusations "vanished into thin air". – I have always looked upon this because it was already not right then.' (Deutscher Bundestag 2013–2014: 51)

In fragment (9), the intertext relates to the time after visiting the U.S. by the Chancellor of Germany Angela Merkel.

4.5 *Quotation marker*

Parliamentarians also often mark explicitly that they use quotations. In this case, the noun *Zitat* 'quotation' or the verb *zitieren* 'quote' is used.

The markers of quotation are usually placed directly before the intertext. In some cases, parliamentarians mark quotation twice: directly before and after the intertext.

(10)   *Norbert Lammert (Präsident des Bundestages): Zitat Andreas Voßkuhle: "Der Bundestag ist und bleibt der Ort, an dem die wesentlichen Entscheidungen für unser Gemeinwesen getroffen werden müssen". Ende des Zitats.*
        Norbert Lammert (The President of the Bundestag): Quote Andreas Voßkuhle: 'The Bundestag is and remains the place where the important decisions for our community

have to be made". Unquote.' (Deutscher Bundestag 2013–2014: 7)

This marker type shows that it is important for parliamentarians to clearly demonstrate that these words belong not to her or him, but to another person that can be judged as impartial and enjoys much more trust. That is why the markers that point to quotation are usually combined with markers that provide us with information about the author of text. Many of these authors are not politicians but journalists, scientists or religious figures.

4.6 *Other types of intertextual markers*

Some other types of meta-intertextual markers are used in German parliamentary discourse in combination with markers of author or hypotext type.

Meta-intertext contains markers of place, though they are not as numerous as the markers of time.

(11) *Norbert Lammert (Präsident des Bundestages): "Zur Verantwortungsübernahme durch das Parlament gibt es keine überzeugende Alternative". So hat es der Präsident des Bundesverfassungsgerichts, den Heinz Riesenhuber schon auf der Tribüne begrüßt hat, nicht nur in Interviews immer wieder festgehalten, sondern auch in einschlägigen Urteilen des Bundesverfassungsgerichts ist das so oder ähnlich nachzulesen.*
'Norbert Lammert (The President of the Bundestag): "To taking responsibility by the parliament there is no convincing alternative". So has the President of the Federal Constitutional Court, who has already been welcomed by Heinz Riesenhuber from the rostrum, pointed out, again and again, not only in the interviews, but also in the relevant judgments of the Federal Constitutional Court is it to read so or in a similar way.' (Deutscher Bundestag 2013–2014: 7)

In example (11) we observe a marker of place *auf der Tribüne* 'from the rostrum'.

Besides using common names that mark the hypotext type, hypotext titles are also sometimes provided: *Roman Kassandra* 'novel Kassandra'; *Flüchtlingsgespräche* 'Refugee Conversations', etc.

Another type of intertextual markers is the name of the recipient of hypotext. This type of markers is used mostly in combination with other markers, especially with the marker of author of hypotext.

5. Conclusion

The study has shown that the members of the German Bundestag use numerous quotations accompanied by meta-intertext to achieve their communication goals – to present their political party and to discredit political opponents. Meta-intertext presents that the author uses intertext, and contains various markers identifying and describing it. The most important are markers of author of hypotext and of hypotext type as well as expressions which evaluate the hypotext.

The findings of the study suggest that the structure of a text fragment containing intertext and meta-intertext can be represented through the sequence of various meta-intertextual markers. This sequence shows a great variation, but it is still possible to detect typical structures that are best suited to achieve the communication goals of the members of

communication in parliament. The first part of these structures is the author of hypotext or hypotext type and the second part is the intertext. If both hypotext type and author of hypotext markers are used, the first is usually placed before the latter. Important is also the sequential structure where the first part is meta-intertextual marker of author of hypotext, the second part is intertext and the third part is the evaluation of intertext. These structures can be extended through such markers as time and place, i.e. when and where the intertext was produced, and recipient of hypotext.

The description of the typology of meta-intertextual markers and their sequential structures provides an opportunity for further study of the pragmatics of using quotations in German parliamentary discourse. Pragmatics analysis will give possibility to compare the use of meta-intertext in different parliamentary cultures. Also, the study of unmarked referential intertextuality is important to reveal the overall picture of the use of intertext in German parliamentary discourse.

References

Adamzik, Kirsten. 2004. *Textlinguistik: Eine einführende Darstellung*. Tübingen: Niemeyer.

Allen, Graham. 2000. *Intertextuality*. London / New York: Routledge.

Antos, Gerd & Tietz, Heike (eds.). 1997. *Die Zukunft der Textlinguistik: Traditionen, Transformationen*. Trends. Tübingen: Niemeyer.

Bakhtin, M.M. 1981. *The Dialogic Imagination: four essays*. Austin: University of Texas Press.

Berrocal, Martina. 2016. Quotations, Intertextual References, Models and Myths in the Presidential Debate of the Check Parliament. *Zeitschrift für Slawistik* 61(1). 119–138.

Broich, Ulrich & Pfister, Manfred (eds.). 1985. *Intertextualität: Formen, Funktionen, anglistische Fallstudien*. Tübingen: Niemeyer.

Burkhardt, Armin (ed.). 2003. *Das Parlament und seine Sprache: Studien zu Theorie und Geschichte parlamentarischer Kommunikation*. Tübingen: Niemeyer.

Constantinescu, Mihaela. 2012. The Use of Quotations in the Romanian Parliamentary Discourse. In L. Ionescu-Ruxăndoiu (ed.), *Parliamentary Discourses across Cultures: Interdisciplinary Approaches*. Cambridge: Cambridge Scholars Publishing. 263–282.

Deutscher Bundestag. *Stenografischer Bericht. 18. Legislaturperiode. Sitzung 1–14*. 2013–2014. – Berlin.

Dörner, Andreas & Vogt, Ludgera (eds.). 1995. *Sprache des Parlaments und Semiotik der Demokratie: Studien zur politischen Kommunikation in der Moderne*. Berlin; New York: Walter de Gruyter.

Fairclough, Norman & Wodak, Ruth. 1997. Critical Discourse Analysis. In T. A. van Dijk (ed.), *Discourse as Social Interaction*. London: Sage. 258–284.

Gavenko, A.S. 2011. Metaintertextuality in Literary Text (on the Material of Russian Short Stories of the 80-s of the 20th Century – 21st Century). *Philology and Human 2*. 36–48.

Genette, Gérard. 1997. *Palimpsests: Literature in the Second Degree*. Linkoln: University of Nebraska Press.

Gruber, Helmut. 2012. Intertextual References in Austrian Parliamentary Debates. A Pilot Study. In A. Fetzer, E. Weizman & E. Reber (eds.), *Follow-Ups across Discourse Domains: a Cross-Cultural Exploration of their Forms and Functions*. Würzburg: Universität Würzburg. 87–106.

Heinemann, Margot & Heinemann, Wolfgang. 2002. *Grundlagen der Textlinguistik: Interaktion – Text – Diskurs*. Tübingen: Niemeyer.

Hermann, Karin & Hübenthal, Sandra (eds.). 2007. *Intertextualität: Perspektiven auf ein interdisziplinäres Arbeitsfeld*. Aachen: Shaker.

Ilie, Cornelia. 2015. Parliamentary Discourse. In K. Tracey (ed.), *The International Encyclopedia of Language and Social Interaction*. Hoboken / New Jersey: John Wiley & Sons. 1113–1127.

Jorgensen, Marianne & Phillips, Louise J. 2002. *Discourse Analysis as Theory and Method*. London: Sage.

Klein, Josef & Fix Ulla (eds.). 1997. *Textbeziehungen: Linguistische und literaturwissenschaftliche Beiträge zur Intertextualität*. Tübingen: Stauffenburg.

Kolesarova, Lucia. 2008. *Das Sprechverhalten eines Abgeordneten in Regierung und in Opposition: Untersuchung zu situationsbedingten Unterschieden*. Wien: Universität Wien. (Doctoral dissertation.)

Kristeva, Julia. 1980. *Desire in Language: A Semiotic Approach to Literature and Art.* New York: Columbia University Press.

Mac, Agnieszka & Szwed, Iwona. 2016. Zu Intertextualitätsrelationen in den Kommentartexten der polnischer Tagespresse. *Germanica Wratislaviensia* 141. 267–283.

Panagiotidou, Maria-Eirini. 2011. A Cognitive Approach to Intertextuality: The Case of Semantic Intertextual Frames. *Newcastle Working Papers in Linguistics* 17. 173–188.

Stasiuk, Oleksandr. 2016. *Metaintertextuelle Markierungen des medialen Interdiskurses in parlamentarischen Texten Deutschlands und Schwedens. Germanica Wratislawiensia* 141. 335–348.

Witosz, Bożena. 2017. Metatext in the Discourse of the Theory of Text, Stylistics and Pragmalinguistics. *Forum Lingwistyczne* 4. 107–112.

Wu, Jiangou. 2011. Understanding Interdiscursivity: A Pragmatic Model. *Journal of Cambridge Studies* 6 (2–3). 95–115.

Oleksandr Stasiuk
Department of Germanic Philology and Translation
Taras Shevchenko National University of Kyiv
Kyiv

Ukraine
o.stasiuk@knu.ua

Galyna Stroganova
Department of Modern European Languages
Kyiv National University of Trade and Economics
Kyiv
Ukraine
gstroganova2016@gmail.com

Cross-dialectal similarity of registers: The case of the sentence across Ghanaian and British newspaper editorials*

George Kodie Frimpong, University of Ghana

The claim of register theory is that there is cross-dialectal similarity in linguistic choices across texts from the same register because language used in a register serves a specific communicative purpose in a given situational context. In this empirical study, the sentence was investigated across a specialized corpus of editorials from Ghanaian and British newspapers with the hope of ascertaining the similarities or otherwise in this aspect of language use across native and nonnative dialects of English. In the light of Biber and Conrad's (2009) model of register, we argue that quantitative dominance of specific sentence types and consistency in their distributional patterns across the two sociocultural contexts are indicative that those dominant sentence types are functional to the editorial register. The data reveal that in spite of a few distributional discrepancies, which reflect regional adjustments in situational and cultural contexts, sentence types are generally similarly distributed across editorials from the two socio-cultural contexts, with declarative and complex sentence types being consistently dominant across the two texts.

Keywords: *register theory, language and function, newspaper editorials, sentence types, hypotactic and rank shifted relations, the sociolinguistic profile of Ghana*

1. Introduction

This study is a corpus-based empirical study. It compares the use of sentence types across editorials from Ghanaian and British newspapers. The sentence, also conceptualized as a special type of clause (Greenbaum 1996; Bloor & Bloor 2004; Halliday & Matthiessen 2004; Downing & Lock 2006), has been argued in the literature as the highest grammatical structure in human language (c.f. Halliday & Matthiessen 2004: 9-10; Eggins 2004: 123; Fawcett 2000: 192). What this means, in the sense of grammatical rank scale or constituency (Fawcett 2000: 233, 237), is that the sentence is the most loaded grammatical unit (Frimpong 2015). It contains in it all the other units (clause, phrase, word, and morpheme) of grammar in a scalar order such that each higher unit embeds the unit immediately below it. Thus, a sentence is made up of at least one clause, a clause is made up of phrases, a phrase is made up of words and a word is made up of morphemes[1] (Bloor & Bloor 2004). The attempt in this paper is to explore the usage dynamics and, thus, the functional motivation (Bloor & Bloor 2004) behind the distribution of sentence types across editorials of four selected Ghanaian and British newspapers (i.e. the *Daily Graphic* and *Ghanaian Times* from Ghana and *The Daily Telegraph* and *The Times* from Britain), taking the newspaper editorial as a unique register whose linguistic choices are influenced by the situational context of the editorial (Biber & Conrad 2009).

Even though more than four newspapers from varied backgrounds could have been included in this study, we considered the four just adequate in enabling us to compare an aspect

[1] These, though, are grammatical potentials since in reality each of these grammatical units may contain at least one of the units immediately below it.

of the English used in Ghana with its historical stock,[2] the British English. The essence of this study is not so much about the variation between a native and nonnative English since there is already ample literature in that regard (Kachru 1981; Schmied 1990; Young & Walsh 2010; Huber 2012; Chien 2014). The essence of this study is about the nature of the variation or otherwise. Especially since this is a study at the confluence of dialectology and register, one is not just interested in ascertaining how language manifests variability and/or uniformity but more importantly, one seeks to establish the nature of the variability or conformity in linguistic choices under the influence of these two external factors of linguistic variability. Thus, the study is premised on the following corresponding research questions:

1. How are functional and structural sentences distributed across editorials from *Daily Graphic, Ghanaian Times*, *The Daily Telegraph*, and *The Times*?
2. What functional motivations influence the distribution of functional and structural sentence types across editorials from the selected Ghanaian and British newspapers?

