A Cross-theoretical and Cross-typological Survey of Lexical Integrity and the Nature of the Morphology–Syntax Interface

Synthesis Examination

Jared Desjardins

Table of Contents

1. Introduction

At its core, this paper is an exploration of the morphology-syntax interface, and whether the combination of linguistic units into words and phrases should be treated as distinct grammatical components or as parts of a uniform grammatical system. Specifically, the present survey takes lexicalism – a long-held theoretical position maintained by morphologists and syntacticians for over forty-eight years (outlined in **Section 1.1** and **1.2**) – as the lens through which to understand not only its role in linguistic theory, but the nature of the morphology-syntax interface and goals of grammatical theory in general. By considering the predictions present formulations of lexicalism (e.g. the Lexical Integrity Principle (Bresnan and Mchombo 1995)) make regarding morphosyntactic interaction (**Section 2**), both cross-theoretically in **Section 3** and cross-typologically in **Section 4**, it is shown that lexicalism does not properly characterize the morphology-syntax interface, and neither do linguistic theories based on lexicalist assumptions (**Section 5**).

1.1. Development of Lexicalism and the Lexical Integrity Principle

Beginning in the 1950s with the advent of early Transformational Grammar, grammar was comprised solely of phonology and syntax; morphology, which at that time was not considered a field of study unto itself, was distributed across morphophonology (part of the phonological component) and syntax, which, in addition to phrasal combination, "[generates] the grammatical phoneme sequences of the language" (Chomsky 1957:32). The domain of morphology reemerged when Chomsky (1970) first attempted to constrain the power of grammatical (syntactic) theory by removing certain linguistic phenomena from the syntactic component and treating it instead in a separate lexical component (i.e., the lexicon itself) (Aronoff 1976:6)[1].

[1] Aronoff (1976:5-6) observes earlier work in generative phonology likewise led to a reintroduction of morphology; specifically, Chomsky and Halle's (1968) attempts to constrain the power of phonological theory revealed that various phenomena (e.g. rules of inflectional morphology) were beyond a formal phonological theory (i.e. part of morphology).

Chomsky (1970) suggests that derivationally complex and idiosyncratic information is present in, and the product of, an expanded lexicon, and is therefore separate from the syntax and immune to syntactic transformations. As evidence, Chomsky cites three differences between English derived nominals, which he views as fully derived prior to lexical insertion and any subsequent syntactic procedures, and gerundive nominals, which he claims are formed syntactically. For the sake of brevity, the present discussion will focus on one of the three[2] points outlined by Chomsky; specifically, the structural behavior of derived nominals, as compared to that of gerundive nominals. Chomsky shows that the lexical category and internal structure of derived nominals is morphologically predetermined, allowing the derived nominal to participate freely in a range of syntactic configurations. This is in comparison to gerundive nominals, whose lexical category and combinatoric potential are syntactically determined, limiting their syntactic distributional properties. In particular, derived nominals are shown to have an NP internal structure (as determined by the lexicon) which allows them to be modified by a determiner or adjective in the syntax (**Example 1a** below), compared to the gerundive nominal in **Example 1b** which lacks this ability. Furthermore, the derived nominal is able to appear with the full range of determiners (**Example 2a**), while gerundive nominals cannot (**Example 2b**). This NP internal structure further allows derived nominals to appear freely in NP structures, where they can function as nominal arguments in, for example, active and passive ditransitive alternations (**Example 3a** compared to the gerundive nominal in **3b**). Chomsky therefore concludes that certain word forms – in this case, the derived nominal – are fully derived prior to lexical insertion and any subsequent syntactic transformations, and are therefore the product of a separate, ordered lexical component.

Example 1

a. The amusement of the children
b. *The amusing of the children

Example 2

a. John's three/several of John's proofs of the theorem
b. *John's three/several of John's provings the theorem

[2] The other two points Chomsky (1970:188-189) discusses include: (i) the productivity of the transformational process that produces gerundive nominals, as opposed to the apparent restrictions on the formation of derived nominals, and (ii) the generality of the relation between each nominal and the associated proposition; namely that derived nominals tend to be semantically idiosyncratic, as compared to the semantics of gerundives.

Example 3

a. John gave Bill advice | advice was given (to) Bill | Bill was given advice

b. *John gave Bill advising | *advising was given (to) Bill | *Bill was given advising

Chomsky's (1970:188) proposal that "we might extend the base [syntactic] rules to accommodate the derived nominal directly", which he terms the "lexicalist position", as opposed to syntactically deriving the nominal (i.e. the "transformationalist position"), established morphology as an autonomous domain distinct from syntax. In order to enforce and restrict the separation between these two grammatical components, linguists have offered various interpretations of Chomsky's original lexicalist position, two of which are provided below:

Interpretations of the Lexicalist Position

1. "[…] transformations do not perform derivational morphology." (Jackendoff 1972:12-13)
2. "Syntactic rules cannot make reference to any aspect of the internal structure of a word." (Scalise 1986:101)

The nature of the division between morphological and syntactic processes and the restriction on intramodular interaction imposed by the lexicalist position was further explicated and refined by Jackendoff (1972), Lapointe (1985[1980]), and Selkirk (1982), in what are referred to as the Extended Lexicalist Hypothesis, Generalized Lexicalist Hypothesis, and Word Structure Autonomy Condition, respectively.

Extended Lexicalist Hypothesis[3] (Jackendoff 1972:13)

"[…] transformations cannot change [syntactic] node labels, and they cannot delete under identity or positive absolute exception. […] the only changes that transformations can make to lexical items is to add inflection affixes such as number, gender, case, person, and tense."

Generalized Lexicalist Hypothesis (LaPointe 1985[1980]:8)

"No syntactic rule may refer to elements of morphological structure."

Word Structure Autonomy Condition (Selkirk 1982:70)

"No deletion or movement transformations may involve categories of both W[ord]-structure and S[entence]-structure."

[3] Spencer (1991:72-73) further elaborates on Jackendoff's (1972) Extended Lexicalist Hypothesis, noting "the content of [the Extended Lexicalist Hypothesis] is that transformations should only be permitted to operate on syntactic constituents and to insert or delete named entities (like propositions). This means that they can't be used to insert, delete, permute or substitute parts of words [which] in turn means that they can't be used in derivational morphology."

Di Sciullo and Williams (1987), in their Atomicity Thesis, emphasize the notion of accessibility of word-internal information, suggesting that not only is word-internal structure inaccessible to syntax but also that word-internal meaning components play no role in semantic composition at the phrasal level:

> **Atomicity Thesis** (Di Sciullo and Williams 1987:49)
>
> "Words are 'atomic' at the level of phrasal syntax and phrasal semantics. The words have 'features', or properties, but these features have no structure, and the relation of these features to the internal composition of the word cannot be relevant to syntax."

As the status of morphology as an autonomous module of a language's grammar remained the subject of debate throughout the late twentieth century, particularly as linguists explored the extent to which, and in what capacity, grammatical theory could be appropriately delimited and constrained, linguists took two divergent directions: on the one hand, in light of the apparent similarities between morphological and syntactic composition[4], developing highly syntacticized models in which word formation processes are handled by the syntactic module itself, and on the other, developing lexicalist approaches to word formation, which supported the view that morphology is largely an independent system similar to phonology (Borer 1998:151). As a result, a lexicalist spectrum[5] began to emerge depending on how distinct word formation is considered to be from syntax, and the related consideration as to how powerful the syntactic component should be with respect to morphology. On one end of the spectrum, all morphological composition is separate from syntax, and word-internal structure is completely invisible to syntactic processes; examples of these strong lexicalist hypotheses include the Atomicity Thesis of Di Sciullo and Williams (1987) above, as well as Anderson's (1992) formulation below:

> **(Strong) Lexicalist Hypothesis** (Anderson 1992:84)
>
> "The syntax neither manipulates nor has access to the internal form of words."

[4] These similarities include the pre-theoretical observation that both involve the combination of linguistic elements into meaningful constituents (words from morphemes, and phrases, clauses and sentences from words). Other similarities are theory-specific, such as the notion that words exhibit hierarchical structures containing heads and their projections (Williams 1981), subcategorization (Lieber 1980), recursion (Josefsson 1998) and incorporation and government (Baker 1985a, b) – all concepts traditionally employed in syntactic description.

[5] The lexicalist spectrum is evidently related to the Split Morphology Hypothesis (SMH) (Anderson 1982), and therefore each approach to morphology/syntax on the lexicalist spectrum will vary depending on the linguist's explicit or implicit assumptions regarding the SMH.

The other end of the spectrum constitutes approaches in which the independence of the morphological and syntactic components is still somewhat maintained, but syntax is in part responsible for morphological processes, or is permitted some degree of access to the internal composition of words. Examples of weak lexicalist hypotheses include Chomsky's (1970) original lexicalist position, as well as the formulations of Anderson (1982:587), who declares "inflectional morphology is what is relevant to syntax", and Aronoff (1976), provided below[6].

(Weak) Lexicalist Hypothesis (Aronoff 1976:8-9)

"[…] derivational morphology is never dealt with in the syntax, although inflection is, along with other such 'morphological' matters such as *Do* support, Affix Hopping, Clitic Rules, i.e. all of 'grammatical morphology'."

Borer (1998) notes that in both its strong and weak forms, lexicalism is enforced by maintaining a distinct lexicon and word formation component separate from syntax, and as a consequence, assumes an ordered relationship between each grammatical component. Borer therefore provides the following statement of lexicalism in terms of ordering, which is in turn closely related to the No Phrase Constraint of Botha (1981), also provided below.

Lexicalism in Terms of Ordering (Borer 1998:152-153)

"The [morphological] component and the syntax thus interact only in one fixed point. Such ordering entails the output of one system is the input to the other."

No Phrase Constraint (Botha 1981:152-153)

"Morphologically complex words cannot be formed […] on the basis of syntactic phrases."

Despite the apparent parallels (e.g. hierarchical structure) and points of contacts (e.g. inflection) between word formation and phrasal syntax, there are clearly operations and categories unique to each component of the grammar. This has led linguists to maintain the fundamental generalization that restricts the interface between the internal structure of complex words and the rest of the grammar. The foregoing principles governing the relationship between morphological and syntactic processes are now commonly grouped together under a single rubric, the Lexical

[6] It should be noted that the weak lexicalism of Aronoff (1976) and Anderson (1982) differ with respect to whether inflection is a syntactic operation (in the case of the former), and whether inflection is a phonological operation (in the case of the latter) (Scalise 1986:102).

Integrity Principle. As articulated by Bresnan and Mchombo (1995), this principle concerns the distinctness of syntactic and lexical composition:

> **Lexical Integrity Principle** (Bresnan and Mchombo 1995:181)
>
> "A fundamental generalization that morphologists have traditionally maintained is the *lexical integrity principle*, which states that words are built out of different structural elements and by different principles of composition than syntactic phrases. Specifically, the morphological constituents of words are lexical and sublexical categories – stems and affixes – while the syntactic constituents of phrases have words as the minimal, unanalyzable units; and syntactic ordering principles do not apply to morphemic structure."

The general notion of lexicalism[7] and the Lexical Integrity Principle (which will henceforth be referred to as Lexical Integrity (LI)) is therefore a modular conceptualization of morphology and syntax, which arose from the historical need for separation between the syntactic system and certain morphological properties, and which was developed through a series of attempts to constrain the types and degrees of interaction between those components.

1.2. LI Today and Its Challenges

While LI and its various formulations[8] have traditionally been applied holistically to morphology and syntax, it has long been understood that LI subsumes two separable points, first noted by Postal (1969): (i) some linguistic phenomena covered by LI involve the relationship between syntax and the lexicon (e.g., situations of apparent non-local dependencies, in which the syntax may access the semantic, sublexical components of words), while others refer to (ii) the relationship between morphology and syntax (e.g., situations in which morphological processes affect phrasal syntax, or syntactic procedures affect word formation). Beginning in the 2000s, these two observations led some to question whether these constraints on the relationship between syntax and the lexicon, and between syntax and morphology, are equally strong. Booij (2005, 2009) distinguishes two kinds

[7] Bruening (2018:2) notes an additional meaning to the term lexicalism, which involves theories with enriched lexical entries that perform important functions in the grammar. An example of such a lexicalist theory is Combinatory Categorial Grammar (CCG) (Steedman 2000)).

[8] There are other variants of lexicalism in addition to the ones presented and referenced in this paper; however, they all have in common some separation between a morphological and syntactic component.

of LI violations[9]: 'accessibility' violations (syntactic and semantic access of word-internal structure), and 'manipulation' violations (syntactic manipulation of word-internal structure). Booij claims that LI should allow for accessibility of word-internal structure to syntax and semantics, while disallowing manipulation, and suggests a revised version of LI:

> **(Revised) Lexical Integrity** (Booij 2005:14)
>
> "[…] syntax may need access to the internal structure of words […]. Hence the part of [LI] that forbids syntax to have access to word-internal structure appears to be incorrect."

Lieber and Scalise (2007:20-21) identify only what they consider to be "the strongest challenges to [LI]", which they broadly categorize in terms of their type of inter-component interaction – (a) morphology has access to syntax, (b) syntax has access to morphology, (c) morphology/semantics interactions, and (d) morphology/phonology interactions[10] – and argue for the following restatement of LI, which acts in tandem with an added formal morphological operation[11] to restrict (on a language-specific basis) the type and degree of morphosyntactic interaction.

> **Restatement of Lexical Integrity** (Lieber and Scalise 2007:21)
>
> "Syntax and morphology are normally blind to each other. However, limited intermodular access may be allowed by virtue of allowing [specific linguistic configurations]."

In line with Booij (2005, 2009) and Lieber and Scalise (2007), Bruening (2018) notes that LI disallows at least two separable types of linguistic phenomena: (1) cases where phrasal syntax can feed word formation (what he refers to as "Error 1" (pp. 2)), and (2) syntactic access of word-

[9] LI violations are formally introduced and discussed in **Section 2**.

[10] Lieber and Scalise (2007:20-21) state only categories (a) morphology has access to syntax, and (b) syntax has access to morphology, relate to the morphology-syntax interface – category (c) instead relates to the morphology-semantics interface and (d) to the morphology-phonology interface. Noting that the original formulation of LI is concerned only with the morphology-syntax interface, Lieber and Scalise therefore question whether any reformulation of LI should be concerned with examples of (b) and (d) at all.

[11] In particular, Lieber and Scalise (2007:21) propose the Principle of Limited Access, stated as "Morphological Merge can select on a language specific basis to merge with a phrasal/sentential unit; there is no Syntactic Merge below the word level", and define the operation Morphological Merge as "Let there be items α, β, such that α is a base and β a base or affix. [Morphological Merge] takes α, β (order irrelevant) and yields structures of the form $< \alpha, \beta > \gamma$, where γ is an X^0, categorically equivalent to α or β, and α or β can be null." The Principle of Limited Access maintains LI by identifying distinct Syntactic and Morphological Merge operations, that can interface on language specific bases, and stipulating that the syntactic form of Merge does not operate below word level. Together with the Principle of Limited Access, Morphological Merge derives words, phrasal compounds, etc., (i.e. γ, functioning as a terminal node X^0) according to language specific, abstract configurations of affixes, roots/stems, words, and/or phrases (i.e. α, β).

internal parts (referred to as "Error 2" (pp. 13)). Bruening shows that situations in which phrasal syntax feeds word formation conflict with a LI view that word formation and phrasal syntax are distinct, linear components since word formation can take fully formed syntactic phrases as input. For example, English permits phrasal derivations as below in **Example 4** (from Lieber and Scalise 2007:9), where the suffix *-ish* attaches to the fully-formed phrases (indicated in square brackets) *self-sufficient* and *Don's Long-Awaited Photo Tutorial*, respectively.

Example 4

a. [self-sufficient]-ish
b. [Don's Long-Awaited Photo Tutorial]-ish Thing

Additionally, situations where syntax accesses word-internal elements conflict with LI, since word-internal structure is held to be opaque to syntactic operations. A transparent example of syntactic access of sublexical parts is found among so-called anaphoric islands (Postal 1969), in which a pronoun may be co-referential with a word-internal element elsewhere in the sentence. In **Example 5** below, the pronouns *him* and *he* refer to the word-internal unit *Reagan,* located inside the derived word form *Reagan-ite*. Since coreference is traditionally understood as a syntactic relation (a binding relationship between NPs in a licit structural configuration), the cases in **Example 5** would appear to be ones in which phrase-level grammar has access to word-internal parts.

Example 5

a. $\text{Reagan}_i\text{ites no longer believe in him}_{is}$
b. he_i no longer believes in $\text{Reagan}_i\text{ites}$

However, Bruening (pp. 15-23) goes further by providing a third error type ("Error 3"), arguing LI incorrectly predicts that morphology and syntax are truly separate components and obey different sets of principles; specifically, examples that are claimed to adhere to LI (e.g. extraction, coordination, and ellipsis) are shown to follow from LI-independent principles, such as the distinction between heads and phrases. Bruening demonstrates that certain processes are only able to affect phrases (namely extraction, coordination, and specific types of ellipsis), while other

processes that can be taken as violating LI affect only heads (e.g. ellipsis in certain coordinate structures, such as word-part ellipsis[12] in *half-brothers and (~~half~~)-sisters*). Bruening therefore concludes that the predictions LI makes are both incorrect, since syntax has been shown to feed word formation and word-internal units can in fact be accessed by syntax, and do no explanatory work, since linguistic phenomena that appear to follow from LI can be accounted for through independent principles. In other words, no separable morphological component appears necessarily useful or essential to linguistic theory and explanation, leading Bruening to suggest a theory in which a single syntactic component is responsible for both words and phrases[13].

While it is clear that LI in its various incarnations is a significant underlying assumption in contemporary linguistic theory, the explanatory power and tenability of LI is poorly understood. Some claim it is essential, with morphology and syntax treated as distinct modules (e.g. Bresnan and Mchombo 1995, among others); some have claimed that, in light of the distinct linguistic phenomena that it covers, the principle must be maintained in certain cases and abandoned in other situations (e.g., Booij 2005, 2009; Lieber and Scalise 2007); and others argue that LI is unnecessary even when distinct classes of violations are recognized (e.g. Bruening 2018). Consequently, the relevance of LI to linguistic theory is nebulous, particularly since it is often simply not clear what linguistic phenomena LI is supposed to disallow, what constitutes a violation of LI, and why. For example, LI violations may manifest at the border of word formation and phrasal combination (e.g. the English phrasal derivations in **Example 4** above), or as issues of interpretation rather than form, as in the case of anaphoric islands outlined above in **Example 5**. LI violations have also been observed in major morphosyntactic patterns, as in the positioning of person markers (also referred to as endoclitics) in Udi (Harris 2002). Udi person markers can occur as a suffix between other bound morphemes on a verb, as for example the third person singular marker *-ne* in **Example 6a**, but also morpheme-internally as in **Example 6b**, in which *-ne-* appears within the verb root *ef* 'keep'. However, person markers also appear on words external to the verb,

[12] So-called word-part ellipsis, and how it can be reanalyzed in terms of suspended affixation, is further discussed below in **Example 7**, and similar cases in **Section 2**.

[13] Müller (2018) provides a reply against Bruening's (2018) strictly syntactic approach, concluding that there is indeed a principled distinction between morphology and syntax.

such as **Example 6c** where *-ne* is attached to the noun *ait* 'word', which functions syntactically as the direct object.

Example 6

a. aš-ne-b-sa
work-3.SG-LV-PRES
'she works'

b. e_i-ne-f_i-sa
$keep_i$-3.SG-$keep_i$-PRES
'she keeps (it)'

c. merab-en ayt-ne ef-sa
Merab-ERG word.ABS-3.SG keep-PRES
'Merab keeps his word'

Therefore, the placement of Udi person markers must be at least partially syntactic, since they can occur at the edge of an apparent VP, in much the same way that a clitic might (e.g. the *'s* genitive 'suffix' in English, which can occur at the right edge of NPs, as in *NP[the person on the street]'s opinion*). It is then unclear what the proper analysis of an inflectional marker that sometimes behaves like an inflectional suffix and infix, and sometimes like a syntactic constituent, should look like, and whether it is in fact a violation of LI. Presenting a somewhat similar problem as the Udi data above, examples of coordination and ellipsis of word-parts, such as **Example 7** below, can be analyzed in terms of a syntactic process whereby entire phrases are conjoined, followed by the deletion of word-parts (Chaves 2008) – in this case, a second *pro* in *pro-choice and (~~pro~~)-gun control*. If viewed from a strictly morphological viewpoint, the ellipsis of word-parts would constitute a syntactic manipulation violation of word structure.

Example 7

Pro-choice and gun control

Conversely, the same phenomena can be viewed in terms of suspended affixation (Spencer 2005:83), where a single affix (*pro-* in the example above) has scope over two or more conjoined words (*choice* and *gun*), constituting an apparent morphological process and syntactic access violation at the phrase-level. Considerations such as the genesis and development of LI, its various (weak and strong) formulations, recent challenges and reformulations, and apparent violations,

suggest that LI is a good lens through which to view the goals of linguistic theory and explanation in general, and the nature and efficacy of the syntax-morphology distinction specifically.

2. Violations of LI

It is important to first note that the very notion of LI violation is enigmatic and theoretically sensitive (a point which is discussed below in **Section 3.2** and again in **Section 5**). Therefore, this section has two preliminary goals, which will serve as the foundational assumptions throughout the remainder of this paper: (i) a formal, working definition of LI will be provided, and (ii) a taxonomy of violation types is introduced, along with linguistic data examples demonstrating how such a definition of LI in (i) facilitates the classification of distinct violation types.

Assuming a strongly lexicalist position in the spirit of Bresnan and Mchombo (1995:181) and Anderson (1992:84), in connection with the ordered modularity presupposed by LI (e.g. Botha 1981:18, Borer 1998:152-153), allows one to formulate certain predictions regarding the morphology-syntax interface. These predictions are presented below, and an attempt is made to relate each with the observations on LI of Postal (1969), Booij (2005, 2009), Lieber and Scalise (2007), and Bruening (2018)[14]:

Predictions Made by LI

1. Syntactic processes (e.g. movement, deletion, etc.) cannot affect a proper subpart of a word (cf. Postal's point (ii), Booij's 'manipulation', Lieber and Scalise's category (a)).
2. Syntactic (and semantic) interpretation rules may not see a proper subpart of a word (e.g. anaphora, semantic scope, agreement between lexical units, etc.) (cf. Postal's point (i), Booij's 'accessibility', Lieber and Scalise's categories (a) through (c), and Bruening's Error 2).

[14] In several cases a direct relation between challenges to LI cannot be determined, largely due to theoretical disagreements on what constitutes a morphological and/or syntactic process, how strongly or weakly lexicalist one's assumptions are, whether semantics is considered separately from morphology and syntax, how fine-grained one chooses to be when distinguishing individual challenges to LI, etc. For example, Lieber and Scalise's (2007) category (a) (morphology has access to syntax) subsumes phenomena which Booij (2005, 2009) treats as belonging to both 'accessibility' and 'manipulation' violations, while Booij's 'accessibility' and 'manipulation' is silent on the ordering relationship noted by Bruening (2018) (Error 1, which is itself related to Lieber and Scalise's category (a)). Additionally, Bruening's Error 3 (morphology and syntax obey separate principles) is addressed by neither Booij nor Lieber and Scalise (but possibly by Postal's initial observations). This should serve to illustrate the tenuous nature of both LI and the precise linguistic phenomena it is intended to prevent.

3. Morphology and syntax is unidirectional (e.g. morphological processes, such as derivational or inflectional affixation, conversion/zero derivation, etc.) cannot affect phrase-level constituents. In the case of weakly lexicalist approaches in which inflection, but not derivation, is determined by syntax, this predicts that inflectional markers should always occur external to derivational markers[15] (Beard 1998:45-46) (cf. Bruening's Error 1 and Lieber and Scalise's category (a)).

The predictions made by a strongly lexicalist, modular view of LI facilitates the development of a taxonomy of apparent LI violations, provided in **Figure 1**[16]. Corresponding to each aforementioned prediction, three general violation types are identified: MANIPULATION violations, which involve syntactic processes that are typically assumed to operate on individual word forms (i.e. terminal nodes) and phrasal constituents; ACCESS violations, which involve bidirectional, intermodular interaction between morphology, syntax, and semantics; and ORDER, which involve situations where observable linguistic forms contradict the supposed linear order of morphology and syntax. In some, if not most cases, specific violations and their overall type can be reanalyzed in terms of an opposite perspective, as noted above regarding word-part ellipsis/suspended affixation in **Section 1.2**, **Example 7**. These points of potential reanalysis are indicated in **Figure 1** by means of the co-indexed superscript symbols †, ‡, and §[17].

[15] Interestingly, this theoretical prediction is at least somewhat typologically supported by Greenberg's (1963) observation that derivational morphemes tend to appear closer to the root than inflectional affixes (Booij 2009:1). However, it is the (very preliminary) hypothesis of the author that this may have more to do with a cognitive distinction between general lexeme (or concept) formation and the overt expression of individual word forms of the same lexeme, as opposed to evidence for LI and the strict modularity of a morphological and syntactic component in relation to overt linguistic forms.

[16] The full taxonomy of LI violations, which provides additional detail regarding individual languages, their specific violation types, and their scholarly sources, is omitted from the present discussion due to size and formatting. However, the full taxonomy is provided in **Appendix 7.2**.

[17] † marks situations in which a syntactic MANIPULATION (word-part ellipsis) violation can be morphologically analyzed in terms of a syntactic/semantic ACCESS violation of word-internal structure. ‡ indicates that specific phenomena constituting apparent phrase-level ORDER violations (e.g. bracketing paradoxes) may be recharacterized as particular forms of ACCESS violations. Finally, § represents the fact that both ACCESS subtypes (syntax/semantics 'sees' morphology, and morphology 'sees' syntax/semantics) could possibly be collapsed into a single ACCESS violation, as opposed to specifying the directionality of the alleged access. For all intents and purposes, each potential violation type and subtype is kept distinct; however, future work will ideally explore the nature of these violation types (and subtypes) in more detail.

Figure 1 Taxonomy of LI Violations Types

1. MANIPULATION		
	a. Movement (into and out of words)	
	b. Word-part ellipsis †	
2. ACCESS ‡		
	a. Syntax/semantics 'sees' morphology §	VP nominalization and constituent order
		Anaphoric islands
		Suspended affixation (*gruppeninflection*) †
		Sublexical indexing
		Semantic scope phenomena
		Focus targeting sublexical units
		Sublexical access for case assignment
	b. Morphology 'sees' syntax/semantics §	Construction dependent morphology
		Agreement and verb position
3. ORDER		
	a. Word-level phenomena	Order of inflectional and derivational affixes
	b. Phrase-level phenomena ‡	Phrasal compounding
		Words zero-derived from phrases
		Words derived from constructions
		Derivational morphology on phrases
		Inflectional morphology on phrases
		Bracketing paradoxes

Furthermore, the general LI violation types MANIPULATION, ACCESS, and ORDER, each comprise at least two subtypes. For example, MANIPULATION may include instances of apparent syntactic movement into words, as in the example of Mohawk noun incorporation (Baker 1985a:38) in **Example 8** below. Baker argues that the relationship between the sentences in **8a** and **8b** is a derivational one, with **8a** functioning as the input, and syntactic (head) movement of the lexical unit *nuhs* 'house' onto the verb root *raky* 'be white' producing the observed output in **8b**[18]. Therefore, instances of noun incorporation and the like (e.g. Udi person markers **Example 6a-c** above) are considered to be a specific subtype of MANIPULATION – particularly, movement (into and out of words).

[18] If certain weakly lexicalist approaches are adopted, this would not theoretically constitute a violation, since it is stipulated (e.g. in Aronoff's (1976:8-9) formulation of the (weak) lexicalist hypothesis) that syntax deals with other situations of movement of morphemes, such as affix hopping and clitic rules.

Example 8

a. ka-rakv ne sawatis hrao-nuhs-aʔ
3.NEUT-be.white DET John 3.MASC-house-SUFF
'John's house is white'

b. hrao-nuhs-rakv ne sawatis
3.MASC-house-be.white DET John
'John's house is white'

A second subtype of MANIPULATION includes linguistic phenomena which can be fully analyzed as either a syntactic or morphological process. As previously noted in **Section 1.2** (**Example 7**), examples such as those presented below in **Example 9**[19] can be characterized as a syntactic process of word-part ellipsis, where a particular morphological element is repeated across words, and subsequently deleted by the syntax. From a strictly morphological perspective, examples of coordination and deletion of word-parts then constitute a form of MANIPULATION violation, and it is necessary to distinguish a second subtype of MANIPULATION violations to include potential instances of word-part ellipsis. However, considering that word-part ellipsis readily yields to a morphological analysis in terms of suspended affixation, where, for example, the Portuguese adverbial ending *-mente* in **9d**, and the Turkish locative suffix *-da* in **9e**, take scope over the preceding conjoined adjectives $_{ADJ}$*[segura] mas* $_{ADJ}$*[lenta]* 'safe but slow' (**9d**) and NPs $_{NP}$*[canın divanı] ve* $_{NP}$*[orhanın yatağın]* 'Can's couch and Orhan's bed' (**9e**), examples of word-part ellipsis must also be recognized as a specific type of syntactic ACCESS violation, in which syntax/semantics accesses word-internal structure.