2. The newspaper editorial register

Many newspapers have special sections for editorials which may be an institutional voice of the newspaper, an opinion of an individual and/or an advice of a virtual counselor (Biber 1988). This study is limited to the institutional editorial which is considered in this study as an independent subgenre of the newspaper macro-genre. This position is in line with arguments in the literature that newspaper editorials, unlike other newspaper subgenres, present opinions on factual information whose communicative purpose is essentially to influence the perception of their readership (van Djik 1989; Wiredu 2012; Frimpong 2017). Biber and Conrad (2009: 110) capture this communicative function of the editorial as follows: "[A]n editorial is meant to express an opinion overtly and persuade readers to that opinion."

Due to its explicit agenda of representing the ideological position of a particular media house and of persuading its audience in a particular way of thinking, the newspaper editorial (NE) is noted to use language strategically. Frimpong (2017), for instance, has observed that editorials from Ghanaian and British newspapers gravitate towards nominal clauses. This, he observes, is a linguistic strategy for communicative elaboration. Other studies on the NE register have identified specific linguistic features that mark them as a unique register. They are seen, for instance, as having similar generic structure (Ansary & Babaii 2009); presenting powerful ideological positions through careful use of language (Le 2009); having similar distribution of transitivity patterns (Shokouhi & Amin 2010), etc.

This study aims to consolidate work on NE from grammatical perspective. Its cross-cultural focus is to affirm the register argument that because texts from the same register are inspired by the same communicative purpose, linguistic choices are largely similar even across different dialectal orientations. This is what Biber and Conrad argue as follows:

> *Regardless 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…*

---

[2] Over a century ago, the British planted the English language on the Ghanaian soil through colonialism. Ghana has continued to use English as an official language even though it gained freedom from British colonial rule in 1957.

There are of course differences in pronunciation and word choice associated with different dialects, and there are even occasional differences in grammar. But these differences are minor when compared to the major linguistic differences among different registers, associated with different situations of use (Biber & Conrad 2009: 12 & 15).

Thus, though the editorials studied in this paper are from different sociocultural contexts with variable dialectal orientations, they should share similar linguistic features since they share similar features in their situational context.

The focus on editorials from Ghanaian and British newspapers is motivated by the desire to compare and describe aspects of the two varieties of English (the Ghanaian and British varieties of English, which coincidentally are historically related) from register perspective. Not only is the English used in Ghana a development of what was once introduced into that geographical context by the British colonial empire but also, the editorial tradition and the newspaper industry in general were introduced into the Gold Coast, now Ghana, by the British colonial empire (Asante 1996; Frimpong 2015). One would expect that this historical connection should influence both language use and newspaper styles of production. However, this is not the case. For instance, as argued below, Ghanaian newspapers have different classification, style of presentation, content coverage, etc. Besides, in terms of the language, the argument in the literature is that when a language migrates to a new location, it reflects the fauna and flora of its new location thereby undergoing a process of nativization (Schneider 2007). Thus, in a sense, this study forms part of the ongoing debate in the discipline that though Ghanaian English emerged from the British stock and that speakers of Ghanaian English are conscious of maintaining the British standard, there is ample evidence of variation in all aspects of the English used in Ghana (Sey 1973; Dako 2003; Huber 2012; 2014). This study extends the argument further to include register variation. Thus, given the argument above that Ghanaian English has significantly distanced itself linguistically from its historical stock, linguistic similarities in aspects of this variety may be safely explained from register perspective.

3. Audiences, ideologies and traditions of newspaper production and circulation across Ghana and Great Britain

Great Britain boasts of one of the oldest surviving newspaper traditions in the world. The English press started in a handwritten pamphlet format in 1622 (Morison 1932). The two British newspapers, *The Times* and *The Daily Telegraph*, were founded in 1785 and 1855, respectively (Morison 1932; Frimpong 2015).

British newspapers over the years have developed along two trajectories which aid their classification. On the one hand, newspapers in Britain have been classified along ideological lines, especially in terms of a newspaper's political inclination. From this perspective, *The Times* and *The Daily Telegraph* are both pro conservative newspapers. This, of course, is a bias that could have been corrected by replacing one of them with a newspaper of an alternative political persuasion. This bias, however, is deemed not so significant to affect the use of grammatical structures. The focus of this paper is on quality newspapers and any of them qualifies to be selected. The second, and perhaps the most popular developmental trajectory, is along the lines of readership and the quality of language use.

Based on a targeted class of readership and quality of language used, British newspapers are again classified into quality newspapers, midmarkets and the tabloids (Chan & Goldthorpe 2007). Newspapers from these categories target upper middle, lower middle and working classes, respectively. In terms of language use, whereas the quality newspapers, including *The Times* and *The Daily Telegraph*, are noted to use formal and serious mode of writing, devoid of sensationalism, the midmarkets and tabloids, however, use less formal and more sensational language (Quinn 2013).

The Ghanaian modern newspaper industry is a relatively nascent one. Until 1950 when the oldest surviving newspaper, the *Daily Graphic*, was established, several attempts had been made unsuccessfully, first by the British colonial administration and then by some local Gold Coasters (Asante 1996). Prior to the advent of the Western-styled print media, the West African sub-region, including Ghana, used various indigenous forms of mass communication such as runners, talking drums, gong gong, fire, etc. to disseminate information to unenumerated audiences (Fosu 2014). It is noteworthy that the Gold Coasters who made the efforts at establishing and maintaining the newspaper industry were motivated primarily by political exigencies. It is not surprising, therefore, that the Ghanaian newspaper industry has developed along different paths. Thus, Ghanaian newspapers are classified into public or private newspapers according to which public newspapers are state funded and private newspapers are funded by private individuals. As state funded newspapers, public newspapers have wider circulation and topical coverage due to their ability to fund both quality production and national representations. The *Daily Graphic* and the *Ghanaian Times*, the two main public newspapers in Ghana, are by far the most voluminous newspapers covering more varied topical contents than their private counterparts.

One key feature that differentiates British newspapers from Ghanaian newspapers is their readership audience. Though newspapers from the two sociocultural contexts are produced in English for English speaking readership, the sociolinguistic backgrounds of readers from the two contexts differ significantly. This is because, whereas readers of British newspapers are native speakers of English, readers of Ghanaian newspapers are non-native second language speakers. Similarly, Fosu (2014) notes that whereas British readers are predictable in their choice of newspaper type, Ghanaian readers are only predictable in their choice of a private newspaper. That is, sympathizers of the major political parties gravitate towards newspapers that promote their party. For public newspapers, however, readers are attracted to them based on the quality of language use. As Fosu argues, readers of public newspapers have variegated ideological backgrounds who are brought together by the common denominator of Standard English. This argument makes sense because in Ghana, Standard English is a mark of identity that brings together educated Ghanaians of all ideological and ethnic backgrounds. Because it is acquired in the classroom, only educated elites, regardless of political or ideological background use this language. Thus, Ghanaian public newspapers target this small constituency of the Ghanaian population as their core reading audience.

4. The sociolinguistic profile of Ghana and the language of Ghanaian newspapers

In a study of this nature, an account on the sociolinguistic situation of Ghana is considered necessary. This is because, language use in Ghana is not as straightforward as it is in a native context. In mother-tongue English communities like the UK, the dominance of English has led to the weakening, even death of minority languages (Simo-Bobda 1997:88). The situation is

different in Ghana. In Ghana, though the English language is the official monolingual language, just as in other West African countries, it co-exists with vibrant indigenous languages making the sociolinguistic situation of Ghana a very complex one (Dakubu 1988; Obeng 1997). As an officially monolingual country, (Dakubu 1988), Ghana uses only one language (English, an exolanguage inherited from British colonial administration (Guerini 2008)) in official domains (Simo-Bobda 1997:87). English is thus used as the language of education, in legal documents, in newspapers and magazines, in parliament and in all official transactions in Ghana (Huber 1999). As the language of education, it is not surprising that most published materials in Ghana are done in English, an observation made by Dakubu decades ago which she foresaw would not change in any foreseeable future (c.f. Dakubu 1988: 10). Thus, though monolingual in official domains, Ghana is highly multilingual (Dakubu 1988) in informal domains, home to over 80 indigenous languages (Ansah 2008) whose speakers live together and share social experiences as friends, neighbours, members of social groups and even as spouses, especially in urban contexts.

The prestige given to English in Ghana by being selected above the 80 plus indigenous languages in all official domains and the social status and power English offers to its users in Ghana have created a kind of a diglossic situation with English as the high language and the local languages as the low and less prestigious languages (Guerini 2008). This prestige influences some parents in Ghana to introduce their children to English very early, even before school age. They rationalise that it is English that has economic power and can gain access for their children into prestigious schools, especially in urban and peri-urban centres.

In spite of the prestige enjoyed by English in Ghana, Akan, one of the indigenous languages has managed to assume a lingua franca status, together with English and Hausa (Obeng 1997), spoken by more than 80% (40% native speakers and about 43% L2 speakers) of Ghana's population (Guerini 2008). Ansah (2008) notes a number of practical and sociocultural reasons underlying the popularity and utility of Akan in Ghana. As she observes, much of private economic activities are engaged in Akan. So too are radio discussions and television programmes conducted mostly in Akan, to the extent that many radio and television stations across the length and breadth of Ghana are Akan-based (Ansah 2008). The prestige and popularity of Akan extends to religious and cultural ceremonies (such as naming, funeral, marriage, Christian worship, etc.) and political discussions.

It is noteworthy that many speakers of Akan and of the other indigenous languages (both native and nonnative speakers) are not literate in these languages. This perhaps explains the popularity of Akan in informal domains only, especially in speech-based domains. This also explains why though just about 40% of Ghanaians can speak and read English, newspaper producers find it profitable to target this minority. It makes sense therefore that newspaper production in Ghana is exclusively done in English. Otherwise, it does not make economic sense to produce for the minority of Ghanaians (<40) who are able to read in English.

5. Language and function

Function has been conceptualized in linguistics from several perspectives (Bloor & Bloor 2004; Payne 2011; Biber & Conrad 2009). It has been a common approach in traditional grammar to analyse grammatical structures from their formal and functional perspectives, based on the grammatical role a grammatical form is playing. Thus, a noun phrase may be a subject, object, complement, etc. (Quirk et al. 1985). Function in language has also been approached from a

semantic perspective where a structure is assigned a participant role of an agent, goal, senser, locative, phenomenon, etc. (Payne 2011). From pragmatic perspective, however, function in language is associated with the speech acts underlying language use. This is what Bloor and Bloor (2004) consider as communicative function. From a more socio-functional perspective, Halliday and Matthiessen (2014) conceptualize function in language as multifaceted in their metafunctional theoretical model. By far, one of the most theorized relationships between language and social function, the metafunctional model argues that language is functional simultaneously in three layers. The first layer of language function is the ideational layer, while interpersonal and textual metafunctions constitute the last two layers (Halliday 2004; Eggins 2004). That is, language reflects realities of its situational context of use at the ideational level, reveals contact and social relations among participants in a communicative event at the interpersonal level and enables information to be organized to reflect a communicative purpose and context of use at the textual level (Eggins 2004).

Though this study is not directly based within the systemic functional theoretical model, already one can imagine how newspaper editorial language can reveal aspects of ideational, interpersonal and textual meanings. For instance, the language of editorials reveals information that is packaged to persuade an unenumerated audience based on its argumentative style. Besides, it may be equally revelatory that the language is written rather than spoken by an unidentified author. These speculations are of course idealistic, taking the situational contexts within which newspaper editorials are written as similar across the two sociocultural contexts. On the contrary, however, the two contexts reflect different sociocultural orientations, which potentially induce linguistic variation in large measures. The orientations of the native/ nonnative context of language use are such that whereas in the former context English is used for both formal and informal activities, in the latter English is used mainly in formal domains. These variations in the use of the English language across the two contexts may reflect in levels of formality across editorials from the two contexts. As Schneider (2007) argues, the sociocultural factor is one crucial parameter that causes a language to be indigenized/nativised to reflect its sociocultural context. Simply put, a language is being used to express cultural and social realities of its new home, especially where the language has migrated as in the case of English in Ghana. Thus, though it is the same language planted on the new land by the native speaker, its novel sociocultural orientations enter the language to engender a new variety of the old language. And this accounts for the dialectal variation postulated in the literature of new Englishes.