Example 9

a. [pre and even to some extent post]-war (economies) *English*

b. [write or print]-able

c. [pré e pós]-guerra *Portuguese*
pre and post-war
'pre and post war'

d. [segura mas lenta]-mente
safe but slow-ADV
'surely but slowly (lit. safely but slowly)'

[19] Language data examples original to: Spencer (2005:82) (English); Booij (2009:88), Vigario (2003:251) (Portuguese); Akkuş (2015:2-3), Lewis (1967:35) (Turkish).

e. [can-ın divan-ı ve orhan-ın yatağ-ın]-da uyu-du-m *Turkish*
Can-GEN couch-3.SG and Orhan-GEN bed-3.SG-LOC sleep-PAST-1.SG
'I slept on Can's couch and Orhan's bed'

f. [zengin ve ünlü]-y-dü-m
rich and famous-COP-PAST-1.SG
'I was rich and famous'

Another example of an ACCESS violation in which syntax/semantics 'sees' morphology is provided in **Example 10**. As discussed by Ackema and Neeleman (2004:11)[20], the order of the verb with respect to its object NP in Quechua is relatively free – the verb can precede the object NP, as in **10a**, or follow it, as in **10b**. However, if the VP is nominalized (denoted by the presence of the nominalizer *sqa*), only the object-verb pattern is permitted (cf. the grammatical acceptability of **10c** with that of **10d**). Therefore, these cases are representative of an ACCESS violation where syntactic ordering is dependent on the presence of specific morphological forms (i.e., syntax apparently accesses morphology in order to determine word order).

Example 10

a. xwan mikhu-n papa-ta
Juan eat-3 potato-ACC
'Juan eats potatoes'

b. xwan papa-ta mikhu-n
Juan potato-ACC eat-3
'Juan eats potatoes'

c. [[xwan papa-ta mikhu-n] sqa-n-ta] yacha-ni
Juan potato-ACC eat-3 NMZ-3-ACC know-1
'I know that Juan eats potatoes'

d. *[[xwan mikhu-n papa-ta] sqa-n-ta] yacha-ni
Juan eat-3 potato-ACC NMZ-3-ACC know-1

Certain varieties of Dutch provide justification for a second subtype to ACCESS violations. In contrast to examples where syntax appears to be conditioned by particular morphological forms, Ackema and Neeleman (2004:11) observe that in East Netherlandic varieties of Dutch[21], one inflectional paradigm is used if the verb follows the subject NP (as evidenced by the *-t* plural

[20] Originally noted by Lefebve and Muysken (1988:18).

[21] Originally noted by van Haeringen (1958).

marker in **Example 11a**), and another paradigm is used if the verb precedes its subject (the *-e* plural marker in **11b**). Examples such as these (as well as instances of construction-dependent morphology as discussed by Booij (2005)), support the possible ACCESS subtype 'morphology sees syntax/semantics'.

Example 11

a. wij speul-t
we play-PL
'we play'

b. dan speul-e wij
then play-PL we
'then we play'

Finally, violations may manifest as a matter of unpredicted linguistic forms in relation to the linear order of grammatical components assumed by LI. **Example 12**[22] provides instances of ORDER violations on the word-level; the Sanskrit example in **12a** demonstrates the derivational function of the preverb *pari-*, which, together with the lexicalized verb root *ṇayat* 'he leads', produces the composite verb form *pariṇayat* 'he leads', while the inflected form in **12b** exhibits the temporal prefix *a-* between the derivational preverb *pary-* and the root/stem. This morphemic order of DERIVATION-INFLECTION-ROOT/STEM violates prediction (3) made by LI of a strict INFLECTION-DERIVATION-ROOT morpheme order (in the case of prefixation in **12b**, and ROOT-DERIVATION-INFLECTION with respect to suffixation). The Georgian verb in **12c** demonstrates a similar (unpredicted) morphemic order of DERIVATION-INFLECTION-ROOT/STEM-INFLECTION.

Example 12

a. pari-ṇayat *Sanskrit*
around-he.leads
'he marries'

b. pary-a-ṇayat
around-IMPF-he.leads
'he married'

c. mo-g-ḳlav-s *Georgian*
PREVERB-2.OBJ-kill-3.SUB
'he will kill you'

[22] Language data examples original to Beard (1998:45-46).

ORDER violations may also appear on the phrase-level, as in cases of phrasal derivation (e.g. **Example 4**, **Section 1.2**) and phrasal compounding[23] where syntax appears to function as the input to word formation. For example, Bruening (2018:4) provides the American English advertisement in **Example 13a**, in which a full VP functions as an adjectival modifier within the NP $_{NP}$*[the* $_{VP}$*[prevent-mold-on-your-shower-curtain-for-up-to-seven-days]* $_{N}$*[spray]]*. Botha (1981:76), Tokizaki (2017:2), and Wiese (1996:185) provide similar examples in Afrikaans, Japanese, and Mandarin Chinese, respectively[24].

Example 13

a. Lysol is the [prevent mold on your shower curtain for up to seven days] spray *English*

b. [ons word gruwelik uitgebuit] houding *Afrikaans*
we are terribly exploited attitude
'attitude indicating a feeling of being unduly exploited'

c. [Nippon-no so'ra-o toboo] kyampe'en *Japanese*
Japan-of sky-ACC fly campaign
'let's fly over the Japanese sky campaign'

d. [bǎi huā qí fàng] yùndòng *Mandarin Chinese*
hundred flower simultaneous bloom campaign
'hundred-flowers-simultaneously-blossom movement'

Up to this point, the term 'word' has been used in reference to morphology, syntax, their interface, and various LI violations. Considering the lack of any clear definition or understanding of the notion 'word' (Haspelmath 2011), and given the present goals of this paper, no attempt at defining it will be made. Rather, following Sag (2012:97-98), 'word' will be considered a cover term based

[23] Phrasal compounding has been formally characterized by Meibauer (2007:234) as a word formation process "of the type YP+X, with YP modifying X semantically", e.g. [$_{VP}$[prevent mold on your shower curtain …] + $_{N}$[spray]].

[24] Phrasal compounding is quite cross-linguistically common, and has been discussed in relation to English (e.g., Bauer 1983, Lieber 1992, Carney 2000, Lieber and Scalise 2007, Bruening 2018, among many others); Afrikaans (e.g., Botha 1981, Savini 1983); Dutch (Ackema and Neeleman 2004); Swedish (Mukai 2006); German (Meibauer 2007); Japanese (Tokizaki 2017); Mandarin Chinese (Wiese 1996, Tokizaki and Kuwana 2008); French and Danish (Bauer 1983). Furthermore, phrasal compounding appears to be highly productive within and across languages, as noted by Bauer (1983), Ackema and Neeleman (2004), Meibauer (2007), Trips (2012), and Tokizaki (2017) (citing Kubozono (1995)). Observations such as these contradict Bresnan and Mchombo's (1995) and Wiese's (1996) attempts to characterize phrasal compounding as a result of lexicalization or quotation; it is not clear how a productive (and in many cases novel, as e.g. demonstrated by Trips (2012)) use of a phrasal compound becomes instantly lexicalized for the purposes of the morphology-syntax interface.

on a given linguistic expression's (i.e. sign, in the sense of de Saussure (1983)) privilege of instantiation in a licit syntactic construction[25].

3. Theoretical Survey

3.1. Overview of Linguistic Theories

It has been observed that the breadth of current morphological and syntactic theories makes a comprehensive evaluation of LI challenging (Lieber and Scalise 2007:13), and the high degree of variation among implementations of LI across different grammatical frameworks only compounds this challenge (cf. **Section 1.1**). Moreover, a proper evaluation of LI is complicated by the fact that certain theoretical frameworks, occupying opposing ends of the lexicalist spectrum, are claimed to conform to LI, albeit at different levels of abstraction. For instance, strongly lexicalist theories can be said to conform to LI at the word level, since both derived and inflected word forms in such theories are the products of a separate morphological component, resulting from a distinct set of morphological rules and principles. These fully inflected and derived word forms are inserted into the syntax as is and are therefore opaque to any subsequent (strictly) syntactic operations. In contrast, weakly lexicalist theories are also claimed to conform to LI, but at the lexemic level (Harley 2015:1141), where a separate morphological component produces derivationally complete word forms (i.e. lexemes), which then receive inflection via syntactic and post-syntactic operations. By stipulating that the interaction of inflectional morphology and syntactic combination is necessarily the result of purely syntactic procedures (as per Anderson (1982:587)'s dictum "inflectional morphology is what is relevant to syntax"), one can craft a version of LI in which the interaction between inflectional morphology and syntax is permitted. In this sense, the uninflected word forms (lexemes) adhere to LI, since their internal, derivationally complex

[25] This more general conceptualization of 'word' is contra Booij (2009), who uses LI and word-cohesiveness in order to define the notion of wordhood, stating "I consider the impossibility of syntactic movement of the constituents of a linguistic unit as a necessary condition for that linguistic unit to be a word" (pp. 85), and in particular regard to LI "[the prohibition on movement] part of [LI] may serve as a basic test to find out if a sequence of morphemes is a word or a phrasal lexical unit" (pp. 86). However, Haspelmath (2011:25) cautions against this very approach to defining wordhood: "The really serious problem resulting from the indeterminacy of word segmentation is that linguists often presuppose the word concept and the morphology-syntax division, and even try to use it for explanatory purposes. Some of the contexts in which the word as a cross-linguistic category and/or the syntax/morphology division is presupposed [include] the Lexical Integrity Hypothesis (e.g. Bresnan and Mchombo 1995)".

structure is invisible to syntactic operations, but at the same time they are available for syntactically conditioned inflectional affixation.

Therefore, given the inherent challenges of a complete, cross-theoretical evaluation of LI, several major linguistic theories are compared along three parameters – FRAMEWORK, MODEL OF GRAMMAR, and TYPE OF THEORY – in order to explore the relationship between LI and each major line of linguistic inquiry as comprehensively and precisely as possible, without becoming entrenched in the minutiae of theory-specific implementations of LI. FRAMEWORK is considered to be the mechanism used in describing syntactic (and potentially morphological) combination and semantic composition. The range of linguistic theories can be categorized as operating within either a UNIFICATION-BASED (NON-DERIVATIONAL)[26] or DERIVATIONAL[27] grammatical framework (Sag et al. 1986). In UNIFICATION-BASED frameworks, linguistic objects are modeled declaratively as feature structures[28]. A feature structure specifies (i.e. constrains) the possible combinations of phonological, syntactic, semantic, and contextual information used to characterize a given linguistic object. The building up of phrase-structural configurations, such as the combination of a VP with its subject NP, are modeled by the application of monotonic[29] constraints of equality over various feature structures. These monotonic constraints specify how the information captured in each feature structure will be merged (through the formal operation of unification) with the result being a complete syntactic construction consisting of all (and only) the information present in that syntactic construction's composite parts (Sag et al. 1986:238-243, Pollard and Sag 1987:7-8). Examples of linguistic theories with a UNIFICATION-BASED foundation include Generalized Phrase Structure Grammar (GSPG) (Gazdar et al. 1985), Head-Driven Phrase Structure Grammar (HPSG) (Pollard and Sag 1987, 1994), Lexical Functional Grammar (LFG)

[26] Also referred to as INFORMATION-BASED or CONSTRAINT-BASED grammar formalisms (Pollard and Sag 1987, Shieber 1992), CORRESPONDENCE-TYPE (Harley 2015:1138-1139) and MODEL-THEORETIC approaches (Müller 2016).

[27] Also referred to as GENERATIVE-ENUMERATIVE approaches (Müller 2016).

[28] Feature structures are used to model different kinds of linguistic objects in different theories, and they are called different things: feature structures in HPSG (Pollard and Sag 1994:8), PATR-II (Shieber 1986:10) and SBCG (Sag 2007, Sag 2012, Michaelis 2012, f[unctional]-structures in LFG (Bresnan et al. 2016:44-48), feature categories or matrices in GPSG (Gazdar et al. 1985:20-27), among others.

[29] See Sag et al. (1986:2-4) concerning the nature of monotonic constraints, compared to structure-changing derivational transformations.

(Bresnan et al. 2016[2001]), Construction Grammar (CxG) (Fillmore 1988; Goldberg 1995, 2006), and the Parallel Architecture (PA) (Jackendoff 1997, 2015).

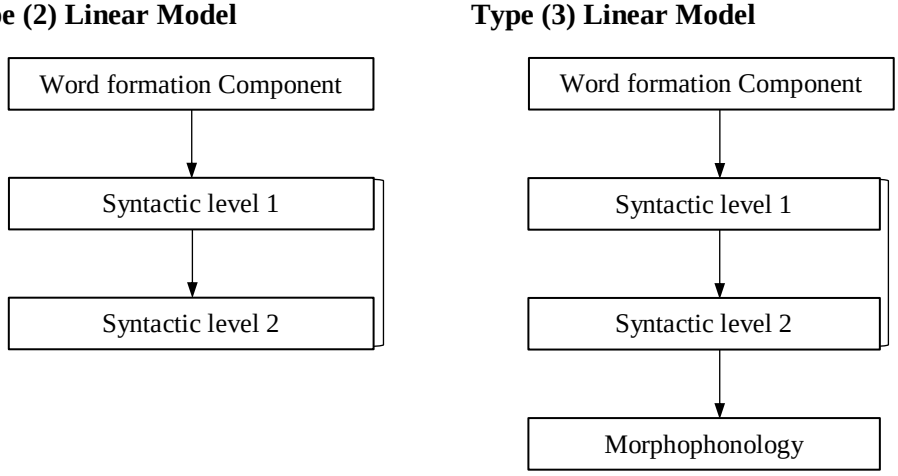
DERIVATIONAL frameworks, by contrast, involve an underlying, fully specified D[eep]-structure, and a derived S[urface]-structure. D-structures are base-generated by the interaction of the lexicon and phrase structure rules, and represent the structural and thematic relations between sentential units. S-structures are derived from their relevant D-structures by the successive application of transformations (e.g. movement, deletion, etc.), according to principles of the grammar (Müller 2016:81-94). For DERIVATIONAL (or TRANSFORMATIONAL) frameworks[30], this derivational relationship between D-structure and S-structure forms the basis of linguistic description and explanation (Sag et al. 1986:2), and for example provides an account of why the English active *Muhammed hugged Li* and passive *Li was hugged by Muhammed* have essentially the same meaning (i.e. they share a common D-structure). Examples of linguistic theories assuming a DERIVATIONAL framework include early Transformational Grammar (Chomsky 1957, 1965), Government and Binding Theory (GB)/The Principles and Parameters Approach (P&P) (Chomsky 1981, 1982), and the Minimalist Program (MP) (Chomsky 1995), as well as morphological theories that are couched within specific DERIVATIONAL frameworks, for example Lieber (1992) (with respect to GB/P&P), and Distributed Morphology (Halle and Marantz 1993) (with respect to Minimalism).

MODEL OF GRAMMAR refers to the overall architecture of the grammar, specifically whether a distinct lexicon, morphological, and syntactic component are assumed, and where word formation occurs with respect to syntax. In general, the various grammatical models are embedded within the two primary DERIVATIONAL and UNIFICATION-BASED linguistic frameworks, with DERIVATIONAL frameworks comprising types of LINEAR MODELS, as well as models which are FULLY SYNTACTIC, and UNIFICATION-BASED frameworks comprising UNIFIED and MODULAR (in correspondence) models. Within many DERIVATIONAL models, LI is enforced by ordering the morphology (comprising a lexicon and a word formation rule component) prior to the syntax,

[30] While D-structure is no longer present in Minimalism (Chomsky 1995), it is still 'derivational' in that sentences are derived by the (external and internal) operation Merge (Müller 2016:124), or specifically, Syntactic Merge (cf. Lieber and Scalise 2007:21).

which "entails that the output of one [component] is the input to the other" (Borer 1998:152-153). This ordering of the morphological and syntactic components is reflected in two types of linear models: TYPE (2) LINEAR MODELS and TYPE (3) LINEAR MODELS (Borer 1998:153). In TYPE (2) LINEAR MODELS, the morphological component, which encompasses both derivational and inflectional processes, precedes D-structure and any syntactic operations, whereas in TYPE (3) LINEAR MODELS, the morphological component is separated from the component that phonologically realizes the morphophonology of words and word-parts. TYPE (3) LINEAR MODELS involve the introduction of categorial feature bundles prior to D-structure, and a separate post-syntactic morphophonological component provides phonology to the feature bundles generated by the lexicon and manipulated by the syntax. TYPE (2) and TYPE (3) LINEAR MODELS are represented below in **Figure 2**, demonstrating the general architecture of the grammar in terms of a series of ordered components.

Figure 2 Linear Models (Borer 1998:153)

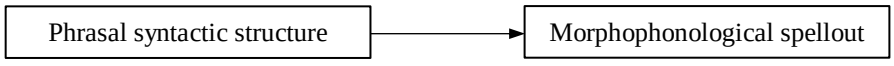


Similar to TYPE (2) and (3) LINEAR MODELS, and still within a DERIVATIONAL framework, are models which are FULLY SYNTACTIC. In contrast to the LINEAR MODELS outlined above, FULLY SYNTACTIC models eliminate a separate morphological component altogether, and instead treat word formation as obeying syntactic constraints and interacting with syntactic rules (Borer 1998:157). FULLY SYNTACTIC models can be further divided based on whether words are incrementally derived in the syntax, or realizationally spelled-out following the syntax. INCREMENTAL theories (Stump 2001:2) assume morphology is "information increasing", in that

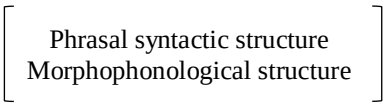
the verb *loves* receives the meanings 'third-person singular subject', 'present tense', and 'indicative mood' by the concatenation of the root *love* and the suffix *-s*. For example, Lieber (1992) presents a FULLY SYNTACTIC INCREMENTAL account of word formation where affixes themselves, as well as roots and stems, are part of the lexicon and contain morphosyntactic features, which are then syntactically combined into complex morphological objects in tandem with phrasal combination (Harley 2015:1145). In contrast, REALIZATIONAL theories (Anderson 1992:12, Stump 2001:1) are so-called in that morpholexical rules and features 'realize' a given word's morphophonological structure and morphosyntactic distribution (Harley 2015:1144-1145). For example, in a REALIZATIONAL account the two English plural forms *cats* and *oxen* are equivalently realized by virtue of the feature [+plural] (i.e. a separate morphophonological component produces the appropriate inflected forms for *cat*[+plural] and *ox*[+plural]) as opposed to an INCREMENTAL account, which would need to specify that the stem *cat* takes the plural *-s* suffix and *ox* the plural *-en* suffix in order to produce the composite plural forms *cats* and *oxen*. FULLY SYNTACTIC REALIZATIONAL theories roughly correspond to what Borer (1998) terms a TYPE (33a) MIXED MODEL (represented below in **Figure 3**), where the word formation component is 'mixed' with the syntax – "the formation of amalgams of functional heads is a non-morphological task [i.e. is fully syntactic], and its output, in turn, feeds into an independent morphophonological component that is syntactically irrelevant" (pp. 180). A FULLY SYNTACTIC INCREMENTAL model would consist solely of a syntactic component that receives stems and roots, as well as affixes, from the lexicon, and assembles both morphologically complex words and syntactic constituents according to well-defined syntactic procedures.

Figure 3 Mixed Models (Borer 1998:180)

Type (33a) Mixed Model



Type (33b) Mixed Model



In contrast to LINEAR and FULLY SYNTACTIC models are architectures that either maintain the modularity of a morphological and syntactic component but put them in correspondence with one another, or provide a unified model of grammar that essentially dissolves the boundaries between a distinct morphological component and syntax completely. Grammatical architectures that maintain the modularity of morphology and syntax, but eliminate the linear ordering of these components, are what Borer (1998:180) refers to as TYPE (33b) MIXED MODELS (**Figure 3**), in which the morphophonological and syntactic components are placed in parallel, or in 'correspondence' (Harley 2015:1138). For instance, in LFG the syntactic component is composed of two structures: a f[unctional]-structure, which handles predicate-argument/grammatical relations, and a c[onstituent]-structure, which handles precedence and dominance relations in terms of classic phrase structure, and which takes fully inflected (and derived) word forms as its terminal nodes (Nordlinger and Sadler 2016:3-6). Morphology is therefore treated as a distinct component that forms the input to syntax (specifically c-structure trees). However, in contrast to LINEAR models, where morphology precedes (i.e. 'feeds') syntax, the parallel relationship between morphology and syntax in LFG has generally been handled in terms of competition between the morphological and syntactic components[31]. UNIFIED models, on the other hand, reject not only a linear relationship between morphology and syntax, but also the view that morphology and syntax comprise distinct modules at all. PA (and Relational Morphology (RM) therein) is one such theory: it sees the lexicon as structured by an inheritance hierarchy and extends this inheritance model to relations among phrasal types – a view that is shared by CxG and Booij's (2007, 2010) Construction Morphology (CM) (Jackendoff and Audring 2016:469). In PA, a word and its possible affix(es) are 'licensed' or 'motivated' by the interface(s) between that word's semantics (the meaning of the word and its part(s)), its morphosyntactic properties (the syntactic functionality of that word and possible affix(es)), and its phonological representation (how that word and potential affix(es) are produced). In this way, morphology is declaratively represented as an interface between phonology, syntax, and semantics, as opposed to being the product of an assembly procedure. Moreover, theories such as PA provide a UNIFIED linguistic model, in which

[31] For example, Andrews (1990:519) "proposes [the Principle of Morphological Blocking] whereby the existence of a more highly specified form in the lexicon [e.g. a synthetic verb exhibiting subject agreement morphology] precludes the use of a less highly specified form [e.g. a corresponding syntactic phrase consisting of an unmarked verb form and a separate subject pronoun]."

language – from phonology and phonotactics to word formation to syntactic composition to semantic interpretation – is characterized as the result of simultaneous constraint satisfaction.

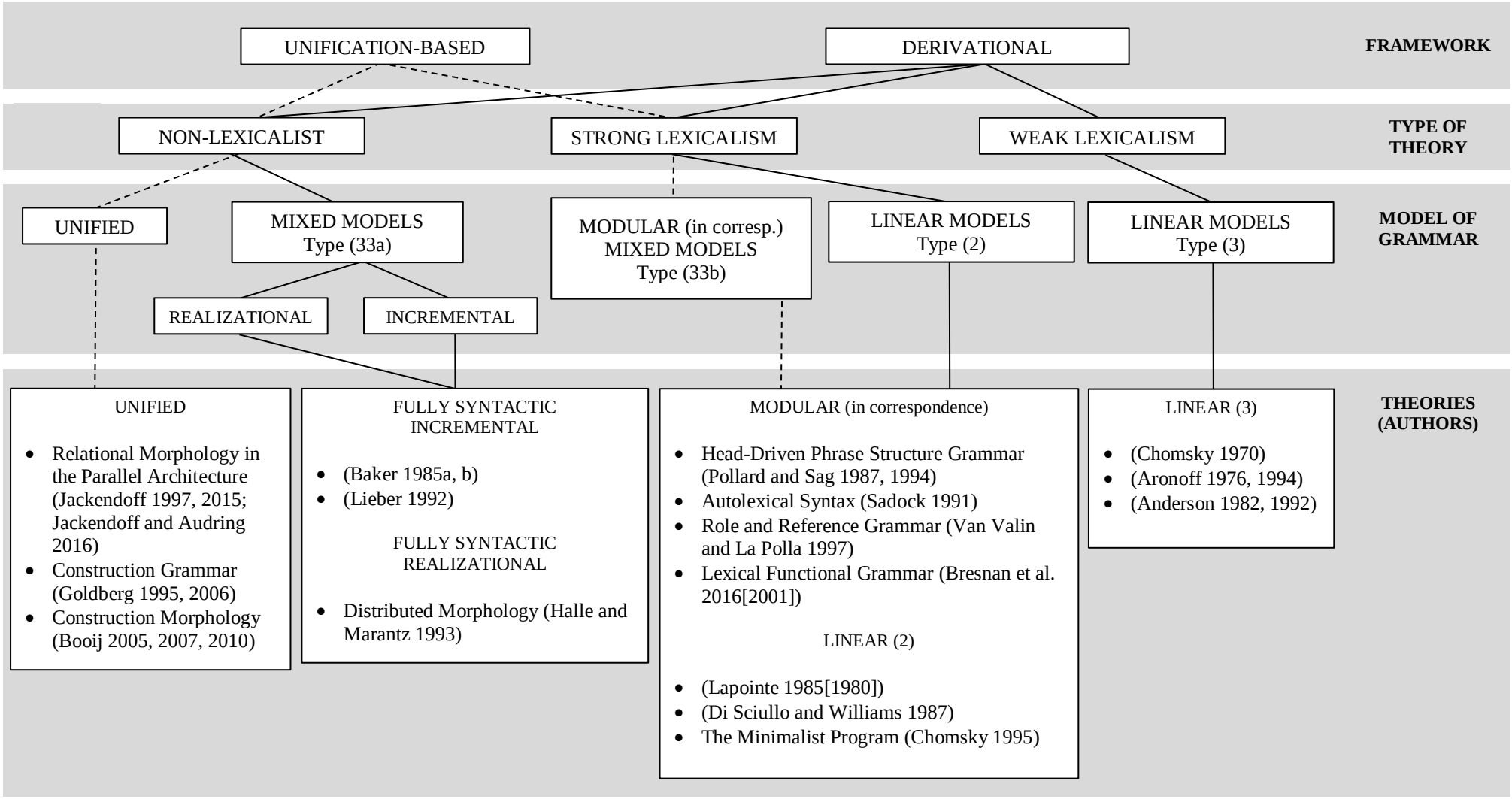
Finally, each theory is organized in terms of its overall 'type'; specifically, whether it is a LEXICALIST or NON-LEXICALIST theory. As outlined in **Section 1.1**, lexicalism (and LI) has its origins in Chomsky's (1970) proposal for a separation between lexical processes (e.g. derivation), comprising the lexical component of grammar, and syntactic processes, comprising the syntactic component. Therefore, in order for a theory to qualify as LEXICALIST, it must assume some degree of modularity (and autonomy) between morphology and syntax (Scalise 1986:20, Spencer 1991:72-73, Beard 1998:44, Toman 1998:306, Harley 2015:1138, Bruening 2018:1). The degree of separation between the morphological and syntactic modules is further represented by the 'strength' of lexicalism assumed by a given theory, or rather which processes (derivational and/or inflectional) are relegated to the lexical component or handled in the syntax. In general, STRONG LEXICALISM considers both derivation and inflection to be the result of a distinct morphological component, while WEAK LEXICALISM splits the two between the morphological and syntactic components, with derivation belonging to the former and inflection to the latter (Scalise 1986:101-102, Spencer 1991:178). In NON-LEXICALIST theories, morphology and syntax are not distinct modules; instead, with regard to SYNTACTICOCENTRIC theories, there is a single syntactic component that assembles both words and phrases, with words and word-parts (or abstract feature bundles, in the case of realizational models) able to occupy terminal nodes in the syntax (Harley 2015:1141). NON-LEXICALIST – UNIFIED theories, on the other hand, declaratively represent both morphological and syntactic processes as constructions (Goldberg 1995; Booij 2007, 2010) or schemas (Jackendoff and Audring 2016), making no distinction between separate grammatical components.

The taxonomy of linguistic theories in **Figure 4** provides specific (groups of) approaches to morphology and syntax in relation to the theoretical assumptions that underlie them. Specifically, each line in **Figure 4** represents a direct relationship between the overall FRAMEWORK (either UNIFICATION-BASED or DERIVATIONAL), and the theory type (NON-LEXICALIST, STRONG-

LEXICALISM, and WEAK-LEXICALISM) and MODEL OF GRAMMAR[32]. Solid lines indicate a relationship between DERIVATIONAL frameworks and other levels of the taxonomy, while dashed lines indicate a relationship to UNIFICATION-BASED frameworks. The relation between each line and level of the taxonomy illustrates that, for example, theories that are generally understood to be distinct (e.g. HPSG and MP) are in fact similar in that both assume a type of STRONG LEXICALISM, but differ with respect to their overall FRAMEWORK and MODEL OF GRAMMAR (UNIFICATION-BASED and TYPE (33b) MIXED MODEL regarding HPSG, DERIVATIONAL and TYPE (2) LINEAR MODEL regarding MP). At the same time, the taxonomy shows that UNIFIED theories (e.g. RM, CxG) are maximally distinct from the work of Chomsky (1970), Aronoff (1976, 1994), and Anderson (1982, 1992).