In this study, function is taken from the perspective of Biber and Conrad (2009), according to which a linguistic feature is deemed functional based on its relationship with the situational context of a particular register. Thus, a linguistic feature is functional if its quantitative representation is more prominent in a particular register than in another. From this functional perspective, users of language are seen as making linguistic choices based on a situational context which is made up of a number of situational features, including the participants of a discourse, the relationship between participants, the topic involved, the communicative purpose of a discourse, the setting of the discourse and the channel or medium of the discourse. For newspaper editorial, one estimates that the participants involved (an authoritative representative of a media house addressing an unidentified mass audience), the medium (written verbal communication), the setting of production and comprehension circumstances (availability of time for planning and editing the editorial before publication and the editor's awareness that the audience have time to process the editorial for comprehension) and the communicative purpose (to persuade its readership to a certain ideological standpoint)

have a high predisposition of influencing linguistic choices. Thus, these variables together with the variable sociocultural backgrounds of the newspapers explored in this study constitute the primary motivation of linguistic variability and/or similarity across the newspapers investigated

6. Sentence types

The sentence has been classified in the literature based on its structural and functional properties. Structurally, a sentence may be either simple sentence (having only one main clause) as in sentence (1) below or non-simple/multiple sentence (having two or more clauses) as in sentence (2) to (4). Among the non-simple sentences, a differentiation is normally made among sentences which have a coordination of two or more independent clauses (compound sentences) as in sentence (2), sentences which have a subordination of one or more dependent clauses to a main clause (complex sentences) as in sentence (3) and sentences which have a coordination of two or more independent clauses and a subordination of one or more dependent clauses (Compound complex sentences) as in sentence (4) (Quirk et al. 1985; Wardhaugh 1995; Biber et al. 1999).

(1)   *Simple sentence*: President Sarkozy has not been arrested. (*The Daily Telegraph*)

(2)   *Compound sentence*: Ghana has scored high marks in its democratic march and the country is seen as the beacon of hope for Africa. (*Daily Graphic*)

(3)   *Complex sentence*: Of course, while the *Times* admits that a greater responsibility rests with Dr. Kwadwo Afari-Gyan and his team on the EC, it also holds the view that there is little the commission can do, without the cooperation of all Ghanaians. (*Ghanaian Times*)

(4)   *Compound Complex sentence*: The Canute-like frustration of politicians at their inability to control the tax affairs of the multinationals is out of proportion to the benefit they bring to the economy - or to the sums at stake which set against the country's vast indebtedness can only be symbolic. (*The Daily Telegraph*)

Functionally, simple sentences are further classified in the literature into declarative, interrogative, imperative and exclamative sentences (c.f. Quirk et al. 1985; Biber et al. 1999). But this classification does not mean that non-simple sentences cannot be classified functionally; for, elsewhere Wiredu (2014) in a study of legal reports, classifies complex sentences, some of which have up to 20 clauses, into functional (predominantly declarative) types and his classification finds support in the literature. Huddleston's (1999) argument about desententialized clauses offers grounds for classifying non-simple sentences into speech acts/ functional types.

Huddleston (1999) has argued that when a clause (finite or nonfinite) is subordinated or coordinated to a main clause the subordinated or coordinated clause is desententialized. This means that the subordinated or coordinated clause has its illocutionary force depleted and therefore its speech act status is tied to the independent clause in the sentence. Based on this argumentation, functional sentences (simple or non-simple types) are classified in this work,

in conformity with Wiredu's (2014) classification, based on the speech act function of the main clause. Thus, a complex sentence, for instance, is declarative if its main clause is declarative, but interrogative if its main clause is interrogative – a principle that is used to classify the following sentences selected across the corpus.

(5)   *Compound Declarative*: The quality of our universities is one of Britain's great strengths and we already have 12 in the top 100, according to the Times Higher Education ranking. (*The Times*)

(6)   *Simple Interrogative*: Are the political parties going about their campaigns in a sensible and mature way? (*Ghanaian Times*)

(7)   *Complex Imperative*: To see where things are going wrong, consider the attempt to introduce elected mayors. (*The Daily Telegraph*)

(8)   *Complex Exclamative*: How cruel death is, to take away the lives of two MPs in a month in an election year! (*Ghanaian Times*)

The four types of functional sentences constitute what (Sadock & Zwicky 1985: 159) refer to as a "mutually exclusive" system according to which membership of one type excludes a sentence from belonging to another type.

Crystal and Davy (1973) note about these four types of sentence that in journalese writing, including editorials, it is the declarative type which is the most preferred. Similarly, Biber et al. (1999) have observed that complex sentences are preferred in the press. But given that the editorial constitutes an independent subgenre within the newspaper, as argued above, with unique situational features, it should be beneficial to our understanding of the linguistic versatility and diversity of the newspaper genre to uncover the usage dynamics of these sentence types in the subgenres of the newspaper such as editorials. In line with register theory (Biber & Conrad 2009), preference for a particular sentence-type is interpreted in this paper as functional. Additionally, this study places emphasis on consistency in the distribution of grammatical features across the two sociocultural contexts of the study. The argument is that if a prominent pattern is functional, it should be consistent in its prevalence across texts from both Ghanaian and British newspapers.

7. Methodology

7.1 *The corpus*

This study is an empirical corpus-based study (Mcenery & Gabrielatos 2006). For the purpose of this study, a specialized corpus of editorials from selected Ghanaian and British newspapers was built. "…smaller, more specialized corpora…focusing on specific registers and genres" (Koester 2010: 66) are becoming increasingly necessary in special purpose research. A specialized corpus was deemed the most appropriate for this study because, as argued by Flowerdew (2004: 21) it is the most suitable method for investigating specific linguistic features and for exploring texts involving particular participants, mode and communicative purpose. For this study, a total of 144 editorial texts from 2 Ghanaian and 2 British newspapers

(i.e. *Daily Graphic* and *Ghanaian Times* from Ghana and *The Daily Telegraph* and *The Times* from Great Britain) produced during 2012 were compiled. For ease of reference, these newspapers are henceforth referred to as DG, GT, DT and TT respectively for the *Daily Graphic*, the *Ghanaian Times*, *The Daily Telegraph* and *The Times*.

The original motivation for conducting this research was to examine an aspect of the English used in Ghana from register perspective and the newspaper editorial came in handy in that in Ghana, it is among the most easily accessible English texts produced by Ghanaian speakers of English. Register studies, according to Biber and Conrad (2009), is better done comparatively and newspaper editorials of any national variety of English could have been compared with the Ghanaian editorials. However, the attempt to compare an aspect of Ghanaian English with a corresponding feature of British English forms part of an ongoing debate in the literature. In his pioneering work on English in Ghana, (Sey 1973) has made the observation that Ghanaian speakers of English are conscious of maintaining British English as an exonormative standard. Over the years, however, scholars have identified tendencies in the English used in Ghana that make it possible for one to argue in favour of a Ghanaian English (c.f. Dako & Huber 2004; Huber 2012). Huber (2012), for instance, argues that the English used in Ghana is unique not in the creation of new words but more importantly in restructuring features of Standard British English. In comparing an aspect of Ghanaian English with British English, one hopes to contribute to the debate on how similar or dissimilar the Ghanaian variety of English is from the British native variety.

The selection of newspaper types was based on convenient sampling. Thus, though there are other equally influential quality newspapers, *The Times* and *The Daily Telegraph* were selected based on convenience. Though one could use the popularity of these two newspapers (c.f. Westin & Geisler 2002) as the motivation for selecting them, in this study they were chosen because they were accessible at the time of the study. On the other hand, the two Ghanaian newspapers were chosen because they are the two main public daily newspapers. The primary focus in selecting these newspapers was mainly comparability. Thus, one was keen at selecting Ghanaian newspapers that were somehow comparable with the British quality newspapers. Generally, the popularity and influence of the two Ghanaian newspapers, based on their wider circulation and assumed objective use of language, influenced their choice.

Work on newspaper circulation in Ghana cites the *Daily Graphic* and *Ghanaian Times* as the most widely circulated (Yankson et al. 2010; Fosu 2014; Frimpong 2017). Fosu (2016: 6) argues that the two Ghanaian newspapers are among the "dominant, credible and influential quality publications in Ghana."

The notion of comparability in this paper is not to denigrate the contextual variability at the background of the newspapers explored. We are aware that people from different dialectal backgrounds use language differently. However, as the theoretical argumentation has it, language used in similar situational contexts reflects similar linguistic choices even across different dialects. It is in the hope of not comparing say tabloids with quality newspapers that we take these precautionary measures which limit our choice of newspapers from the Ghanaian context to the two public newspapers.

Due to the large number and the varied topical scope of issues produced by the four newspapers during the period of the study and in the bid to control the variables compared, the selection of texts into the corpus was further limited to texts which have political content. These texts were not difficult to select from the British newspapers since their online archives have categorized texts into various thematic areas, including politics. The problem was with the Ghanaian newspapers which were uncategorized. One could observe, however, that most of

the Ghanaian texts had political content in the sense that they almost always argued or commented on political debates, election processes and results, political campaigns and rallies, etc. Thus, in order to obtain adequate texts into the Ghanaian corpus, we focused on political editorials. Political editorials were operationalized in this study to mean texts which are about or comments on national elections and campaigns, national registration of voters, parliamentary and presidential debates, activities of government departments and agencies, speeches and activities by members of the four wings of government[3], etc. These parameters equally influenced our selection of the types of political editorial texts that were compiled into the British corpus. The headlines of editorials for both British and Ghanaian newspapers selected for this study are listed in Appendix I.

Texts selected from the four newspapers were the first three texts of each of the 12 months of 2012 whose contents suit the definition of political editorial above. A summary of the size of the corpus and the number of sentences and words are presented in Table 1 below.

Table 1: Background information about the corpus

	DG	GT	DT	TT	TOTAL
Number of Texts	36	36	36	36	**144**
Number of Sentences	700	534	774	1118	**3126**
Number of words	17666	14004	18027	21925	**59022**

Table 1 shows that though the same number of editorial texts was chosen from the four newspapers, the number of sentences attested varies from one newspaper to the other. This is not just because of stylistic variation but also because of the variable length of editorials from the four newspapers. One observes that the British newspapers typically have longer texts than the Ghanaian newspapers. Besides, there is internal variation in the length of editorials even from the same sociocultural context. For instance, *The Times* is about 3000 words longer than *The Daily Telegraph* and so is the *Daily Graphic* about 3000 words longer than the *Ghanaian Times*. To compensate for this, a normalization formula (c.f. Biber & Conrad 2009) was used to standardize the results (the detail of which is presented in section 7.2.2 below).

7.2 *Analytical framework*

The analytical model adopted for this study is that of Biber and Conrad's (2009) register model. According to their framework, a register-based study is a three-layered engagement, which requires a mixed methodological approach of quantitative and qualitative analysis. These include a description of the situational context, a description of the linguistic categories explored and an analysis of the functional relationship between the distributional patterns and the situational context. These tasks are engaged at the various sections of this study. For the purpose of analysis, the corpus of the study was tagged and the attestations normalized to facilitate comparability.

---

[3] This operational definition is based on *Oxford English Dictionary's* (OED) definition of *political* as "Of, belonging to, or concerned with the form, organization, and administration of a state, and with the regulation of its relations with other states."

7.2.1 *Tagging of corpus*
Since the editorial corpus was compiled purposely for this study, care was taken to annotate it with the grammatical categories studied in this work – the sentence. This task of annotation, which is very time-consuming (Mcenery & Gabrielatos 2006) because it is done manually, is very useful in facilitating the retrieval and counting of attestation. Afterwards, attested sentence types were retrieved using Microsoft Word search engine. In order to ensure that the attestations were not influenced by subjective biases, a national service personnel was trained as a checker who verified the tagging and the attested sentence patterns.

7.2.2 *Normalizing attestations*
One principle that is crucial in a comparative register study is standardization of attested results either through percentage terms or normalization (Biber 1988; Biber & Conrad 2009). Standardization is essential because, usually, texts compared come in variable lengths, making the raw scores methodologically problematic. For instance, it is misleading to use the raw scores reported in Table 2 below. This is because, since the texts are of unequal lengths, it will be wrong to conclude based on these results that *The Times,* for example, used the highest number of declaratives. This is why it is necessary to standardize the scores to facilitate comparison. For the purpose of this study, normalization was used to standardize attested results.

Table 2: Spread of functional and structural Sentence types in raw scores

Functional	DG	GT	DT	TT	Structural	DG	GT	DT	TT
Declarative	690	520	751	1086	**Simple**	62	80	150	260
Interrogative	6	5	19	31	**Complex**	585	409	548	784
Imperative	4	8	2	1	**Compound**	18	15	26	34
Exclamative	0	1	2	0	**Comp compl**	35	30	50	40
Total	700	534	774	1118	Total	700	534	774	1118

Using Biber and Conrad's (2009) framework, normed rates were computed out of the raw scores in Table 2 above with a fixed amount of text. For the purpose of this study, these raw counts were converted to rates of 100 sentences per newspaper. This is the rate at which a sentence type occurs in every 100 sentences.

The formula for the conversion, according to Biber and Conrad (2009: 62), is:

Normed rate **=** (raw count ÷ total sentence count) **x** the fixed amount of text

Now using our fixed amount of 100 sentences per newspaper, the normed rate for the occurrence of declaratives in DG will be:

(690 (sentences) ÷ 700 (total)) x 100 (sentences) = 98.57.