[32] The relationship between a given theory's type of lexicalism and MODEL OF GRAMMAR is closely linked, and given the state of the art in linguistic theory, that relationship can be considered in two ways: either one's assumption regarding the type of lexicalism will constrain the model of grammar in which one operates, or the MODEL OF GRAMMAR one adopts will influence the type of lexicalism that is assumed. These two levels are kept distinct in the taxonomy in order to demonstrate, among other things, that despite the individual variation between grammatical models and associated theoretical approaches, each is related to a more general type of lexicalism.

Figure 4 Taxonomy of Linguistic Theories



3.2. Theoretical Survey Observations

In order to address both the cross-theoretical nature of LI violations, as well as whether a particular type of violation is correlated with specific theoretical assumptions, individual implementations of each linguistic theory are considered generally with a specific focus instead on how each group of theories, and their inherited theoretical traits, relate to LI and its violations. The proceeding groups of linguistic theories (as outlined above in **Section 2**, and provided toward the bottom of **Figure 4**) and LI violation types (**Section 3.1**, **Figure 1**) are compared:

Groups of Linguistic Theories	**LI Violation Types**	
UNIFIED	1. MANIPULATION	a. *movement* b. *word-part ellipsis* †
FULL SYNTACTIC INCREMENTAL		
FULL SYNTACTIC REALIZATIONAL	2. ACCESS ‡	a. *syntax/semantics 'sees' morphology* †, § b. *morphology 'sees' syntax/semantics* §
MODULAR (in correspondence)		
LINEAR (2)	3. ORDER	a. *word-level* b. *phrase-level* ‡
LINEAR (3)		

LI violations are entirely theoretically dependent, and as a result, exhibit no true cross-theoretical validity. However, it will be shown that specific types of violations are directly correlated with particular parameters of the linguistic theory one adopts. **Table 1** provides a general comparison of each theoretical group (indicated in bold), in relation to the LI violation types (and subtypes) that the linguistic phenomena listed in **Figure 1**, **Section 2** present for those theoretical approaches[33]. By considering each theory group in terms of its FRAMEWORK (UNIFICATION-BASED or DERIVATIONAL), lexicalist type (STRONG LEXICALISM, WEAK LEXICALISM, or NON-LEXICALIST), and MODEL OF GRAMMAR (TYPE (2) LINEAR MODELS, FULLY SYNTACTIC REALIZATIONAL, TYPE (33a) MIXED MODELS, etc.), in relation to each LI violation type, clear interrelationships emerge between specific properties each group of theories holds and the LI violations with which they contend.

[33] There is a high degree of variability regarding each individual's work. Accordingly, individual linguistic phenomena (within a given violation type) that may constitute a violation to one linguist may not to another, even when operating within the same theory. To reiterate, the present observations are therefore intended in a general manner, in order to illustrate the various interrelationships among LI and linguistic theory.

Table 1 Theoretical Parameters and LI Violation Types

FRAMEWORK – TYPE – **MODEL/THEORY GROUP**	**LI Violation Types**					
	1a	1b	2a	2b	3a	3b
UNIFICATION – NON-LEXICALIST – **UNIFIED**	(CM)					
DERIVATIONAL – NON-LEXICALIST – **FULLY SYNTACTIC**						
UNIFICATION – STRONG – **MODULAR (in correspondence)**	?	?	?	?	?	?
DERIVATIONAL – STRONG – **LINEAR MODELS (2)**	X	X	X	X		X
DERIVATIONAL – WEAK – **LINEAR MODELS (3)**	?	?	?	?	X	X

LI Violation Type Abbreviations

1a	MANIPULATION (movement)
1b	MANIPULATION (word-part ellipsis) †
2a	ACCESS (syntax/semantics 'sees' morphology) †, ‡, §
2b	ACCESS (morphology 'sees' syntax and semantics) ‡, §
3a	ORDER (word-level)
3b	ORDER (phrase-level) ‡

ORDER violations, and both its subtypes, are specifically correlated with a given theory's grammatical model (indicated in dark grey shading and white text in **Table 1**). TYPE (2) and (3) LINEAR MODELS, which position a word formation component as separate from and preceding syntax, are problematic for phrase-level LI violations since examples such as phrasal compounding, phrasal derivation, etc., contradict this ordering relationship. However, word-level ORDER violations are more closely associated with TYPE (3) as opposed to TYPE (2) LINEAR MODELS; given that TYPE (3) LINEAR MODELS integrate inflectional processes into the syntax, and word formation (i.e. derivation) precedes syntax, then instances in which the reverse order is observed violate this parameter. This is not the case for TYPE (2) LINEAR MODELS, since both derivational and inflectional affixation, independent of order, may be handled by separate word formation rules prior to syntax. Despite the ability of TYPE (2) LINEAR MODELS to treat apparent word-level LI violations, the fact that LINEAR (2) type theories assume both a TYPE (2) LINEAR MODEL and (forms of) STRONG LEXICALISM present several potential issues when considered in relation to the other violation types. Given the firewall imposed by STRONG LEXICALISM, in connection with the modular, linear architecture assumed by a TYPE (2) LINEAR MODEL (indicated within a solid black rectangle in **Table 1**), both MANIPULATION and ACCESS violation types (and

their subtypes) should not be possible; particularly, syntactic operations should not MANIPULATE word-internal elements (whether via movement or word-part ellipsis) by virtue of the ordered, procedural nature of word formation with respect to syntax. Moreover, as a result of the barrier enforced by STRONG LEXICALISM, in coordination with the architecture of a TYPE (2) LINEAR MODEL, ACCESS of morphological structure by the syntax (and vice versa) (e.g. anaphoric islands, focus targeting sub-lexical units, construction dependent morphology, etc.) should be equally impossible. Therefore, LI violations present the most serious, inherent problems for LINEAR (2) type theories (e.g. Di Sciullo and Williams 1987, Chomsky 1995), in that nearly every violation type is not fundamentally amenable to these theoretical approaches.

The cross-theoretical permeation of the lexicalist spectrum (i.e. variable implementations of STRONG and WEAK LEXICALISM, indicated in light grey shading in **Table 1**) appears to be predominantly responsible for the broad potential for LI violations. These violations do not appear to necessarily correlate with the FRAMEWORK or MODEL OF GRAMMAR individually, but rather depend more on the linguist's attitude toward LEXICALISM in relation to either parameter. For example, STRONGLY LEXICALIST – UNIFICATION-BASED theories maintain distinct morphological and syntactic components, however, given one's opinion of what constitutes morphology and syntax proper, various linguistic phenomena can be principally integrated into one component over the other. And, considering the UNIFICATION-BASED nature of these STRONGLY LEXICALIST approaches, and the fact that the related grammatical model (TYPE (33b) MIXED MODEL) may position the morphological and syntactic components in a non-linear fashion (e.g. parallel, or in correspondence), what are then defined as distinct morphological and syntactic processes can interface more openly through unification and constraint satisfaction. Accordingly, MANIPULATION, ACCESS, and ORDER violations might indeed constitute violations, if one assumes a maximally strong conceptualizations of LI, morphology, syntax, as well as a TYPE (33b) MIXED MODEL as in LFG (e.g. Bresnan and Mchombo 1995)[34]; or they might not constitute violations, if certain linguistic phenomena are relegated to the lexicon/morphology and/or syntax (e.g. lexicalization of phrasal/derivational compounds, cliticization, suspended affixation, noun

[34] While STRONGLY LEXICALIST – UNIFICATION-BASED theories do not assume the procedural derivation of TYPE (2) and (3) LINEAR MODELS, the fact that "the interaction [between morphology and syntax] is still only one way" (Bruening 2018:2, f.n. 3) in such UNIFICATION-BASED approaches results in the potential for each LI violation types.

incorporation, etc., i.e. potential types of ORDER and MANIPULATION violations), which can then interface (i.e. ACCESS) more directly via unification. Likewise, MANIPULATION and ACCESS violations are variably problematic for WEAKLY-LEXICALIST – DERIVATIONAL theories for two reasons: first, certain processes may be deemed a morphological as opposed to syntactic procedure; and second, depending on one's treatment of the pre-syntactic word formation and post-syntactic morphophonological components in the corresponding TYPE (3) LINEAR MODEL, ACCESS and MANIPULATION violations may be avoided by introducing abstract feature bundles in word formation, which are syntactically manipulated and accessed prior to morphophonological realization. LINEAR (3) type theories (e.g. Chomsky (1970), Anderson (1992), Aronoff (1994), etc.) are thus faced with serious challenges regarding ORDER (phrase-level) violations, due primarily to the particular MODEL OF GRAMMAR, while MANIPULATION and ACCESS violations may or may not be problematic, depending on theory-specific treatments of LI, morphology and syntax, and the grammatical architecture of TYPE (3) LINEAR MODELS. Similarly, MODULAR (in correspondence) type theories (e.g. LFG, HPSG, etc.) may be faced with MANIPULATION and ACCESS violations, due to the modularity of the grammatical model and the strength of LI assumed; however, such linguistic approaches may also accommodate each LI violation type, since (as noted) certain phenomena can be nested in either the morphological or syntactic component, and unification provides a higher degree of intermodular interaction.

Finally, those theories which essentially eliminate the LEXICALISM parameter (indicated within a dashed black rectangle in **Table 1**) are faced with little challenge when evaluated in terms of LI. Since each LI violation type involves some sort of morphosyntactic interaction, whether structurally or semantically, it follows that by removing the received barrier between each component, violations of the morphology-syntax interface turn out to be non-violations. In the case of UNIFIED theories (e.g. CM[35], CxG, RM and PA, etc.), morphological and syntactic processes are declaratively represented in terms of constructions, or schemas, and relate and interface with

[35] Apparent examples of MANIPULATION would remain violations in CM, since Booij (2009) leverages that point of LI in defining the notion 'word' (as noted in **Section 2**, f.n. 25), and further argues that "We need the prohibition on the movement of word constituents for explaining why in Dutch and German the rule of Verb Second that places finite forms of verbs in second position in root clauses cannot strand the prefix of a complex verb […], whereas the particle in particle verbs […] can be stranded" (pp. 86). However, given that CM assumes the framework of CxG (Booij 2013), which declaratively represents distinct syntactic patterns, it is not clear why Booij maintains a prohibition on syntactic movement in his analysis of the Dutch and German Verb Second pattern and 'stranding'.

one another via inheritance and relation. In slight contrast to UNIFIED approaches, which largely reject a formal distinction between morphology and syntax in particular, and words and rules (i.e. grammar) more generally (Goldberg 1995, Jackendoff and Audring 2016), FULLY SYNTACTIC (INCREMENTAL and REALIZATIONAL) theories (e.g. Lieber 1992, DM, etc.) treat morphology *as* syntax, with syntactic rules responsible for both word and sentence formation. However, what UNIFIED and FULLY SYNTACTIC types of theories share is the elimination of a traditional morphology-syntax distinction, and consequently, the elimination of violations of that interface.

The present state of LI, as a culmination of its near fifty years of development and further redevelopment, is therefore manifest in at least two distinct theoretical parameters assumed within linguistic theory, and to varying extents: the degree and form of LEXICALISM imposed, and the architecture assumed by the MODEL OF GRAMMAR[36]. Also of note is that the FRAMEWORK (i.e. the general method used in describing the combination of morphological and syntactic units and their semantic composition) appears to bear little direct relation on the notion of LI and its violations, suggesting that either a UNIFICATION-BASED or a DERIVATIONAL approach can be equally descriptive, and possibly explanatory, depending on specific conceptualizations of LEXICALISM and the MODEL OF GRAMMAR.

Reflecting upon the present state of LI (**Section 1.2**), the empirical data and problems they pose to LI (i.e. violations, outlined in **Section 2**), and the theoretical sources of those problems, one straightforward conclusion at this point might be to remove the boundary between morphology and syntax, as in UNIFIED and FULLY SYNTACTIC types of theories, thereby eliminating both LEXICALISM as well as linear and modular architectures grammar. The question then becomes one of how best to treat morphology and syntax within a unified grammar (if conceived of separately), whether morphology and syntax form a single (unified) morphosyntactic system, or if morphology is truly just an extension of a single syntactic module. Any potential answer(s) to these questions

[36] In some cases, it is likely that the MODEL OF GRAMMAR influenced the formulation of the type of LEXICALISM (e.g. Chomsky 1970), while in others, it is likely that LEXICALISM influenced the development, or refinement, of the MODEL OF GRAMMAR (e.g. Aronoff 1976, Anderson 1982). However, seeing that LEXICALISM is often taken as an implicit assumption that is inexorably interconnected to specific MODELS OF GRAMMAR (e.g. (TYPE (2) and (3) LINEAR MODELS, MODULAR (in correspondence)), determining the precise relationship between a theorist's grammatical architecture and lexicalist assumptions (especially those theories that have been influenced by, but developed in reaction to, Transformational Grammar (Chomsky 1957, 1965, 1981, 1982) (such as GPSG, HPSG, LFG, etc.) is a quixotic enterprise.

will have profound ramifications for the descriptive and explanatory goals of the theory; these possibilities are briefly discussed in **Section 5**.

4. Typological Survey

4.1. Approaches to Morphological Typology

Each language exhibiting at least one LI violation (as briefly discussed in **Section 2**, and listed in **Appendix 7.2**) was surveyed in terms of two approaches to morphological classification: (1) the traditional morphological types AGGLUTINATIVE, FUSIONAL, and ISOLATING (Comrie 1989[1981], Sapir 1921, etc.), and (2) degree of semantic density (or inflectional synthesis) and type of phonological fusion (Bickel and Nichols 2007, 2013a, 2013b). The latter approach is considered in addition to the former traditional scale since, as Bickel and Nichols (2013a, para. 1) observe, "such a scale conflates many different typological variables and incorrectly assumes that these parameters covary universally". Therefore, both methods are employed in the event that a particular typological property, as potentially conflated by the traditional scale in (1), is relevant in relation to the various LI violation types.