This normalization formula above is applied to all the functional and structural sentence types and the normed rates of occurrence are reported in Tables 3 and 6 under sections 8.1 and 8.2 below. For purposes of comparability, a Chi-square test of independence is employed at alpha = 0.05.

8. Results and Discussions

8.1 *Functional Sentence types*

The normed rates of occurrence for functional sentence types reported in Table 3 below present a very fair basis upon which comparison was based across the four newspapers investigated. To facilitate comparability, the Chi-square test was run on individual sentence types. This is because, the data variates the assumptions in the Chi-square distribution of having about 66.7% of the cells less than expected counts of 5.

Table 3: Normed rates of occurrence for functional sentence types

Sentence type	DG	GT	DT	TT	χ	*P-value*
Declarative	98.57	97.38	97.03	97.14	0.0308	0.9986
Interrogative	0.86	0.94	2.45	2.77	1.7416	0.6321
Imperative	0.57	1.50	0.26	0.09	1.9847	0.5756
Exclamative	0	0.18	0.26	0	0.4528	0.9291
Total	**100**	**100**	**100**	**100**		

These results in Table 3 indicate that the functional sentence types are largely similarly distributed across the four newspapers – the declarative sentence being the dominant type (attesting >97), followed by the interrogative (hovering between 1 and 3) and the imperative (<2) consistently across the four newspapers. The least attested type (the exclamative sentence) attested almost nothing across the four newspapers. These patterns of distribution enable one to argue that choices of these sentence types are motivated by the situational context. That is, something about the newspaper editorial genre must have influenced the similarity in these linguistic choices. However, before we get to that argument, there is an observation that needs to be made. Though the Chi-square test run on all the four sentence types confirms that the distribution of the sentences does not depend on newspaper type since observable differences are statistically insignificant, one detects a pattern in the distribution of interrogative sentences which seems to distinguish Ghanaian editorials from British editorials. In the light of the register theory, results such as those for the interrogatives are deemed insignificant due to their low attestations. However, the regularity in their occurrence serves as a prod to investigate the interrogatives a little closely for dialectal reasons.

A detailed analysis of the interrogatives reveals interesting observations. It is observable, from functional perspectives that the few attested interrogatives presented in Table 4 below are mainly the rhetorical type, especially for the Ghanaian editorials. This pattern of distribution of rhetorical interrogatives in the Ghanaian editorials seems to reflect the sociocultural context. It indicates a conscious effort of an editor to conform to norms of Standard English. That is, since Standard English is used in Ghanaian newspapers as an ideological plank to attract educated elite, as argued above, efforts are being made at carefully selecting grammatical patterns in conformity with Standard English. This variation is obvious when comparison is made with attestations in the British editorials.

Observably, the British editorials used non-rhetorical types in combination with rhetorical interrogatives. This too may be explainable based on the sociocultural context. This use of interrogatives, perhaps, reflects interactivity, a phenomenon that may be attributed to

the relatively more active presence of British newspapers online. During 2012 when the data were compiled, the two Ghanaian newspapers were not actively present online. *The Times* and *The Daily Telegraph*, however, were actively online and had created an interface for their readers to interact with the editors. And so the active presence of these newspapers online might have raised their consciousness of the need to draw closer to their audience through linguistic devices including interrogatives.

Table 4: Spread of functional interrogative types across the four newspapers

Functional	DG	%	GT	%	DT	%	TT	%
Rhetorical Ques	6	100	5	100	12	75	24	75
Non-Rhetorical Ques	0	0	0	0	4	25	8	25
Total	**6**	**100**	**5**	**100**	**16**	**100**	**32**	**100**

A careful study of the corpus reveals that the dominant rhetorical questions seem to perform two communicative functions in the editorial data across the Ghanaian and British newspapers. They, first of all, raise polemical issues, which often serve as a plank to launch onto the next level of argumentation; a function that most of the attested interrogatives perform in the editorials. The set of interrogative sentences in the following excerpt from the data performs this function:

(9)   What is the ultimate impact of the crisis in the eurozone? How are we going to be able to afford all the necessities of an ageing population? And, looming above all these, how is Britain to equip its people with the resources to compete with economies that are already growing rapidly? These are the genuine leadership questions that Mr Cameron will face in Birmingham. (*The Times*)

But this argumentative function performed by rhetorical questions in the data is quite different from the second function observed by Crystal and Davy (1973) according to which interrogatives are used to keep articles from dragging in newspaper reporting. The following example of interrogative sentence from the data is of the second type:

(10)   Had he, as the minister deciding on the purchase by News Corporation of the outstanding shares in BSkyB, acted inappropriately and favoured the company for political reasons? (*The Times*)

In this rhetorical question, just like in the first above, the writer does not expect a direct answer from his/her readers; neither does he/she provide an answer himself/herself. The difference, however, is that the type of rhetorical question in sentence (10) is not used as a plank to launch onto an argument. Similar rhetorical questions found in the Ghanaian editorials include the following:

(11)   What do we seek to gain by preventing eligible persons from registering? (*Daily Graphic*)

(12)   Are the political parties going about their campaigns in a sensible and mature way? The answer lies in the incessant appeals to them to avoid acts which negate the gains so far made in our democratic pursuits. (*Ghanaian Times*)

Equally noteworthy about the distribution of the interrogative sentences is the observation, captured in Table 5 below, that the most preferred structural patterns are the ones which do not just require *YES* or *NO* for an answer. They are the ones which have the potential of engaging their reading audience.

Table 5: Spread of structural interrogative types across the four newspapers

Structural	DG	%	GT	%	DT	%	TT	%
Yes/No Ques	2	33.3	5	100	6	38	15	47
WH-Ques	4	66.7	0	0	10	62	17	53
Total	**6**	**100**	**5**	**100**	**16**	**100**	**32**	**100**

Though the usage patterns are not consistent across the four newspapers as seen in Table 5 (with GT as the odd one, conforming, perhaps, to a house style), it is clear that there is preference for *WH*-Questions. The following are examples of WH-*Questions* attested in the corpus.

(13)   What…would the protesters in Tahrir Square make of a nation in which more than 70 per cent of the upper chamber are political appointees, in which the legislature is "stuffed full of friends and colleagues of party leaders"? (*The Daily Telegraph*)

(14)   What happens to applicants who are unable to register at a centre whose equipment break down for three days? (*Daily Graphic*).

Other equally important observations about the distribution of interrogative sentences, especially unique to the British newspapers, which have the highest representation of this type of sentence, include the following:

- They do not seem to be evenly distributed. They tend to be used in successive order of two or three in particular texts whereas others do not have any interrogatives at all. In the example below, seven questions (Q1, Q2, Q3, Q4, Q5, Q6 and Q7) are asked one after the other in one text; and this is not an isolated pattern, especially among the British editorials.

(15)   **(Q1)** How about, for instance, appointing Boris Johnson as Party Chairman? **(Q2)** Does Mr Miliband want to follow a Blairite strategy of trying to keep Britain at the heart of decision-making, as implied by his party's relentless urging of David Cameron to 'stay in the room'? **(Q3)** How does that fit with directing his party to vote for a position on the EU budget that would isolate the UK and diminish our influence? **(Q4)** And how does it affect Labour's policy towards the Liberal Democrats in the run-up to the next election? **(Q5)** Would Labour fight to repatriate powers from Brussels and if so, which? **(Q6)** What approach does he think Britain should take to a eurozone that is changing fast and in which a second division will soon be created: does he think the UK should be part of that second division? **(Q7)** If not, what role should it play? (*The Times*)

- Interrogatives tend to appear as the last sentence of their texts.
- Some interrogatives are premised upon preceding declarative sentences. As a result, they do not really make complete sense; neither do they appear to be complete sentences in isolation. For instance, in example (16) below, 'But why?' lacks both subject and predicate and is meaningless unless it is analyzed within its broader co-text. Similarly, 'Why?' in sentence (18) makes sense only when it is put after the preceding declarative.

(16)   But why? (*The Daily Telegraph*)

(17)   Yet it was overwhelmingly rejected. (*The Daily Telegraph*)

(18)   Why? (*The Daily Telegraph*)

It is not surprising that all the examples of interrogatives that do not constitute complete sentences are from the British editorials. In the Ghanaian editorials, the grammatical rule of keeping sentences complete in order to avoid fragments seems to be among the guiding principles. This therefore seems to reflect the second language sociocultural context where English is used for formal purposes and the Standard English is an ideological link between the editor and his/her audience.

Considering the patterns of distribution of the functional sentence types, the question that is outstanding at this point is what accounts for the regularities in the distribution of these grammatical structures in texts from different sociocultural contexts? In other words, what makes the declaratives similarly preferred across the four newspapers and the interrogatives, imperatives and exclamatives similarly dispreferred?

These patterns of distribution are indicative of functionality in the framework of Biber and Conrad's (2009) register model. They support the central claim of the register theory that linguistic features are similarly distributed across dialects of the same language. This pattern of distribution may be interpreted to be functional by arguing that declaratives relate with the situational context of the newspaper editorial register. In other words, the declarative sentence is dominant across the four newspapers because it is useful for the realization of the communicative function of editorials in propagating ideological positions. In relying predominantly on the declaratives, not only is the Ghanaian editor conforming to norms of Standard English. In doing so, the Ghanaian editor, just like the British editor, is making claims and statements about beliefs and opinions.

Typically, it is the declarative sentence which is capable of performing a whole range of speech act functions, including expressing claims, hypotheses, descriptions, suggestions, promises, pledges, threats, assessment, facts, etc. (Finegan 2012). These are functions, one wants to argue, that are appropriate for the formulation of ideological positions in newspaper editorials. That is, the communicative functions of editorials to represent the ideological positions of their press houses are expressible through a variety of speech acts, which are realizable through several linguistic means, including the declarative sentence. The low representation of the non-declarative types (especially imperative and exclamative sentences), perhaps, suggests that they are not very essential linguistic strategies for formulating ideological positions.

8.2 *Structural Sentence Types*

The structural sentence types – simple, complex, compound and compound complex sentences – are distributed in the editorials across the four newspapers in the patterns reported in Table 6 below; these are normed rates calculated using the formula under section 7.2.2 above. Here too, the Chi-square test was run on each of the sentence types.

Table 6: Normed rates of occurrence for structural sentence types

Sentence type	DG	GT	DT	TT	χ	*P-value*
Simple	8.86	14.98	19.38	23.25	6.8705	0.0765
Complex	83.57	76.59	70.80	70.13	1.5403	0.6730
Compound	2.57	2.81	3.36	3.04	0.1138	0.9901
Compound complex	5	5.62	6.46	3.58	0.8571	0.8358
	100	**100**	**100**	**100**		

Evidently, the results in Table 6 also reveal some level of consistency in distribution. Obviously, the complex sentence is consistently the dominant type across the four newspapers followed by the simple sentence, though after a wide margin. The compound complex and compound sentences are the least used structural types across the four newspapers. In spite of these patterns of consistency, one observes in the details that some variability exists in terms of how much of a particular structural sentence type was realized in a particular newspaper. For instance, though complex sentences in the two British newspapers are similarly distributed, the difference between attestations of the complex sentence for the two Ghanaian newspapers is too wide for one to argue for regional sway. The same dissimilarity exists in the distribution of the simple sentence across the four newspapers. However, despite these dissimilarities, the confirmatory Chi-square tests identify no statistically significant variation in the distribution of these sentence types across the four newspapers.

In a sense, these distributional patterns support Hopper and Traugott's (2003) observation:

> *…to our knowledge human languages have had complex sentence structure available throughout recorded history. But reorganization of complex combinations is well evidenced…as is the association of certain complex sentence types with certain genres, especially of planned discourse" (Hopper & Traugott 2003: 177).*

This is one reason for the argument in this paper that the distribution of these sentence types is basically register-based, especially due to the level of consistency observed. In explanation, we argue that the communicative function and other situational features of newspaper editorials attract complex sentences among other linguistic devices as relevant structural tools

In spite of these patterns of similarities, one observes strands of structural diversity across texts from the two sociocultural contexts. One observes, for instance, variability in the organization of the structural sentences across newspapers. It is observed that the Ghanaian editorials are typically texts whose paragraphs are predominantly made up of a single complex sentence. This is sharply contrasted by the British editorials whose paragraphs are made up of a number of complex sentences interspersed by one or two simple, compound or compound complex sentences. This may be argued to be indicative of linguistic versatility on the part of

the native speaker at proficiently managing complex linguistic data, an assumption that is relevant to the sociocultural context. However, an observable variability in the situational context may be argued to be accountable for this variation in language use. One observes that though newspapers from the two contexts of this study may equally operate with some restrictions in the space allocated for their institutional editorials, the variation in the number of words (see Table 1 above) and space dimensions of editorials across Ghanaian and British newspapers are indicative of some of the forms of linguistic variation we have observed in this study.

But structural variability is also evident in the way the immediate constituents of the dominant complex sentences are used. A close investigation of the complex sentences attested in the data reveals the following three patterns:

- **Pattern 1**: Complex sentences in which a subordinate clause is an element of the sentence, functioning as a subject, object, complement or an adverbial. For instance, the underlined clauses in sentences (19) and (20) are adverbial and nominal clauses functioning as clause elements:

(19) When the Budget speech is still leading the news three weeks after delivery, something has gone awry. (*The Times*)

(20) But to surrender at this stage would send out the appalling message, which is precisely Labour's aim. (*The Daily Telegraph*)

- **Pattern 2**: Complex sentences in which a subordinate clause is a post-modifier of a phrase, functioning as qualifier or phrase complement. The underlined clauses in excerpts (21) and (22) are a *That*-Relative clause functioning as a post-modifier (in excerpt (21)) and a *That*- nominal complement whose head is the noun '*way*' (in excerpt (22)).

(21) These can be called bye-laws or codes of ethics that serves as a guide for acceptable behavior or conduct. (*Daily Graphic*)

(22) Of course, there should be some amount of humour to make proceedings in the House a bit interesting, but this should be done in such a way that discipline would not be undermined. (*Ghanaian Times*)

- **Pattern 3**: Complex sentences which blend the first two. In sentence (23) below, we have, first, an adverbial clause '*Although this arrangement is good*' which depends on the main clause which, itself, a matrix clause containing the complement clause '*that…the entire exercise would end in confusion*' which complements (and therefore depends on) the adjectival head '*obvious*':

(23) *Although the arrangement is good*, it is obvious *that* without proper education, *the entire exercise would end in confusion*.

Of these, Pattern 1 Complex sentences ((19) and (20) above) instantiate a hypotactic relationship. That is, clauses of unequal status (i.e. a main clause and a subordinate clause

which depends on the main clause, rather than on a phrase) have been conjoined in one sentence (Quirk et al. 1985). Pattern 2 Complex sentences ((21) and (22) above) instantiate embedding or rank shifting.

Our data reveal that there are more complex sentences with rank shifted/embedded subordination (Pattern 2 Complex sentences) than those with hypotactic ones across the four newspapers as Table 7 shows.

Table 6: Normed representations of hypotactic and rank shifted relationships

	DG	GT	DT	TT
Hypotactic	42.1	38.1	46	42.9
Embedded	57.9	61.9	54	57.1
	100	**100**	**100**	**100**

To arrive at these normed rates in Table 7, the raw attestations were converted to rates of 100 sentences per newspaper (c.f. 7.2.2 above). This is the rate at which a hypotactic or rank shifted relationship occurs in every 100 sentences. The evidence in Table 7 shows that the distribution of hypotactic and embedded/ rank shifted subordinate clauses in the editorials is independent of newspaper type. The consistent dominance of clause embedding across the four newspapers is perhaps an indication of the nature of the structural complexity of the newspaper editorial register. It seems rankshifting is among the linguistic means through which newspaper editorials achieve their communicative purpose. Put properly, rankshifting is more suitable to the situational context of the editorial register. However, the prevalence of rankshifted clauses is not as outstanding as that of the dominant declarative-complex sentence types. It is observable from Table 7 that even hypotactic clauses are appreciably attested across the four newspapers. This is, perhaps, an instructive indication that with ample data, new insights may emerge in the use of hypotactic and embedded structures within and across registers.

9. Implications of the findings to media discourse

The findings of this study largely support the argument in the literature about the language of newspaper editorials. Until the advent of the World Wide Web, which allows newspapers to publish online, many aspects of the newspaper, including editorials, were deemed formal, elaborated and abstract (Biber 1988; Wiredu 2012; Fosu 2016; Frimpong 2017) not only because they were written texts but also because there was no direct interaction between the addresser and the addressee. And so complex declarative sentences were suitable for the situational context of newspaper editorials.

During 2012 when the data were collected, the two Ghanaian newspapers studied in this paper were not actively online. Their British counterparts, however, were online with columns for their reading audience for correspondence. This crucial differentiation in the situational context of editorials from the two sociocultural contexts may be among the factors that have accounted for the disparity in the distribution of sentence types (especially simple and interrogative sentences). Thus, the higher representations of simple and interrogative sentences in the British newspapers, perhaps, are reflections of the consciousness of British editorialists of feedback from their reading audience. And this observation is important because as newspapers get closer to their audience, especially through the medium of the Internet and

other social media platforms, adjustments in situational context is expected to reflect adjustment in situational context.

10. Limitations of the study

It must be acknowledged that these sentence patterns have been compared in spite of variability in a number of contextual dynamics. And that is the audacity of the register theory which argues that in spite of these variations, choices of major linguistic elements in a register across different dialects of a language are normally similar (Biber & Conrad 2009).

11. Conclusion

This paper has investigated the usage dynamics of the sentence in editorials from four newspapers across Ghanaian and British sociocultural contexts from register perspectives. It has been observed that the declarative and complex sentences are the dominant types across editorials from the four newspapers. This pattern of distribution, it has been argued, supports arguments in the register theory that since texts from the same register are produced in similar situational contexts, linguistic in texts from the same register are largely similar across dialects of a language. Though the Chi-square tests confirmed that any observable variation in distribution is statistically insignificant, we note that minor irregularities observed in the distribution of simple and interrogative sentences across the two sociocultural contexts reflect adjustments in the situational and sociocultural contexts of the editorials investigated.

References

Ansah, Gladys. 2008. Linguistic Diversity in the Modern World: Practicalities and Paradoxes. *The International Journal of Language Society and Culture* 26.

Ansary, Hasan, & Babaii, Esmat. 2009. A Cross-cultural Analysis of English Newspaper Editorials: A Systemic-Functional View of Text for Contrastive Rhetoric Research. *RELC, Vol. 40* (2), 211-249.

Asante, Clement E. 1996. *The Press in Ghana: Problems and Prospects.* New York: University Press of America.

Biber, Douglas. 1988. *Variation across speech and writing.* Cambridge: Cambridge University Press.

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.

Chan, Tak Wing & Goldthorpe, John. 2007. Social Status and Newspaper Readership. *American Journal of Sociology, Vol 112* (No. 4), 1095 - 1134.

Chien, Shouchun. 2014. Varieties of English: Taiwanese attitudes and perceptions. *Newcastle and Northumbria Working Papers in Linguistics*, Vol 20, 1-16

Crystal, David & Davy, Derek. 1973. *Investigating English Style.* London: Longman.

Dako, Kari. 2003. *Ghanaianisms - a Glossary*. Accra. Ghana University Press.

Dako, Kari & Huber, Magnus. 2004. Ghanaian English: morphology and syntax. In: Bernd Kortmann and Edgar W. Schneider (Eds), *A Handbook of Varieties of English*. A Multimedia Reference Tool. Vol. 2: Morphology and Syntax, 854 – 865. Berlin: Mouton de Gruyter.

Daily Graphic. 2017. Graphic Communications Group Limited. http://www.graphic.com.org.gh.

Dakubu, Mary Esther Kropp. 1988. *The Languages of Ghana*, London, Kegan Paul International.

Denkabe, Aloysius & Gadzekpo, Audrey. 1996. *WHAT IS FIT TO PRINT? Language of the Press in Ghana* . Accra: FES.

Downing, Angela & Lock, Philip. 2006. *English Grammar: A University Guide* (2nd Edition ed.). London: Routledge.

Eggins, Suzanne. 2004. *An Introduction to Systemic Functional Linguistics* (2nd Edition ed.). London: Continuum International Publishing Group.

Fawcett, Robin. 2000. *A Theory of Syntax for Systemic Functional Linguistics.* Amsterdam: John Benjamin.

Finegan, Edward. 2012. *Language: Its Structure and Use* (6th ed.). United States: WADSWORTH CENGAGE.

Flowerdew, Lynne. 2004. The Argument for Using English Specialized Corpora to Understand Academic and Professional Language. In U. Connor and T. Upton (eds) *Discourse in the Professions: Perspectives from Corpus Linguistics*. Amsterdam: John Benjamins, pp. 11–33.

Fosu, Modestus. 2014. *The Press and Political Participation: Newspapers and the Politics of Linguistic Exclusion and Inclusion in Ghana*. The University of Leeds, Institute of Communication Studies. Unpublished PhD Thesis.

Fosu, Modestus. 2016. A linguistic description of the language of Ghanaian newspapers: implications for the readability, comprehensibility and information function of the Ghanaian press. *Ghana Journal of Linguistics* 5.1: 1 - 36.

Frimpong, Kodie George. 2015. *A Comparative Register Analysis of Editorials from Ghanaian and British Newspapers.* University of Ghana, Department of English. Accra: Unpublished PhD Thesis.

Frimpong, Kodie George. 2017. Subordination Across Ghanaian and British Newspaper Editorials: A Register Perspective. *Ghana Journal of Linguistics* 6(1). 72 - 115.

Frimpong, Kodie George. 2018. The complex sentence across written genres from native and nonnative contexts; a corpus-based study. *SKASE Journal of Theoretical Linguistics* 15. 3: 185 - 203

Ghanaian Times. 2017. New Times Corporation Ltd. http://www.ghanaiantimes.com.gh

Greenbaum, Sydney. 1996. *Oxford English Grammar.* Oxford: Oxford University Press.

Guerini, Federica. 2008. Multilingualism and language attitudes in Ghana: a preliminary survey. *ETHNOREMA* IV. 4: 1- 33

Halliday, Michael A.K. & Matthiessen, Christian. 2004. *An Introduction to Functional Grammar* (3rd Edition.). London: Hodder Education.

Halliday, Michael 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. 1999. *Ghanaian Pidgin English in its West African Context*. Amsterdam: Benjamins

Huber, Magnus. 2012. Syntactic and variational complexity in British and Ghanaian English: Relative clause formation in the written parts of the International Corpus of English. In B. Kortmann, & B. Szmrecyani (Ed.), *Linguistic Complexity: Second Language Acquisition, Indigenization, and Contact* (pp. 218-242). Berlin: Walter 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* (pp. 86-106). Amsterdam: John Benjamin.

Huddleston, Rodney. 1999. Sentence Types and Clause Subordination. In K. Brown, & J. Miller (Ed.), *Concise encyclopedia of grammatical categories* (pp. 329- 343). Amsterdam: Elsevier.

Kachru, Braj (1981). The Pragmatics of Non-Native Varieties of English. In Smith L. E. (ed.). English for Cross-Cultural Communication (15 - 39). London: Palgrave Macmillan.

Koester, Almut. 2010. Building small specialised corpora. In A. O'Keefe & M. McCarthy (Eds.) *The Routledge Handbook of Corpus Linguistics* (pp. 66 - 79). Routledge: London and New York.

Kuhi, Davud, Mojood, Manijheh. 2014. Metadiscourse in Newspaper Genre: a Cross-linguistic Study of English and Persian Editorials. *Elsevier: Procedia - Social and Behavioral Sciences* 98, 1046 – 1055.

Le, Elisabeth. 2009. *Editorials' genre and media roles: Le Monde's editorials from 1999 to 2001. Journal of Pragmatics* 41 1727–1748.

McEnery, Tony, & Gabrielatos, Costas. 2006. English Corpus Linguistics. In B. Aarts, & A. McMahon (Ed.), *The Handbook of English Linguistics* (pp. 33-71). Malden: Blackwell Publishing.

Morison, Stanley. 1932. *The English Newspaper, 1622 - 1932: An Account of the Physical Development of Journals Printed in London.* Cambridge: Cambridge University Press.

Obeng, Samuel. Gyasi. 1997. An Analysis of the Linguistic Situation in Ghana. In *African Languages and Cultures,* Vol. 10, No. 1 (1997), 63-81.

Oxford University Press. 2015. *Oxford English Dictionary.* Oxford: Oxford University Press.

Payne, Thomas. E. 2011. *Understanding English Grammar: A Linguistic Introduction.* Cambridge: Cambridge University Press.

Quinn, Anthony. 2013. *Magforum.* Retrieved March 3, 2015, from Magforum Web site: http://www.magforum.com.

Quirk, Randolf & Greenbaum, Sydney & Leech, Geoffrey & Svartvik, Jan. 1985. *A comprehensive grammar of the English language.* Edinburgh: Longman Group Limited.

Sadock, Jerrold & Zwicky, Arnold. 1985. Speech Acts Distinctions in Syntax. In S. T. (Ed.), *Language Typolohy and Syntactic Description. Vol. 1.* Cambridge: Cambridge University Press.

Sey, Kofi. 1973. *Ghanaian English: An exploratory survey.* London: Macmillam.

Schmeid, Josef. 1990. Corpus linguistics and non-native varieties of English. *World Englishes* Vol 9:3, 255 - 266.