Traditional approaches to classifying the morphology of a given language include three prototypical categories (AGGLUTINATIVE, FUSIONAL, and ISOLATING), each of which are defined along two dimensions: degree of synthesis (i.e. the degree to which a language deviates from an ideal analytic type, or the amount of segmentable affixation a given language demonstrates), and degree of fusion (i.e. the degree to which a language deviates from prototypical agglutination, or the amount of grammatical information that is denoted by single, non-segmentable morphemes) (Comrie 1989:42-49). AGGLUTINATIVE languages are characterized by higher degrees of synthesis and low degrees of fusion, as demonstrated below in **Example 14a** and **14b**; both the English verb *destabilize* and Turkish noun *adamları* are single words exhibiting higher degrees of synthesis (each is composed of three morphemes), and each morpheme and associated semantic category is easily segmentable (i.e. a low degree of fusion). On the other hand, FUSIONAL languages are defined in terms of (potentially) higher degrees of synthesis and higher degrees of fusion, as shown in **Example 14c** and **14d**. While the Latin nouns *tempora* and *mores* in **14d** are slightly synthetic

(comprised of two morphemes), the suffixes *-a* and *-es* simultaneously encode the categories 'plural number', 'accusative case', and 'neuter gender'. In contrast with the AGGLUTINATIVE, one-to-one relationship (and straightforward segmentation) of morphemes and semantic categories observed in Turkish (**Example 14b**, e.g. the individual expression of case and number), Latin exhibits FUSIONAL traits in that it deviates from such prototypical forms of agglutination. Similar FUSIONAL tendencies are present in the English verbal agreement ending *-s*, indicated in **Example 14c**. Finally, as opposed to the synthetic examples in **14a** through **14d**, the English and Vietnamese data in **Example 14e** and **14f** represent canonical analytic forms in that each word form is monomorphemic (i.e. a low degree of synthesis) and associated with one semantic concept (i.e. a low degree of fusion), indicative of ISOLATING morphology.

Example 14[37]

a. de-stabil-ize
   REV-stable-VZR

b. adam-lar-ı *Turkish*
   man-PL-ACC
   'men'

c. the dog bark-s
   DET dog bark-3.SG.PRES

d. o tempor-a o mor-es *Classical Latin*
   oh time-NOM.PL.NEUT oh custom-NOM.PL.NEUT
   'oh what times, oh what customs'

e. the dog will sit in the park
   DET dog FUT sit LOC DET park

f. khi tôi dên nhà bạn tôi chúng tôi bắt dāu làm bài *Vietnamese*
   when I come house friend I PL I begin do lesson
   'when I came to my friend's house, we began to do lessons'

Bickel and Nichols (2007) propose alternative means for describing morphological typology, whereby languages are not characterized along the traditional indices of synthesis and fusion, but rather according to typological parameters such as phonological fusion and semantic density[38].

[37] Non-English data examples original to: Comrie (1989:43-44) (Vietnamese and Turkish); Marcus Tullius Cicero (Orationes in Catilinam 1.1) (Classical Latin).

[38] Bickel and Nichols (2007:17) identify a third parameter – flexivity – which refers to the degree of variance, or allomorphy, of a particular grammatical marker in a specific language. For example, Latin demonstrates a higher

Phonological fusion is defined as the degree to which grammatical markers[39] are phonologically bound to their host (pp. 13), and can manifest in terms of one of the following three types:

1. *Isolating* – grammatical markers are single, free phonological words (isolating in this sense is therefore very similar to its traditional construal, e.g. **Example 14e** and **14f** above).
2. *Concatenative* – grammatical markers are phonologically bound to their host word (under the present definition, concatenative includes both the AGGLUTINATIVE and FUSIONAL traditional morphological types, e.g. **Example 14a** through **14d** above).
3. *Non-concatenative* – grammatical markers are added by means of modifying the host word, as opposed to the addition of segmentable, linear sequences of linguistic units[40]. Since this survey failed to identify languages exhibiting both an LI violation and productive non-concatenative morphology, this specific type of phonological fusion is not considered in the overall typological survey.

In addition to a language's phonological fusion type, Bickel and Nichols (2007:21) identify two typological characteristics with respect to semantic density: exponence (density at the level of individual grammatical markers) and synthesis (density at the level word-level). Exponence is a measure of the number of individual grammatical categories expressed by a single grammatical marker, which provides a more quantitative evaluation of the traditional notion of fusion, e.g. **Example 14c** and **14d** above. Synthesis, which is also related to the traditional index of synthesis outlined above, measures "the number of [grammatical markers] and lexical roots that are bound together in one word" (pp. 22), ranging from ANALYTIC – SYNTHETIC – POLYSYNTHETIC. Given that the closely related traditional scale of ISOLATING – AGGLUTINATIVE – FUSIONAL is presently considered, under which polysynthesis may be considered a specific form of agglutination (or compounding), only exponence is reflected in the measure of semantic density. Furthermore, while semantic density is a measure of what is commonly considered to be inflectional grammatical

degree of flexivity with respect to nominative case marking, as evidenced in the two Latin nominative endings *-a* and *-es* in **Example 14d** above. Since the present survey is concerned with the structural interaction of full morphological forms and syntax, and apparent syntactic/semantic access of specific morphological elements (i.e. the morphology-syntax interface), instances of phonological and paradigmatic variation (flexivity) are not considered.

[39] Bickel and Nichols (2007:4) distinguish FORMATIVES (which generally correspond to bound morphemes) and WORDS (i.e. phonologically independent linguistic units). However, the more general term 'grammatical marker' will be used in place of FORMATIVE, in an attempt to avoid introducing additional terminology.

[40] An example of productive, non-concatenative morphology includes the templatic (or root-and-template) morphology found among the Semitic (and certain other Afro-Asiatic) languages.

categories, and LI violations involve both so-called inflectional and derivational properties[41], it is included in the survey simply as a demonstration of how much semantic information a given word may morphologically or lexically encode for a specific language.

Utilizing information provided in Bickel and Nichols (2013a, 2013b), in conjunction with information from individual reference grammars (sources provided in **Appendix 7.4**), the traditional morphological type(s)[42], average semantic density, and phonological fusion type(s)[43] was determined for each language[44] exhibiting at least one LI violation (**Appendix 7.2**). Each language and corresponding typological classifications are provided in **Table 2**, in approximate ascending order of least to most synthetic.

[41] This distinction assumes the Split Morphology Hypothesis (SMH) (Anderson 1982) (briefly noted in **Section 1.1**, f.n. 5). See Booij (1993) for an argument against the SMH.

[42] In situations where a language demonstrates multiple (traditional) morphological types (e.g. English, as evidenced in **Example 14a**, **14c**, and **14e** above), each individual type is provided. Future work would address matters such as: whether specific typological tendencies are associated with specific lexical domains (e.g. verbal, nominal, etc.), whether the typological variation is generally derivational or inflectional in nature (assuming the SMH), explore the potential relationship between the aforementioned considerations and LI violation types (and specific manifestations of each type), etc.

[43] Average semantic density was calculated by first determining the individual inflectional semantic density of nouns and verbs in a particular language, and then deriving the average between the two measurements. Similarly, the overall phonological fusion type was identified in relation to the individual classifications of nouns and verbs. In the majority of cases, the fusion type between nouns and verbs matched, providing a straightforward categorization. In cases where verbal morphology (e.g. concatenative) is distinct from nominal morphology (e.g. isolating), both are noted as the overall type. **Appendix 7.3.1** provides a list of each language under consideration, in relation to their traditional morphological type(s), individual and average semantic density measurements, phonological fusion type(s), and genetic classification.

[44] In a very limited number of cases, information for a genetically related language (typically within the same genus) if information for a particular language exhibiting an LI violation was unable to be located.

Table 2 Semantic Density and Morphological Types of Languages Exhibiting LI Violations

Language	Traditional Morphological Type(s)	Phonological Fusion Type(s)	Average Semantic Density
Mandarin Chinese	I	IS/C	1
Tagalog	A/I	C – IS	3
Warlpiri	A	C	3
Japanese	A	C – IS/C	5
Afrikaans	A/F/I	C	5
English	A/F/I	C	5
Swedish	A/F/I	C	5
Dutch	A/F/I	C	7
German	A/F/I	C	7
Hungarian	A/F	C	7
Italian	F	C	7
Portuguese	F	C	7
Spanish	F	C	7
Sanskrit	A/F	C	7
Udi	A/F	C	7
Korean	A	C – IS/C	7
Tamil	A	C	7
Estonian	A/F	C	9
Georgian	A	C	11
Mohawk (Akwesasne)	A	C	11
Turkish	A	C	11
Quechua (Huallaga)	A	C	13
Walapai	A	C	15

Traditional Typological Abbreviations	
A	AGGLUTINATIVE
F	FUSIONAL
I	ISOLATING

Phonological Fusion Type Abbreviations	
C	CONCATENATIVE
IS	ISOLATING

4.2. Typological Survey Observations

Adopting a similar methodology as that employed in the preceding theoretical survey (**Section 3.2**), specific manifestations of each LI violation are considered in terms of their general type(s) (MANIPULATION, ACCESS, and ORDER) along with their subtypes. In order to explore the cross-linguistic nature of LI and identify any potential correlations between individual violation types and specific morphological traits, the typological information provided in **Table 2** immediately above (and **Appendix 7.3.1**) is organized in relation to each LI violation (sub)type, presented

below in **Table 3** through **Table 5**[45]. Phrase-level ORDER violations are the most cross-typologically present, occurring among all morphological types, while MANIPULATION violations (and its subtypes), which tend to occur among more synthetic languages, also manifest across a broad typological spectrum. For example, **Table 3** and **Table 4** show that phrase-level ORDER violations are attested among all typological profiles, whether characterized as AGGLUTINATIVE (A), FUSIONAL (F), and/or ISOLATING (I), or concatenative (C) and/or isolating (IS) phonological fusion.

Table 3 Traditional Morphological Typology and LI Violation Types

Traditional Type(s)	**1. Manipulation**		**2. Access**‡		**3. Order**	
	Movement	Word-part ellipsis†	Syntax/semantics 'sees' morphology†,§	Morphology 'sees' syntax/semantics§	Word-level	Phrase-level‡
I	x					x
A/I						x
A/F/I		x	x	x		x
A/F	x				x	x
F	x	x	x			x
A	x	x	x		x	x

Table 4 Overall Type of Phonological Fusion and LI Violation Types

Fusion Type(s)	**1. Manipulation**		**2. Access**‡		**3. Order**	
	Movement	Word-part ellipsis†	Syntax/semantics 'sees' morphology†,§	Morphology 'sees' syntax/semantics§	Word-level	Phrase-level‡
IS/C	x					x
C – IS						x
C – IS/C		x	x			x
C	x	x	x	x	x	x

[45] **Table 3** through **5**, in conjunction with **Appendix 7.3.2** through **7.3.5**, were consulted in order to identify the present observations. The more fine-grained analyses in each case (e.g. **7.3.2** through **7.3.5**) support the observed patterns in the calculated averages (**Table 3** through **5**); therefore, only the average results are included in the present discussion. Furthermore, frequency of individual violation types is not taken into consideration since (i) the information was gathered from languages demonstrating LI violations in the first place, and (ii) identifying clearly delineated classes of violation types is largely arbitrary (**Section 1.2** and **2**). Given the small sample size (only twenty three individual languages are considered), it could be the case that the present observations are based on typological profiles that are disproportionately represented in the linguistic literature. Future work will ideally address such concerns.

Furthermore, phrase-level violations are observed among all languages with low to relatively high measures of semantic density (**Table 5**), which also tends to correlate with the aforementioned synthetic morphological types, but not necessarily so[46].

Table 5 Average Degree of Semantic Density and LI Violation Types

Average Semantic Density	**1. Manipulation**		**2. Access**[‡]		**3. Order**	
	Movement	Word-part ellipsis[†]	Syntax/semantics 'sees' morphology[†,§]	Morphology 'sees' syntax/semantics[§]	Word-level	Phrase-level[‡]
1	x					x
3						x
5		x	x			x
7	x	x	x	x	x	x
9	x					x
11	x	x	x		x	x
13			x			
15			x			

MANIPULATION violations instead seem to associate more with synthetic properties (e.g. AGGLUTINATIVE/FUSIONAL and concatenative tendencies, cf. **Table 3** and **4**), and specifically mid-degrees of semantic density (cf. **Table 5**). The data point in the upper left corner comes from an apparent LI violation in Mandarin Chinese (Huang 1984:64) concerning the separability of verb-object compounds. Since compounding can be considered a type of word formation, and compounds often function syntactically, semantically, and phonologically as independent words (Fabb 1998, Bauer 2009), the separability of these verb-object compounds could be conceived as a type of MANIPULATION (movement) violation. Adopting a diachronic perspective, compounding can also be viewed as a grammaticalization process resulting in derivational morphology (Hopper and Traugott 2003:40), in which case this particular example of compounding in Mandarin Chinese could be viewed as a slight shift toward synthetic properties, supporting a general connection between MANIPULATION violations and synthetic/concatenative properties. ACCESS violations (namely situations in which syntax/semantics 'sees' morphology) are similarly observed among the more synthetic typological systems (**Table 3** and **Table 4**), however, in slight contrast

[46] For example, the English possessive pronoun *their* is monomorphemic (a low degree of synthesis, i.e. ISOLATING), but represents a higher degree of fusion, in that it singly expresses the semantic concepts 'possession', 'person', and 'number' (a semantic density measure of at least three).

with MANIPULATION violations, ACCESS violations tend to correlate with mid-to-high degrees of semantic density (**Table 5**).

Considering the general patterns above, specific interrelationships between each LI violation type and individual typological parameter provide the most intuitive connections between a given language's typological profile, apparent violations of LI, and the nature of the morphology-syntax interface. Languages with low degrees of semantic density per grammatical marker and predominantly analytic (i.e. ISOLATING) morphology tend to exhibit syntactic methods for grammatical expression (e.g. word order and grammatical relations) and word formation (e.g. compounding). Since in some cases individual linguistic units may serve as function or content words in the syntax, or as derivational components in word formation (as in the case of Mandarin Chinese verb-object compounds noted above, English verb-particle constructions, and so on), it therefore stands to reason that languages with significant ISOLATING tendencies may appear to contain MANIPULATION violations (cf. the top row of **Table 3**, **4**, and **5**), where loosely bound linguistic forms such as clitics, preverbs, particles, etc., and potentially affixes, might appear to be syntactically manipulated.

Relatedly, languages with more analytic tendencies also exhibit phrase-level ORDER violations, where syntax appears to serve as the input to morphology. Given such languages tend to rely on syntax (construed as the 'input' in the ordering relationship) for grammatical expression to a significantly greater degree than synthetic languages, it follows that the predominant system for grammatical expression would be extended into the traditional domain of word formation (e.g. productive phrasal compounding). Furthermore, since all languages exhibiting a phrase-level ORDER violation also demonstrate at least some degree of concatenative tendencies (cf. **Table 4**), situations of productive inflectional and derivational morphology appearing on phrases could be the result of specific morphological properties interacting with the fundamental role of syntax (analytic means of linguistic expression) cross-linguistically. Other cases of apparent phrase-level ORDER violations, such as words zero-derived from phrases, words derived from constructions, and potentially non-productive morphology on phrases, may simply be the result of lexicalization and grammaticalization processes, whereby phrases become words, and historically independent linguistic units become affix-like.