Schneider, Edgar. W. 2007. *Postcolonial English: Variations around the World.* Cambridge: Cambridge University Press.

Silverman, David. 2011. *Interpreting Qualitative Data.* London: SAGE.

Simo-Bobda, Augustine. 1997. English in a multilingual society. In M. E. Kropp Dakubu (Ed.), *English in Ghana.* Accra: Ghana English Studies Association

Shokouhi, Hossein. & Amin, Forough. 2010. A Systemist 'Verb Transitivity' Analysis of the Persian and English Newspaper Editorials: A Focus of Genre Familiarity on EFL Learner's Reading Comprehension. *Journal of Language Teaching and Research*, Vol. 1, No. 4, 387-396

The Daily Telegraph 2017. Telegraph Media Group Limited. www.telegraph.co.uk

The Times. 2017. Times Newspapers Limited. http://www.thetimes.co.uk

Van Dijk, Teun. 1989. Race, riots and the press: An analysis of editorials in the British press about the 1985 disorders. *Gazette* (43), 229 - 253.

Wardhaugh, Ronald. 1995. *Understanding English Grammar.* Oxford: Blacwell.

Westin, Ingrid & Geisler, Christer. 2002. A multi-dimensional study of diachronic variation in British newspaper editorials. *ICAME Journal* (No 26), 133 - 152.

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.

Yankson, Isaac Kofi & Browne, Edmund; Tagor, H; Donkor, Peter; Quansah, Robert; Asare, George Ernest; Mock, Charles & Ebel, Beth. 2010. Reporting on road traffic injury: Content analysis of injuries and prevention opportunities in Ghanaian newspapers. *Injury Prevention.* 16. 194-197.

Young, Johnstone Tony & Walsh, Steve. 2010. Which English? Whose English? An investigation of 'non-native' teachers' beliefs about target varieties. *Language, Culture and Curriculum*, Vol 23: 2, 123-137

APPENDIX 1: Editorial Headlines and other corpus information

Newspaper	No of words)	Editorial Headline
The Daily Telegraph		
DT01	443	Osborne bears good tidings to the East

DT02	697	Rich or poor, we are all paying too much tax
DT03	434	A shabby episode that Cameron may regret
DT04	441	These Tory MPs are in tune with the country
DT05	686	This is not the time for surrender on the NHS
DT06	441	Faith must not be driven from Britain's public life
DT07	435	A tax that stalls the engines of recovery
DT08	715	A pointless, politicized plan for Lords reform
DT09	446	A budget for Britain, not for the Lib Dems
DT10	442	Voters have no time for political manoeuvring
DT11	434	We must stand up to the Strasbourg court
DT12	714	Britain's cities should say yes to mayors
DT13	754	Re-elect Boris – then give him more powers
DT14	719	There is still a route to a Conservative majority
DT15	433	Stability has been the Coalition's key success
DT16	445	The baroness and a case of double standards
DT17	440	The comfortable course is no longer an option
DT18	421	Regional pay plan scuppered by politics
DT19	711	The corrosive crisis of trust in our institution
DT20	445	Tories must resist Clegg's ridiculous blackmail
DT21	448	Clegg's foolish fantasy comes crashing down
DT22	437	Open government is here to stay, like it or not
DT23	448	Clegg shows his true colours on equal votes
DT24	439	Clegg cares more about revenge than fairness
DT25	437	A new team – but the problems don't change
DT26	441	More can still be done to get Britain growing
DT27	434	This defence deal is fraught with danger
DT28	447	Labour cannot be trusted to rebuild Britain
DT29	593	So much for the Tories' 'secret weapons'
DT30	441	Sniping at the rich is futile and damaging
DT31	436	Nick Clegg is on the wrong side of history
DT32	710	Until voters feel involved, localism is a lost cause
DT33	443	Vince's mansion tax rises from the dead
DT34	438	Osborne is still spending more than we can afford
DT35	711	A lot of froth over Starbucks and tax
DT36	456	Leaders' debates are good for democracy
The Times		
TT01	555	University Challenge
TT02	620	Devolutionary Road
TT03	561	Capitalism and Labour
TT04	615	No Escape
TT05	627	The Politics of Power
TT06	558	Revising Legislation
TT07	640	Court Out
TT08	634	A Budget for Aspiration

TT09	638	Special Effects
TT10	545	Petrol Gauge
TT11	633	Charity Case
TT12	668	The Trouble Ahead
TT13	559	The Choice for London
TT14	637	Leading the Argument
TT15	612	The Message
TT16	628	The Hot Potato
TT17	564	Cracking the Code
TT18	562	In, Out, Shake It All About
TT19	620	Politicians and Bankers
TT20	643	House of Cards
TT21	618	A Risible Delay
TT22	571	Avoiding the Issue
TT23	623	Judo Diplomacy
TT24	652	Taxing the Rich
TT25	617	Changing Places
TT26	565	A Sorry Affair
TT27	631	The Party is Over
TT28	637	Vintage Labour
TT29	650	Party Politics
TT30	621	Defensive Merger
TT31	620	Labour and Europe
TT32	558	Working Together
TT33	637	A Plane Mistake
TT34	614	Aid and Abetting
TT35	659	Home Truths
TT36	543	The Difficult Decade
Daily Graphic		
DG01	495	Everybody's Peace
DG02	561	Preparing for Election 2012
DG03	558	Strengthening internal Democracy
DG04	544	Good Move but…
DG05	624	PNC at the Crossroad
DG06	577	Polarization Destroys Ghana
DG07	489	Bright prospects await Ghana
DG08	551	Guaranteeing continuity in change
DG09	532	Working towards peaceful polls
DG10	647	More Room for Improvement
DG11	568	Sanity must Prevail
DG12	496	Resolving Electoral Blues

DG13	564	Well done, EC, but…
DG14	569	Strengthening Ghana's democratic system
DG15	479	Another feather in Ghana's democratic cap
DG16	560	Towards peaceful transition
DG17	502	EC, fix this anomaly
DG18	536	Media shouldn't declare election results
DG19	505	Let's die for Ghana, not for…
DG20	577	A shot in the foot
DG21	511	Ghana in tears
DG22	566	Time to bond
DG23	528	We are of age
DG24	504	Towards credible polls
DG25	481	Patronise voters register exhibition
DG26	677	Peace before anything else
DG27	612	Make room for jaw-jaw, not…
DG28	527	Put Ghana first
DG29	739	Upholding campaign promises
DG30	557	Let's play by the rules
DG31	553	Deepening Ghana's democracy
DG32	497	Let's learn from the US elections
DG33	514	Let's support the NCCE
DG34	511	Let's lead by example
DG35	522	Resolving electoral disputes
DG36	577	Politicians should stop funding macho men
Ghanaian Times		
GT01	388	Protecting our infant democracy
GT02	379	The President's portraits and matters arising
GT03	390	Lessons from the primaries
GT04	442	Enough of the vain promises!
GT05	366	Time for mature politics!
GT06	377	Respect the president!
GT07	647	Ensuring a credible voters register
GT08	389	Getting all voters to register (1)
GT09	378	Getting all voters registered (2)
GT10	399	Making district assemblies efficient
GT11	442	A good decision, Mr Minister!
GT12	375	A worthy decision!
GT13	398	Who should resign
GT14	437	Let's have a united CPP
GT15	343	Registration of prisoners
GT16	373	Parliamentary fraud
GT17	340	Ending the political mudslinging

GT18	324	The dignity of the presidency
GT19	422	By-elections in election year?
GT20	350	Taking parliamentary business serious!
GT21	342	Dialogue is the answer!
GT22	350	Time to reflect on the economy
GT23	318	The President must rest
GT24	380	Ending the political mudslinging
GT25	383	Towards a credible voters register
GT26	340	Don't toy with our education!
GT27	361	Politicians and promises
GT28	384	Election 2012 and small arms proliferation
GT29	353	Don't sabotage our education!
GT30	375	Rid the electoral process of jokers
GT31	364	The issue of minors in election 2012
GT32	359	Keeping peace alive
GT33	339	The heat is on!
GT34	259	All must play by the rules
GT35	447	All must accept outcome of election
GT36	777	Election 2012; decision day today!

George Kodie Frimpong
Department of English
P. O. Box LG129
University of Ghana
Legon – Accra
Ghana – W/A
gfrimpong@ug.edu.gh
frimpong.kodie@gmail.com

Swear Words Used by Jordan Belfort in *The Wolf of Wall Street* Movie

Barli Bram & Puguh Kristanto Putra
Sanata Dharma University, Indonesia

This study analysed the swear words that were used by the main character Jordan Belfort in The Wolf of Wall Street movie. Data, consisting of 322 occurrences of swear words, were collected from the movie. The data source was the transcripts of The Wolf of Wall Street movie and discourse analysis was applied in this study to identify the language use in the movie. The data were analysed based on Pinker's (2007) typology. Results showed that Jordan Belfort used five types of swear words, namely descriptive (twelve times; 3.7 per cent), idiomatic (19 times; 5.9 per cent), cathartic (26 times; 8.1 per cent), abusive (29 times; 9 per cent) and emphatic swearing (236 times; 73.3 per cent).

Keywords: *main character, discourse analysis, language use, swear word, The Wolf of Wall Street*

1. Introduction

Swearing is a form of language which is often used by people in daily conversation. People swear in various contexts (see Culpeper, 1996 & Beers Fägersten, 2007), and when they are in a conflict, for instance, they will more likely use linguistic expressions which are often referred to as bad language because some people think that it is taboo (Karjalainen, 2002, p. 6; see also Stone, McMillan, & Hazelton, 2015; Van Lancker & Cummings, 2016; Andang & Bram, 2018; Kristiano & Ardi, 2018). Ljung (2007) notes that although swear words are associated with crime, rawness, and lack of raising and poor education (p. 35), they have become regular in people's speech. Nevertheless, in some countries, for instance, in Indonesia, using swear words is highly prohibited because they are considered as taboo – or at least impolite.

In his monologue entitled Class Clown, Carlin (1972) states that there are seven words categorized as swear words. The seven words are banned on television because they are considered aggressive words and they may hurt other people's feelings. The words are *shit, piss, fuck, cunt, cocksucker, motherfucker,* and *tits*. Since the monologue contains so many swear words, there appeared numerous complaints about it. Thus, the Federal Communications Commission (FCC), an agency to regulate communication in all 50 states, included in the District of Columbia and U.S. territories, responded to the complaints by declaring that the monologue is considered as indecent.

In the education world, swearing is also restricted. Pyle (2001, p. 616) explains that different contexts may give different rights when speaking in public school. However, Stenström, Gisle and Hasund (2002, p. 203) argue that the use of swear words is a way to establish group identity. Rayson, Leech, and Hodges (1997, p. 10) add that swear and taboo words are linguistic devices to establish boundaries between all members in a group and it is to affirm social norms for language use.

By exploring *The Wolf of Wall Street* movie scripts as the object of the study, the current researchers attempted to identify various aspects related to the use of swear and taboo words. The main reason for choosing the movie was the characters of the movie used many swear words when communicating with one another. In other words, the movie scripts were a rich

source of data containing swear words. Furthermore, the second reason for choosing the movie was that it was popular as it became the Oscar nominee for five categories in 2014 and won the Golden Globe for the best performance by an actor in a motion picture in 2014. However, the current researchers only focused on the swear words used by Jordan Belfort, the main character of the movie. The main character was chosen because this movie is a first-person perspective movie, so the main character has the biggest role in the movie. Moreover, Jordan Belfort used swear words in most of his daily conversations, even in formal ones. The fact that this character uses swear words in any situation does not mean that other characters do, or even tolerate that; the same is true of any real life situation as well (see De Klerk, 1991, 1992 & Beers Fägersten, 2012).

Based on the research background above, the research question was formulated as below: What types of swear words are used by Jordan Belfort in *The Wolf of Wall Street* movie? The objective of this research was to identify and analyse the types of swear words used by Jordan Belfort in *The Wolf of Wall Street* movie. Besides, this study aimed to help students to understand that there are some different types of swear words so that they can become wiser in using swear words and preventing misunderstandings in formal and informal contexts of conversations.

2. Literature review

2.1 *Swear words and taboo words*

Swear words are considered taboo, and therefore, it is important to know the concept of taboo itself. Freud (2004: 3) defines taboo in the context of ancient Aboriginal Australian, and explains that Aboriginal Australians profess totemism as their system of tribal life. Totems can be a form of animals, plants, and some natural phenomena which are believed as their "guardian spirit" protecting them from danger by its oracle. The totem symbolises a power to prevent incest among the totemic clans.

Hughes (2006: 462) says that taboo in present context transforms into some prohibitions and restrictions varying in all societies. It means that taboo in particular places will be different from other places. Hughes (2006: 463) adds that taboo is only strictly prohibited in printed words and broadcasting but observed less frequently in the conversations among men. Thus, it can be concluded that taboo is a prohibition based on people's belief and culture in a particular place.