Another noteworthy correlation is that between higher degrees of semantic density (**Table 5**), languages exhibiting synthetic morphology (the presence of concatenative, AGGLUTINATIVE, FUSIONAL traits (**Table 3** and **4**)), and ACCESS violations, where syntax/semantics 'sees' morphology. In contrast with prototypical analytic types, languages with higher degrees of semantic density per grammatical marker, and more synthetic typological profiles, tend to utilize morphological (e.g. affixation) and lexical means (e.g. suppletion) for expressing grammatical content. Considering that a highly semantically dense word form simultaneously expresses multiple semantic concepts, all of which may bear relevance to syntax, syntax and semantics would accordingly require ACCESS to specific (word-internal) information.

Mixed traditional typological profiles (middle of **Table 3**), exclusively concatenative phonological fusion (**Table 4**), and mid-ranges of semantic density (notably, an approximate degree of seven) (**Table 5**), correlate with the widest spectrum of LI violation types. While the measure of semantic density likely underscores each MANIPULATION, ACCESS, and ORDER violation for reasons noted above, each can be further associated with specific typological trends present in the mixed profiles. Certain MANIPULATION and ACCESS violations seem related to the presence of concatenative morphological processes and some sort of AGGLUTINATIVE and/or FUSIONAL (synthetic) properties, in coordination with mid-degrees of semantic density. Since such languages tend to form words in terms of linear sequences of bound morphemes, and each morpheme and/or word might encode multiple semantic concepts, apparent instances of syntactic MANIPULATION (e.g. word-part ellipsis) and ACCESS (e.g. syntax/semantics 'sees' morphology, and vice versa) may be observed. ORDER violations are correlated with specific analytic and synthetic tendencies within each (mixed) typological profile, with word-level ORDER violations most closely related to strictly synthetic word-formation (**Table 3** and **4**). In these cases, since grammatical markers typically manifest in terms of segmentable, linear sequences which are oftentimes bound to their host, then morpheme orders that contradict the received predictions of TYPE (2) LINEAR MODELS can be taken as ORDER violations on the word-level. On the other hand, phrase-level violations among these

languages likely manifest from an interplay of language-specific synthetic properties with co-present analytic traits, as discussed above[47].

Specific LI violations are thus connected to particular, interrelated typological properties within and across languages. Furthermore, it is likely that the various violations observed across each language and associated typological profile are a result of that language's particular state of diachronic change, and the role of morphology, syntax, and semantics therein. It goes without saying that languages are not simply static systems but undergo diachronic change (e.g. de Saussure's (1983) distinction between synchrony and diachrony, various applications of the comparative method (e.g. Baldi 1990), and earlier philological scholarship). Moreover, it has been observed that languages are in a constant "cycle of change" (Dixon 1994:182) between various typological profiles, and the complications these typological profiles (and attested violations) present for the morphology-syntax interface and LI demonstrate this. **Figure 5** below represents each violation type, along with its correlated traditional typological profile(s), in terms of Dixon's cycle of language change (pp. 183)[48].

[47] This could also be related to Jackendoff's observations on linear concatenation of symbols (e.g. 2002:246-252), compounding (2010:423), and hierarchy of grammars (2014), in order to explain language acquisition, processing, emergence, evolution, etc.

[48] Dixon (1994:182-183, f.n. 1) notes that an earlier proposal of diachronic change (e.g. Jesperson 1922:421-425) argues for a unidirectional progression of synthetic to analytic, in contrast with the cyclic conceptualization presented above. However, it is the impression of the author that even this cycle is too narrow in terms of unidirectional, albeit cyclic change, in that it assumes a categorical progression of FUSIONAL to ISOLATING, ISOLATING to AGGLUTINATIVE, AGGLUTINATIVE to FUSIONAL, ad infinitum. Such a delineation between stages of diachrony obfuscates the relationship between individual typological parameters (particularly those typically assumed to belong to one morphological type over another), and the degree and type of morphosyntactic interaction observed across languages (given one's theoretical assumptions). This is particularly relevant in respect to languages with highly mixed typological profiles, and which exhibit a wide spectrum of LI violation types.

Figure 5 The Cycle of Change and LI Violation Types

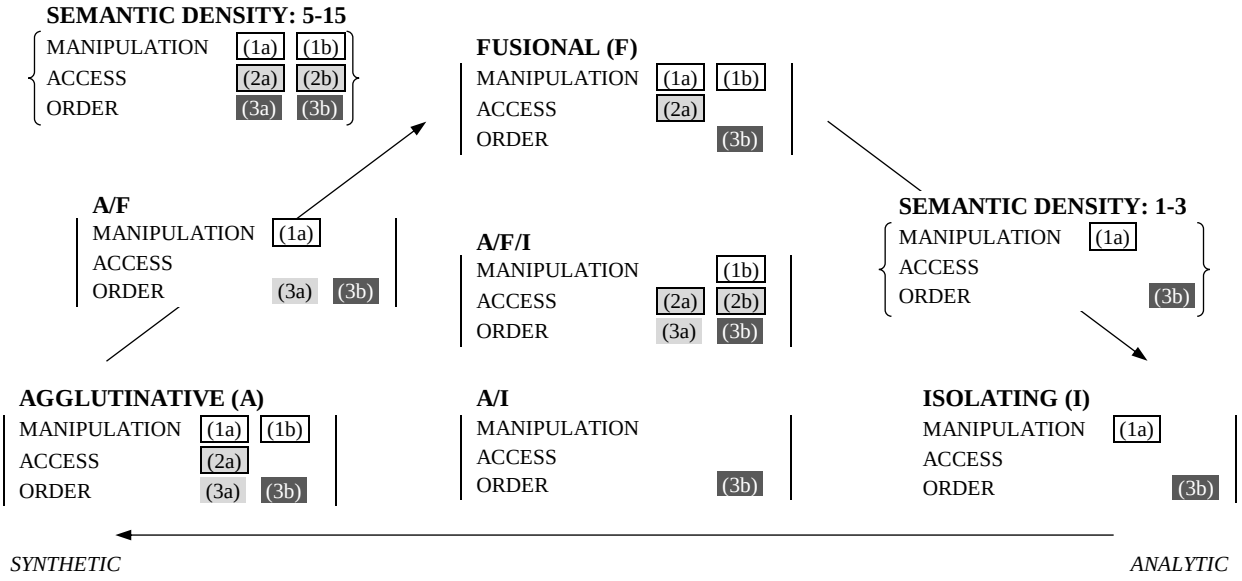


Figure 5 Key

1a	MANIPULATION (movement)	concatenative/synthetic properties
1b	MANIPULATION (word-part ellipsis) †	
2a	ACCESS (syntax/semantics 'sees' morphology) †, ‡, §	isolating/analytic properties
2b	ACCESS (morphology 'sees' syntax and semantics) ‡, §	
3a	ORDER (word-level)	general properties of language (e.g. grammaticalization)
3b	ORDER (phrase-level) ‡	

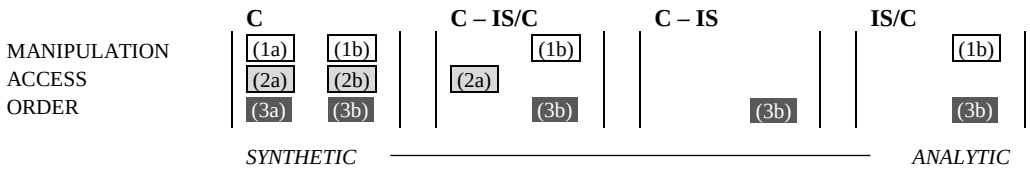
The lower right corner of the cycle represents highly analytic languages exhibiting ISOLATING morphology and very low degrees semantic density (indicated in curly brackets), along with their closely linked MANIPULATION (movement) and phrase-level ORDER violations outlined above. Furthermore, the dark gray shading represents a violation with a direct correlation to ISOLATING typological properties, specifically predominantly syntactic methods of linguistic expression. The violation outlined in black represents one whose source is argued to be a general property of language and language change, namely the function of syntax and semantics at various stages of grammaticalization and lexicalization (e.g. the movement of more loosely linguistic units, such as particles and clitics in more analytic languages, and the movement of more tightly bound elements in more synthetic languages).

As languages become more AGGLUTINATIVE (clockwise from the lower left in **Figure 5**), phrase-level ORDER violations persist due to the continued presence of analytic tendencies and general

role of syntax. However, as languages continue to undergo processes of grammaticalization and lexicalization (AGGLUTINATIVE to FUSIONAL), they begin to exhibit more LI violations, all of which are closely associated with concatenative morphology (indicated in light gray shading) and higher degrees of semantic density (indicated in curly brackets in the upper left of **Figure 5**) – specifically, ACCESS, MANIPULATION (word-part ellipsis, also re-analyzable as suspended affixation, i.e. ACCESS), and word-level ORDER violations. ACCESS violations are indicated as a general (semantic) property of language, since syntax undoubtedly requires access to information encoded in various word forms, while at the same time considered a largely synthetic association, since semantic information encoded on a single word form, complex morphology, and the potential for word-internal syntactic ACCESS seem positively correlated.

This relationship between LI violations and various stages of morphological change is also supported when considered in terms of overall type of phonological fusion and the linear scale of unidirectional change, from synthetic/concatenative (C) to analytic/isolating (although still demonstrating some concatenative tendencies (IS/C)) (**Figure 6** below). As languages become less synthetic and exhibit more clearly identifiable isolating tendencies, fewer violations (as correlated to higher semantic density and complex morphology) are observed, leaving only phrase-level ORDER and MANIPULATION (movement) violation types on the analytic end.

Figure 6 Phonological Fusion, LI Violation Types, and Linear Change



Languages with highly mixed typological profiles (represented in the center of **Figure 5**) therefore, and unsurprisingly, demonstrate essentially every violation type, as correlated with each of their typological tendencies, general linguistic properties, and various stages of language change. To complete the cycle, as languages undergo further diachronic processes (such as semantic bleaching and further lexicalization) they tend to become more ISOLATING, and accordingly exhibit analytic-type LI violations.

LI, and the theoretical assumptions that underlie it, are thus based on an ideal typological trait, particularly analytic tendencies toward syntactic expression. This is demonstrated by the few violation types manifested across highly ISOLATING languages, and the types that are observed likely arise from the effects of diachronic change (i.e. syntax more or less becoming morphology, and vice versa). The correlation between LI violations and synthetic morphology in general, and the between each type and specific synthetic traits (i.e. high degrees of semantic density, concatenative phonological fusion, and mixed typological profiles) further show LI and its foundations lie in a syntacticocentric view of grammar and the morphology-syntax interface, since the ACCESS of semantically dense words, MANIPULATION of complex word forms, and syntax-morphology ORDER violations present significant inherent challenges to theories based on such a conceptualization of grammar. Therefore, an proper grammatical theory of the morphology-syntax interface should be able to accommodate the diachronic effects on morphology and syntax, and more crucially model and describe the synchronic interaction between words and word-parts, phrases, and clauses, along with their meanings, regardless of a language's typological characteristics.

5. Concluding Remarks: On the Nature of the Morphology-Syntax Interface

It turns out that LI can be quite useful when actually applied as a hypothesis, as various titles suggest (e.g. Extended Lexicalist Hypothesis (Jackendoff 1972:13), Generalized Lexicalist Hypothesis (LaPointe 1985[1980]:8) etc.), as opposed to a constraint or principle to be maintained. And after a careful evaluation of the predictions a maximally strong version of LI makes (i.e. violations) in relation to the empirical data, theoretical assumptions, and typological considerations, it can be concluded that LI, as both a theoretical parameter and conceptualization of the morphology-syntax interface, is incorrect.

The tenability of LI is at least superficially challenged by the overall difficulty in defining its corresponding predictions (i.e. violations) in any universal manner; as noted in **Section 1.2** (and **Section 2**, f.n. 17), certain potential violations can be recharacterized depending on whether a morphological or syntactic analysis is chosen. However, assuming distinct classes of LI violations

(**Section 2**), the correlations between each violation (sub-)type and specific theoretical parameters observed in **Section 3.2** present a greater challenge to LI and traditional assumptions of grammar. In particular, the relationship between STRONG and WEAK LEXICALIST assumptions and LINEAR MODEL grammatical architectures suggest theories inheriting such parameters (e.g. Lapointe 1985[1980], Aronoff 1994, etc.) provide the least ideal characterization of the morphology-syntax interface. This is also true for theories with STRONGLY LEXICALIST – MODULAR (in correspondence) assumptions (e.g. LFG), since a STRONGLY LEXICALIST viewpoint typically forces one to consider challenges to LI as strictly morphological or syntactic phenomena, assigned to distinct (parallel) syntactic and morphophonological components.

Furthermore, the correlations between each LI violation type and particular typological traits in **Section 4.2** help identify which types of empirical phenomena a proper theory of the morphology-syntax interface should accommodate and reveals the narrow typological foundations of LI and its associated theoretical parameters. Certain MANIPULATION (movement) and phrase-level ORDER violations, which correlate with analytic tendencies (e.g. typically low degrees of semantic density and syntactic methods of grammatical expression), demonstrate that any descriptive and explanatory theory of the morphology-syntax interface should yield to such instances of apparent bidirectional interaction. More importantly, however, the interrelationship between each MANIPULATION, ACCESS, and ORDER violation type and subtype, specific synthetic typological properties (e.g. higher semantic density and concatenative morphology), and diachronic effects on morphosyntax shows a fully descriptive and explanatory grammatical theory cannot be based on select analytic typological characteristics alone, but must be cross-typologically applicable.