Taboo words are part of language and they are a product of society. Lehtonen (2000: 10) explains that language can be an important factor to identify a group, group solidarity, and it can be the signal of difference. Language is not only a medium to communicate with each other, but also it is an inseparable part of being humans (see also Jay & Janschewitz, 2008). Nevertheless, Karjalainen (2002: 18) states that although swear words are taboo, not all taboo words are swear words. Ljung (2011: 4) further elaborates that taboo and swear words are two different cases, and proposes the following four criteria of swear words:

1. Swearing is the use of utterances containing taboo words.
2. The taboo words are used with non-literal meaning.

3. Many utterances that constitute swearing are subject to serve lexical, phrasal, and syntactic constraints which suggest that most swearing qualifies as formulaic language.
4. Swearing is emotive language: its main function is to reflect, or seem to reflect, the speakers' feeling or attitude (Ljung 2011: 4).

Based on Ljung's (2011) explanations above, it is clear that swear words can be used as metaphors to refer to anything else which may not be taboo. For example, the word *fuck,* which refers to "having sex with someone", but in the utterance "it is fucking brilliant!", the word *fuck* can be the medium or tool to emphasise something. Based on the fourth point, Ljung (2011: 4) states that swear words can be used as a tool to show the emotion or feeling of the speaker, especially their feeling about what they are talking about. As a symbol of emotion, swear words can be expletive interjections (to show anger, disappointment, and joy), emphasisers (to emphasise the meaning of the words that the speakers want to convey), and expletive slot-fillers.

2.2 *Types of swear words*

According to Pinker (2007: 219), there are five different ways which people can choose to express or use swear words, namely descriptively, idiomatically, abusively, emphatically, and cathartically.

a. Descriptive
Descriptive swearing is the use of swear words to convey their literal meaning. Giora (1997: 185) says literal meaning is the meaning of the word or sentence which is computed from the lexical meaning automatically associated with any entries before any extra interferences based on contextual assumptions have been derived. For instance, in the phrase or expression *oh shit,* the word *shit* refers to 'excrement'.

b. Idiomatic
Idiomatic swearing is the use of swear words as idioms, where idioms refer to expressions in the form of phrases which have a different meaning if they are translated as individual words. Therefore, an idiom cannot be separated or translated word by word, because it will have a different meaning. For instance, in the utterance *I am fucked up,* where *fucked up* means that the speaker is drunk. Nevertheless, the idiom *fucked up* cannot be separated as two words (*fucked* and *up*), because the separation may result in a different meaning from what the speaker intends to convey originally, namely *drunk* or *intoxicated*.

c. Abusive
People use swear words in order to offend or abuse other people. Mostly, this kind of expression is used to express the speakers' emotion, especially anger. The example of this type of swear word is the use of *fuck you!* by the speaker when talking to the addressee.

d. Emphatic
The role of this type of expression is to give a stronger emotion from the speaker. For instance, "your idea is *fucking* brilliant!" In this case, the speaker appreciates the idea, not having a negative thought or intention even though the speaker uses a swear word.

e. Cathartic

Pinker (2007: 223) argues that swear words are used to spontaneously express the emotional burst of the speaker in response to something that happens to him/her. People tend to swear when they are surprised or shocked. In this case, the speaker does not address their swearing to others. For instance, a person can swear by stating *oh, fuck!* to express a sudden pain or frustration (Pinker, 2007: 226).

3. Research method

This study was qualitative research. Onwuegbuzi and Weinbaum (2017: 364) say that qualitative research focuses on understanding social phenomena from the perspective of human participants in the study. Qualitative research provides detailed, in-depth, and complete information about the object being studied. This research focused on the use of language in a movie or linguistic study of a movie, and it was then categorized as discourse analysis. Stubbs (1983: 1) states that in discourse analysis, researchers investigate the use of language, and it is related with the interrelationship between language and society. The current research used the transcripts of the conversations in *The Wolf of Wall Street* movie as data. The researchers also watched the movie in order to know the expressions of Jordan Belfort when using swear words. The acting and the transcripts helped the researchers to feel the emotion of the actors and actresses, and it can help to analyse the types of swear words in the movie transcripts.

In order to collect the data, the current researchers took several steps. The first one was listing the swear words from the transcripts retrieved from https://subscene.com/subtitles/the-wolf-of-wall-street/english/1247746. Then, the researchers checked the utterances which contain swear words to identify their functions and meanings. From the utterances, the researchers classified the swear words based on the context. The researchers used Pinker's (2007) typology, namely descriptive, idiomatic, abusive, emphatic and cathartic in order to classify the swear words used in the movie.

To analyse the collected data, the researchers used Pinker's (2007) typology, which consists of descriptive, idiomatic, abusive, emphatic and cathartic categories. The analysis aimed to identify the categories of swear words uttered by the main character, namely Jordan Belfort. Then, the researchers made a list of swear words based on their categories.

4. Findings and discussion

The results showed that there were 322 occurrences of swear words used by Jordan Belfort, as presented in Table 1 below.

Table 1. Types of swear words used by Jordan Belfort

No.	Types of swearing	Occurrence	Percentage (%)
1	Descriptive swearing (D)	12	3.7 %
2	Idiomatic swearing (I)	19	5.9 %
3	Cathartic swearing (C)	26	8.1 %
4	Abusive swearing (A)	29	9 %
5	Emphatic swearing (E)	236	73.3 %
	Total	322	100 %

Based on Table 1, it can be seen that the type of swear word that was mostly used by Jordan Belfort, as the main character in *The Wolf of Wall Street* movie, was emphatic swearing (E), namely the use of swear words to convey a stronger emotion, followed by abusive swearing (A). In the third position was cathartic swearing (C). Then, idiomatic (I) and descriptive (D) swearing were used less than 10%. The results showed that Jordan Belfort – when communicating with people around him – used mostly emphatic swearing to express his feelings or stronger emotions.

The main character of *The Wolf of Wall Street* movie likes to use swearing expressions in his daily conversations as can be seen from the frequency of his swearing speeches. The current researchers might not contend that a social status limits the use of swearing expressions in this movie since almost all the characters are of the same social status. Nonetheless, the researchers only focused on the analysis of the main character. Furthermore, swearing expressions can be used by and directed to people from both the same and different social statuses.

Swearing expressions can also be expressed among people from different social statuses in their communication. For instance, people with higher social status can even swear to people from lower social status or vice versa. The following is the discussion on the types of swear words used by Jordan Belfort.

a. Descriptive swearing

The first type of swearing found in *The Wolf of Wall Street* is descriptive. Giora (1997) states that descriptive swearing is the use of swear words to convey its literal meaning (as cited in Pinker, 2007, p. 219). The conversation below exemplifies this type of swearing.

(1) Flight attendant*: May I see your boarding passes, sir?*
Jordan: *We have the boarding passes. Jesus Christ, I want to **fuck** her.*
(01:38:54 – 01:38:59)

Contextually, the word *fuck* itself has various functions and meanings. For example, when used in a phrase, the word *fuck* can also function to emphasize a point, as in *Ron is annoying as fuck*. Based on that example, the phrase *as fuck* conveys that Ron is enormously annoying since the *as fuck* phrase was inserted in the sentence.

The dialogue from *The Wolf of Wall Street* above shows that Jordan had the desire to make love with the flight attendant since he was under the influence of drugs he had taken

before. Jordan was intoxicated. Subsequently he was unable to think clearly, and as a result, he used those words when talking to the stewardess.

(2)   Donnie: *I hate that fucking dog.*
       Jordan: *Yeah, he is getting old and decrepit. Starting to **shit** in the house again.*
       (02:32:12 – 02:32:17)

The dialogue above happened when Donnie came to Jordan's house. He saw Jordan's pet and tried to play with him. Subsequently, Donnie did not like the dog as he might have his preferences for an animal. The word *shit*, which was used by Jordan, belonged to descriptive swearing as it implied the actual meaning since Jordan wanted to convey that his dog no longer can defecate properly in its cage as it got older.

b. Idiomatic swearing
       The second type of swear words that can be seen in the movie script was idiomatic. Fromkin, Rodman and Hyams (2003) contend that idioms are the expressions with fixed phrases, consisting of more than one word, which cannot be concluded from the meanings of the individual words. Consequently, they cannot be separated word by word since they will have a different meaning. The following dialogue illustrates the exertion of this type of swearing in the movie.

(3)   Naomi: *And you are still acting like an infant!*
       Jordan: ***Goddamn it**!*
       (01:13:15 – 01:13:18)

In this dialogue, it can be seen that Jordan used the word *goddamn it* as an expression of his anger since Naomi kept bursting the water into his face while asking him several questions related to Venice, a name which Jordan mentioned during his sleep. Jordan tried to recall his memory of Venice because he went home in the early morning and took several drugs. As a result, he did not remember about Venice and Naomi provoked his reaction and made him become angry and swear. The word *goddamn* itself belongs to idiomatic swearing since it will have a different meaning if it is detached word by word.

(4)   Jordan: *We liked to get as **fucked up** as possible during our business powers in order to stimulate our free-flowing ideas, which is why we were popping these ludes like they were M&M's.*
       (00:51:17 – 00:51:21)

The monologue above took place when Jordan told his team that tried to maintain their good work by consuming alcohol and drugs which would provide them with more ideas. The word *fucked up* in this monologue is classified as idiomatic swearing since it cannot be separated word by word.

c. Cathartic swearing
       Pinker (2007) explains that cathartic swearing is the use of swear words to express the speakers' emotional burst in response to something that happens to him or her (p. 215). It means

that it does not require addressees when the speakers are swearing. This type of swearing can be seen in the following dialogue.

(5) Robbie: *The brother, you put a skateboard on him, you strap him to a skateboard, you toss him down an alley at some pins.*
Jordan: *No, **shit**. That is interesting.*
(00:44:53 – 00:44:59)

Based on the conversation above, it can be seen that Robbie elaborated the plan which they were going to execute towards their competitor. He mentioned that he would tie Stratton to the skateboard. After paying attention to Robbie's explanation, Jordan automatically stated, 'No, shit. That is interesting'. In that sentence, the word *shit* did not have any certain meaning since there was no coherence to the following sentence. It is exemplified in order to express his stunned reaction to Robbie's plan.

(6) Donnie: *Actually, I got you a little present.*
Jordan: *You got me a present?*
Donnie: *I got you something.*
Jordan: *Oh, **fuck**. You are sweet.*
(00:26:04 – 00:26:08)

In the dialogue above, Donnie gave a little gift to Jordan because their company ran well. As a response to Donnie, Jordan said *oh, fuck. You are sweet*. In that sentence, the word *fuck* was used in order to illustrate that Jordan was surprised since Donnie gave him a present. The word *fuck* belonged to cathartic swearing because it was used in order to express Jordan's feeling in response to Donnie.

d. Abusive swearing
        Abusive swearing is the use of swearing in order to offend or abuse other people. This type of swearing is used to express the speakers' feeling, and it is directed to the addressees.

(7) Naomi: *Kiss you?*
Jordan: *Yeah, give me one… **Fuck** you!*
(01:14:57 – 01:15:01)

Abusive swearing was employed in the dialogue above. Jordan pronounced the words *fuck you* in order to swear at his wife. In the story, Jordan was angry at Naomi because she threw the water right into his face all over again. Naomi showed her anger since Jordan mentioned a name, Venice during his sleep. Jordan used the swear word to express his anger, and he thought that it might insult his wife since the word chosen was offensive.

(8) Chester: *Captain who?*
Jordan: *Captain Ahab. From the fucking… The book, **motherfucker**. From the book.*
(00:32:34 – 00:32:42)

The dialogue above happened when Jordan tried to highlight the method in order to get clients for their company. He had the insight from *Moby Dick*, a classic American book.

Unfortunately, his colleagues were unfamiliar with the book which in contrary to what Jordan imagined before. He thought that they would understand it since the book was widely known as the American masterpiece. Jordan stated the word *motherfucker* because he was angry since none of his colleagues knew *Moby Dick*. The word *motherfucker* itself belonged to abusive swearing as the chosen word was unpleasant.

e. Emphatic swearing

The fifth type of swearing that is portrayed in the *Wolf of Wall Street* is emphatic. This type of swearing is to give a stronger emotion from the speaker. It is employed to emphasise the meaning delivered to the addressee. The following conversation is the example of this type of swearing.

(9)   Donnie: *Your turn.*
Jordan: *I am not **fucking** doing this. You are out of your **fucking** mind.*
Donnie: *Smoke this shit, bro.*
Jordan: *No.*
(00:26:21 – 00:26:25)

In the dialogue above, Donnie persuaded Jordan to smoke marijuana since he knew that Jordan was stressful. In order to release the stress, Donnie offered Jordan to smoke the weed with him. However, Jordan was trying to refuse the offer as he knew the effect of smoking it. Although weed was legal in several states of America, people who consumed weed still have to hide from the authority. Subsequently, Jordan contemplated that fact and tried to decline the offer by using the word *fucking* in order to emphasise that he did not want to smoke that. The word *fuck* is often associated with a negative meaning. As a result, people often use this word to express their feeling especially when they are not in a good mood.

(10)   Denham: *I'm wearing the suit three days in a row… Yeah, you bet I do. I have thought about it before. Who would not, right?*
Jordan: *Right, who **fucking** would not?*
(01:33:00 – 01:33:05)

The conversation above took place when Jordan found out that FBI agent Denham wanted to have a broker's license as he dreamed of having a better life. Consequently, he ended up being a federal agent for the bureau and had not achieved his goal of being a broker in Wall Street. Jordan mentioned the word *fucking* while he gave his response to Denham as he wanted to highlight that no one had ever thought of being a broker in Wall Street.

5. Conclusion

Results showed that Jordan Belfort, the main character in *The Wolf of Wall Street* movie, used five types of swearing, namely descriptive swearing, idiomatic swearing, cathartic swearing, abusive swearing, and emphatic swearing. The least frequent swearing was descriptive, occurring 12 times (3.7 per cent) and the most frequent one was emphatic, appearing 236 times (73.3 per cent). In other words, based on the analysis, descriptive swearing, namely using swear words to convey literal meanings, was the least productive, and emphatic swearing, that is

using swear words to express stronger emotions and emphasise the meanings conveyed to addressees, appeared the most productive.

Acknowledgements

References

Andang, Kristina, & Bram, Barli. 2018. Swear words and their implications for English language learning-teaching. *LLT Journal: A Journal on Language and Language Teaching*, 21 (Suppl), 43-49. Retrieved from http://e-journal.usd.ac.id/index.php/LLT/article/view/1195

Beers Fägersten, K. 2007. A sociolinguistic analysis of swearword offensiveness. *Saarland Working Papers in Linguistics*, 1, 14–37.

Beers Fägersten, K. 2012. *Who's searing now?: The social aspects of conversational swearing*. Cambridge: Cambridge Scholars Publishing.

Carlin, George. 2012, July 27. *George Carlin – Class Clown* [Video file]. Retrieved from https://www.youtube.com/watch?v=j7YDCXS5gc4

Culpeper, J. 1996. Towards an anatomy of impoliteness. *Journal of Pragmatics*, 25(3), 349-367.

De Klerk, V. 1991. Expletives: Men only? *Communications Monographs,* 58(2), 156– 69.

De Klerk, V. 1992. *How taboo are taboo words for girls? Language in Society,* 21(02), 277–89.

Freud, Sigmund. 2004. *Totem and taboo*. London: Rutledge.

Fromkin, Victoria, Rodman, Robert, & Hyams, Nina. 2003. *An introduction to language* (7th Edition). Boston, MA: Thomson and Heinle.

Giora, Rachel. 1997. Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics* 8(3), 183-206.

Hughes, Geoffrey. 2006. *An encyclopedia of swearing: The social history of oath, profanity, foul language, and ethnic slurs in the English-speaking world.* New York: M.E.

Jay, Timothy & Janschewitz, Kristin. 2008. The pragmatics of swearing. *Journal of Politeness Research* 4, 267-288.

Karjalainen, Markus. 2002. Where have all the swearwords gone? An analysis of the loss of swearwords in two Swedish translation of J.D. Salinger's Catcher in the Rye. Pro gradu thesis, University of Helsinki). Retrieved from http://ethesis.helsinki.fi/julkaisut/hum/engla/pg/karjalainen/wherehav.pdf

Kristiano, Johan, & Ardi, Priyatno. 2018. Swear words in Bad Boys II: A semantic analysis. *LLT Journal: A Journal on Language and Language Teaching* 21(2), 191-198. Retrieved from http://e-journal.usd.ac.id/index.php/LLT/article/view/1586

Lehtonen, Mikko. 2000. *The cultural analysis of texts*. London: Sage.

Ljung, Magnus. 2007. *Swearwords*. Copenhagen: Nörhaven Paperback A/S.

Ljung, Magnus. 2011. *Swearing: A cross-cultural linguistic study*. London: Palgrave Macmillan.

Onwuegbuzie, Anthony, & Weinbaum, Rebecca. 2017. A framework for using qualitative comparative analysis for the review of the literature. *The Qualitative Report* 22(2), 359-372. Retrieved from https://nsuworks.nova.edu/tqr/vol22/iss2/1

Pinker, Steven. 2007. *The stuff of thought: Language as a window into human nature*. London: Penguin.

Pyle, Jonathan. 2001. Speech in public schools: Different context or different rights. *U. Pa. J. Const. L.* 4, 586 – 636.

Rayson, Paul, Leech, Geoffrey, & Hodges, Mary. 1997. Social differentiation in the use of English vocabulary: Some analyses of the conversational component of the British National Corpus. *International Journal of Corpus Linguistics* 2(2), 133-152.

Stenström, Anna-Brita, Andersen, Gisle, & Hasund, Ingrid Kristine. 2002. *Trends in teenage talk: Corpus compilation, analysis and findings*. Amsterdam: John Benjamins.

Stone, Teresa Elizabeth, McMillan, Margaret, & Hazelton, Mike. 2015. Back to swear one: A review of English language literature on swearing and cursing in Western health settings. *Aggression and Violent Behavior* 25, 65-74.

Stubbs, Michael. 1983. *Discourse analysis: The sociolinguistic analysis of natural language* (Vol. 4). Chicago: The University of Chicago Press.

Van Lancker, Diana & Cummings, Jeffrey L. 2016. Expletives: Neuro-linguistic and neurobehavioral perspectives on swearing. *Brains Research Review* 31, 83-104.

Barli Bram
Sanata Dharma University
Yogyakarta
Indonesia
barli@usd.ac.id

Puguh Kristanto Putra
Sanata Dharma University
Yogyakarta
Indonesia
puguhkris21@gmail.com

Interview with
Bas Aarts

PS
What brought you to linguistics?

BA
Well, linguistics runs in my family. My father worked as a professor of English at Nijmegen University, and his love of the English language was infectious! He tells me that he first became interested in English during the war. At the time my grandparents, who were living in a village near Maastricht in the Netherlands, hid English pilots whose planes had been shot down. The Netherlands was occupied by the Nazis at the time. My father had been particularly impressed by a pilot called Ernest Jackson whose RAF plane crashed in a wood not far from my father's village on 14 February, 1943. My grandfather took Jackson in, but the pilot's presence became known to the Germans, who arrested him. My father recalled the event, at a ceremony in 2017 to commemorate the shooting down of the plane:

> *Before the Germans arrived, Jackson gave my father some of his belongings: his watch, his boots, a compass and his signet ring with the letters E. J. My father passed that ring onto me and I have been wearing it for more than 40 years. I remember when the Germans came to make Jackson a prisoner of war. He made the Victory sign when they drove off. I was an eight-year-old boy at the time.*

This event made a great impression on my father and was probably one of the factors that led him to study English at University. And as they say: the apple doesn't fall far from the tree.

PS
You are Director of the Survey of English Usage. To what degree is it a managerial and to what degree a linguistic work?

BA
It's a mix of the two. With my team I need to spend a lot of time applying for grants. We've been quite successful in this, and over the years the grants have allowed us to build two major corpora the British Component of the *International Corpus of English* (ICE-GB) (www.ucl.ac.uk/english-usage/projects/ice-gb) and the *Diachronic Corpus of Present-Day Spoken English* (DCPSE) (https://www.ucl.ac.uk/english-usage/projects/dcpse/). These are grammatically analysed collections of language materials that contain a combined total of around 1.8 million words of written and spoken English. Although they are small by today's standards, these corpora are fully tagged and parsed (i.e. grammatically analysed). The SEU has one of the largest collections of *parsed* spoken English currently available anywhere in the world. Users can conduct complex and detailed searches of the corpus with the exploration software ICECUP (International Corpus of English Corpus Utility Program; https://www.ucl.ac.uk/english-usage/resources/icecup/) which can be used to find and extract instances of grammatical and lexical constructions in the corpora, and to track linguistic trends including changes in the contemporary language, as well as the use of contextual and frequency data for a vast array of linguistic features.

More recently we built the *Englicious* website (www.englicious.org) which makes a wide variety of innovative teaching materials freely available to pupils and teachers in primary and secondary schools. These include lesson plans, interactive exercises, projects, glossaries,

etc., as well as background materials for teachers to improve their subject knowledge of English grammar, e.g. videos. The example sentences used on the site come from the corpora based at the Survey of English Usage.

This year we celebrate the 60th anniversary of the Survey of English Usage, which was founded by Randolph Quirk. See https://www.ucl.ac.uk/english-usage/about/history.htm. We are organising a memorial event for him on 9 July 2019 (https://www.ucl.ac.uk/english-usage/events/quirk19.htm).

PS
You have significantly contributed to the compilation of the Diachronic Corpus of Present-Day Spoken English (DCPSE). How does this work gear with your research into grammatical phenomena?

BA
As mentioned above, both ICE-GB and DCPSE are tagged and parsed, and you can search them not only for individual words or parts of speech, but also for grammatical patterns, using the ICECUP software. In my research I have made extensive use of these resources, as they are a good way of finding authentic examples. To give you an example, some time ago I wrote an article about what I called 'binominal constructions' such as *a fool of a policeman* (https://bit.ly/2SffDTa). These are quite common in many languages, e.g. *een schat van een kind* (Dutch), *o pobre do Manuel* (Portuguese)*, la tonta de Juana* (Spanish). The question of which noun in these constructions is the head is a fascinating syntactic puzzle. In most of the scholarly work that has been published on this construction the data have been made up, and so somewhat artificial. In my work I used the Survey's corpus resources. A particularly interesting example that I came across is *a bitchy iceberg of a woman*. It's not a very politically correct example, but it's authentic, and grammatically it's interesting because here we have an adjective (*bitchy*) placed in front of a noun (*iceberg*) that it does not modify. I used authentic examples like this to argue that in binominal noun phrases the second noun is the head.

PS
You were Vice-President of the International Society for the Linguistics of English. Could you introduce this society to our readers? In what respect does it differ from the European Society for the Study of English (ESSE), or are these two societies complementary in their scope of activities?

BA
ISLE was founded because it was felt by many scholars working in the field of English linguistics that there was no society exclusively for them. We do of course have the Linguistic Society of America, the Linguistics Association of Great Britain, the Societas Linguistica Europaea and so on, but these tend to be very large organisations with a wide range of linguistic interests, whereas ISLE is exclusively focused on English linguistics from Old English to Modern times. ESSE is also focused exclusively on English, but it is predominantly about literature. More recently ESSE has become more concerned with linguistics, especially because its President is a linguist. However, personally I feel more 'at home' at ISLE conferences than I do at ESSE conferences.

PS
You are founding editor of the English Language and Linguistics journal. What are your experiences with this work? This is connected with, inter alia, problems of finding well-

qualified reviewers willing to participate in the peer-reviewing process. In general, how do you evaluate the situation in linguistic journals?

BA

I founded *English Language and Linguistics* in 1997 with David Denison and Richard Hogg because we felt that there was no journal that catered exclusively for the field of English linguistics. I was editor for almost 15 years and then reviews editor for another few years. This was hard work because you have to read a large number of manuscripts, send them out to readers and, if accepted, prepare them for publication. Towards the end of my stint I often found it difficult to find reviewers. This isn't surprising because reading manuscripts for journals is a lot of work. I found that colleagues differ enormously in how much time they are prepared to spend on manuscripts. A small handful were always very helpful, and even prepared to read revised version of papers. Others were less obliging. I think that the shortest review I ever received was just two lines. These days there are so many developments in journal publishing, especially open access publishing, that it's hard to keep up. We may well see changes in the way journal articles are reviewed.

PS
What is the main current focus of your linguistic research?

BA

In my current research I'm working on the lexical item *for*. In a book chapter that was recently published (https://bit.ly/2SN25nb) I argue that all instances of *for* are prepositions, including the *for* that is regarded as a subordinating conjunction (as in e.g. *We are happy for you to call us.*). I am next working on what I call the predicative use of *for* in constructions such as *I took him for a fool*.

PS
As it follows from our interview you are no doubt a versatile academic with a range of activities. Is there one you like most of all?

BA

I particularly enjoy conference visits because they give me the opportunity to meet new colleagues in the field and to catch up with old friends.

PS
Given your workload and your multifarious activities the question of free time necessarily arises. Do you have any free time? If so, how do you spend it?

BA

I really love reading about history, so when I go on holiday with my family I research the history of the place where we are going, and I start looking at maps, for which I have a passion. My children are not always as interested as I am in the castles and museums that my wife and I drag them to!