Synthesizing across the typological survey observations (**Section 4.2**), the specific theoretical parameter(s) underlying each violation type (**Section 3.2**), and the development and current state of LI (**Section 1.1** and **1.2**) reveals the basic problem with, and circular nature of, LI. English served as the basis for Chomsky's (1957, 1965) original conceptualization of grammar as a syntactically driven system; specifically, particular variations in English word order were taken as the method by which to understand linguistic "competence", or "the underlying system of rules [i.e. grammar] that has been mastered by the speaker-hearer and that [s]he puts to use in actual performance" (1965:4). As a result, largely analytic typological characteristics were considered in

both the formation and continued development of the theory and its various parameters (a derivational FRAMEWORK and LINEAR model, by virtue of syntactic transformations being the theoretical foundation). When particular synthetic properties of English were subsequently taken into account, a pre-syntactic word formation component was proposed (Chomsky 1970), resulting in the gradual formulation of LEXICALISM and LINEAR MODELS. As additional synthetic and analytic typological traits were considered, the lexicalist spectrum emerged, along with its various modular implementations. STRONG LEXICALISM and TYPE (2) LINEAR MODELS were developed in order to maintain syntax and morphology as distinct components obeying separate grammatical rules – the syntactic component based on analytic typological properties, particularly word order, and the morphological component (and lexicon) based on synthetic tendencies in English and other select languages (e.g. Di Sciullo and Williams (1987), Lapointe (1985[1980]). Unsurprisingly then STRONGLY LEXICALIST – TYPE (2) theories are presented with the most serious threat when evaluated in terms of potential LI violations, since intermodular morphosyntactic interaction, and synthetic typological characteristics, were not viewed as integral to the theoretical goals in the first place.

WEAKLY LEXICALIST – TYPE (3) theories are the result of maintaining the syntacticocentric theoretical foundations, but selectively integrating some morphological phenomena into the syntactic component (e.g. instances of FUSIONAL morphology and high semantic density in French and Italian verbal inflection (Rizzi 1997)), and otherwise working to treat distinct derivational processes morphologically while assuming inflection to be part of syntax (e.g. Aronoff 1976). Other theoretical approaches were developed as reactions to, but still adopting specific assumptions made in early Transformational Grammar (e.g. Chomsky 1957, 1965, 1970). For instance, as Van Valin (2009:2) notes, "a leading idea in the development of [GPSG] was to determine whether natural language syntax could be adequately described in terms of a context-free phrase structure grammar" (cf. Chomsky 1957, 1964), while LFG involved "applying the formalism of unification grammar to natural language phenomena and showing that lexical rules were superior to transformational rules" (pp. 2). In the case of the former (as well as its descendant HPSG), the syntacticocentric typological viewpoint was maintained by virtue of its focus on phrasal combination, and in the case of both GPSG (later, HPSG) and LFG, strong LEXICALISM persisted, with morphology treated via separate word formation rules (LFG) and generally

considered realizational in nature (HPSG) (Nordlinger and Sadler 2016). Therefore, LI violations are variably correlated with MODULAR (in correspondence) theories, based on how a particular theorist treats the inherited (problematic) theoretical parameters in relation to instances of morphosyntactic interaction. Although still assuming a form of UNIFICATION-BASED – MODULAR (in correspondence) analysis, Role and Reference Grammar (RRG) was developed in recognition of diverse typological structures (e.g. Lakhota, Tagalog, Dyirbal, and Barai), and seeks to understand how the interaction between syntax, semantics, and pragmatics across such diverse typological profiles can be modeled and explained (Van Valin 2009:2; see also Van Valin and LaPolla 1997). Given the theoretical goals and general focus on the interrelationship between syntax, semantics, and pragmatics particularly, future work will hopefully explore the precise nature of LI and the morphology-syntax interface within specific implementations of RRG.

As noted in **Section 3.2**, those theories which remove the boundary between morphology and syntax fare the best when considered in terms of each LI violation type, with the question then being whether morphology is just an extension of a single syntactic component, or if morphology and syntax truly form a single (unified) morphosyntactic system. The extension of syntactic rules (such as head movement) to account for highly synthetic typological tendencies in FULLY SYNTACTIC – INCREMENTAL approaches stem from, for example, noun incorporation in Mohawk (Baker 1985a) and choice data examples from English, Tagalog, French, Dutch, and several other languages (Lieber 1992)[49]. This seems a bit strange – considering such theories assume the basic narrow conceptualization of grammar (Chomsky 1957, 1965) as strictly formulated on English's tendency toward syntactic (analytic) means of grammatical expression, it therefore does not seem appropriate to describe and explain all aspects of grammar – morphological, syntactic, or otherwise – in terms of these generative approaches based one typological trait. A FULLY SYNTACTIC – REALIZATIONAL theory such as DM[50], which treats derived and inflected word forms as the result of syntactically manipulated feature bundles, that are then phonologically spelled-out via lexical insertion late in the syntactic derivation (Marantz 1995:379), suffers from the same fundamental

[49] The syntacticization of various morphological phenomena, whether through REALIZATIONAL or INCREMENTAL methods, is intimately related to the assumptions underlying "the Great Agglutinative Fraud" (Hockett 1987:82-84).

[50] DM is based on insights drawn from non-concatenative, templatic morphological patterns (e.g. largely non-categorical roots) predominant among Semitic languages (Jackendoff, personal communication, September 16, 2017)

issue as FULLY SYNTACTIC – INCREMENTAL theories; it seeks to model all aspects of linguistic expression in terms of theoretical assumptions based on very narrow typological considerations.

It can therefore be concluded that LI is incorrect in both theory and principle, and probably so too are the predominant conceptualizations of grammar in which it was developed and subsequently maintained. The relationship between morphology and syntax is likely much more fluid than previously conceived, since the notion of LI and a modular morphology-syntax interface is (i) evidently violable, (ii) completely theoretically sensitive, (iii) developed from a strictly ISOLATING typological viewpoint. Accordingly, any proper grammatical theory should not simply seek descriptive and explanatory adequacy of a single linguistic domain, from which to extrapolate in an attempt to understand the system in its entirety, but rather it should seek to characterize a given linguistic system in terms of all its relevant properties, whether phonological, morphosyntactic, semantic, pragmatic, or cognitive in nature. Current UNIFIED linguistic theories based on CxG (e.g. Goldberg 1995) and PA (e.g. Jackendoff 1997, 2015), which assume no division between syntax and morphology, or between the grammar and the lexicon, and model both phrasal and productive morpheme combinations as various interrelationships between linguistic forms and their associated semantics, are the most amenable to the present considerations and observations outlined in this paper. Ultimately, if the goal is a proper characterization of the morphology-syntax interface, then one must abandon the notion of LI as well as the theoretical foundations from which is arose, and look instead to the insights of CxG[51] and PA.

[51] Syntactic theories such as Sign-Based Construction Grammar (SBCG; Sag 2007, 2012) may be well equipped to model both morphology and syntax, and the manner in which they interact. SBCG models language as an infinite set of signs, represented as typed feature structures, that syntactically combine according to combinatoric constructions. Since words, like phrases, are modeled as typed feature structures, both words and phrases can be mother nodes in constructions—the only difference being that words are the mother nodes of inflectional and derivational ('lexical rule') constructions, while phrases are the mothers of phrasal constructions. A declarative constructional/SBCG model of morphology and syntax would have three desirable results for LI and linguistic theory: (i) it would permit instances where syntax appears to ACCESS word-internal information, since feature percolation can ensure that this information is locally available (e.g., in an argument-structure list, for the purpose of stating binding constraints); (ii) it would rule out syntactic MANIPULATION of word-internal components because the sub-lexical units are not possible daughters in phrasal constructions; and (iii) it would license phrases inside words (alleged ORDER violations) because both phrases and lexemes can be daughters in derivational constructions, like that which licenses nominal compounds. Full integration will hopefully be considered in future work.

6. References

Ackema, P., & Neeleman, A. (2004). *Beyond Morphology: Interface Conditions on Word Formation*. New York, NY: Oxford University Press.

Ackerman, F., & Webelhuth, G. (1997). Complex Predicates and Lexicalism. In *A Theory of Predicates* (pp. 1–36). Stanford: CSLI Publications.

Akkuş, F. (2015). Suspended Affixation with Derivational Suffixes and Lexical Integrity. *On-Line Proceedings of the Tenth Mediterranean Morphology Meeting (MMM10)*, 1–15.

Anderson, S. (1982). Where's Morphology? *Linguistic Inquiry*, *13*(4), 571–612.

Anderson, S. (1992). *A-morphous Morphology*. Cambridge, United Kingdom: Cambridge University Press.

Andrews, Avery. 1990. Unification and Morphological Blocking. Natural Language and Linguistic Theory 8(4):507–558.

Aronoff, M. (1976). *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.

Aronoff, M. (1994). *Morphology by Itself: Stems and Inflectional Classes*. Cambridge, MA: The MIT Press.

Artstein, R. (2004). Focus below the Word Level. *Natural Language Semantics*, *12*(1), 1–22.

Baker, M. (1985a). *Incorporation: A Theory of Grammatical Function Changing* (Dissertation). Massachusetts Institute of Technology, Cambridge, MA.

Baker, M. (1985b). The Mirror Principle and Morphosyntactic Explanation. *Linguistic Inquiry*, *16*(3), 373–415.

Baldi, P. (Ed.). (1990). *Linguistic Change and Reconstruction Methodology*. New York, NY: Mouton de Gruyter.

Bauer, L. (1983). *English Word-Formation*. New York, NY: Cambridge University Press.

Beard, R. (1998). Derivation. In A. Spencer & A. Zwicky (Eds.), *The Handbook of Morphology* (pp. 44–65). Hoboken, NJ: John Wiley & Sons.

Bickel, B., & Nichols, J. (2007). Inflectional Morphology. To appear in T. Shopen (Ed.), *Language Typology and Syntactic Description* (2nd ed., pp. 169–240). New York, NY: Cambridge University Press.

Bickel, B., & Nichols, J. 2013a. Fusion of Selected Inflectional Formatives. In M. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved from http://wals.info/chapter/20

Bickel, B., & Nichols, J. 2013b. Inflectional Synthesis of the Verb. In M. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved from http://wals.info/chapter/22

Booij, G. (1993). Against Split Morphology. In G. Booij & J. Van Marle (Eds.), *Yearbook of Morphology 1993* (pp. 27–50). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Booij, G. (2005). Construction-dependent Morphology. *Lingue E Linguaggio*, *4*(2), 163–178.

Booij, G. (2007). Construction Morphology and the Lexicon. *Selected Proceedings of the 5th Décembrettes: Morphology in Toulouse*, 34–44.

Booij, G. (2009). Lexical Integrity as a Formal Universal: A Constructionist View. In S. Scalise, E. Magni, & A. Bisetto (Eds.), *Universals of Language Today* (pp. 83–100). Dordrecht, The Netherlands: Springer.

Booij, G. (2010). Construction Morphology. *Language and Linguistics Compass*, *4*(7), 543–555.

Booij, G. (2013). Morphology in Construction Grammar. In T. Hoffmann & G. Trousdale (Eds.), *The Oxford Handbook of Construction Grammar* (pp. 255–273).

Borer, H. (1998). Morphology and Syntax. In A. Spencer & A. Zwicky (Eds.), *The Handbook of Morphology* (pp. 151–190). Hoboken, NJ: John Wiley & Sons.

Botha, R. (1981). A Base Rule Theory of Afrikaans Synthetic Compounding. In M. Moortgat, H. van der Hulst, & T. Hoekstra (Eds.), *The Scope of Lexical Rules* (pp. 1–77). Dordrecht, The Netherlands: Foris Publications.

Bresnan, J., Asudeh, A., Toivonen, I., & Wechsler, S. (2016[2001]). *Lexical-Functional Syntax* (2nd ed.). West Sussex: John Wiley & Sons, Inc.

Bruening, B. (2018). The Lexicalist Hypothesis: Both Wrong and Superfluous. *Language*, *94*(1), 1–42.

Carnie, A. (2000). On the Definition of X0 and XP. *Syntax*, *3*(2), 59–106.

Chaves, R. (2008). Linearization-Based Word-Part Ellipsis. *Linguistics and Philosophy*, *31*(3), 261–307.

Chomsky, N. (1957). *Syntactic Structures*. New York, NY: Mouton de Gruyter.

Chomsky, N. (1964). *Current Issues in Linguistic Theory*. New York, NY: Mouton de Gruyter.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1970). Remarks on Nominalization. In J. Roderick & P. Rosenbaum (Eds.), *Readings in English Transformation Grammar* (pp. 184–221). Boston, MA: Ginn and Company.

Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht, The Netherlands: Foris Publications Holland.

Chomsky, N. (1982). *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, MA: The MIT Press.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. New York, Evanstown, and London: Harper & Row, Publishers.

Comrie, B. (1989[1981]). *Language Universals and Linguistic Typology* (2nd ed.). Chicago, IL: University of Chicago Press.

de Saussure, F. (1983). *Course in General Linguistics*. (C. Bally & A. Sechehaye, Eds., R. Harris, Trans.). Chicago, IL: Open Court Publishing Company.

Di Sciullo, A.-M., & Williams, E. (1987). *On the Notion of Word*. Cambridge, MA: The MIT Press.

Dixon, R. M. W. (1994). *Ergativity*. New York, NY: Cambridge University Press.

Felíu, E., & Fábregas, A. (2003). Phrasal Scope and Argument Constraints of Spanish inter-Prefixation. *Topics in Morphology. Selected Papers from the Third Mediterranean Morphology Meeting (MMM3)*, 165–176.

Fillmore, C. (1988). The Mechanisms of "Construction Grammar." *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*, 35–55.

Gazdar, G., Klein, E., Pullum, G., & Sag, I. (1985). *Generalized Phrase Structure Grammar*. Cambridge, MA: Harvard University Press.

Goldberg, A. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago, IL: University of Chicago Press.

Goldberg, A. (2006). *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

Greenberg, J. (1963). Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In J. Greenberg (Ed.), *Universals of Language* (pp. 73–113). London: The MIT Press.

Haeringen, C. B. van. (1958). Vervoegde voegwoorden in het Oosten. *Driemaandelijkse Bladen*, *10*, 115–124.

Halle, M., & Marantz, A. (1993). Distributed Morphology and the Pieces of Inflection. In K. Hale & S. J. Keyser (Eds.), *The View from Building 20* (pp. 111–176). Cambridge, MA: MIT Press.

Harley, H. (2015). The Syntax/Morphology Interface. In T. Kiss & A. Alexiadou (Eds.), *Syntax, Theory and Analysis: An International Handbook* (Vol. 2, pp. 1128–1154). Boston, MA: Mouton de Gruyter.

Harris, A. (2002). *Endoclitics and the Origins of Udi Morphosyntax*. New York, NY: Oxford University Press.

Harris, A. (2006). In Other Words: External Modifiers in Georgian. *Morphology*, *16*(2), 205–229.

Haspelmath, M. (2011). *The Indeterminacy of Word Segmentation and the Nature of Morphology and Syntax*. Presented at the Morphology of the World's Languages, University of Leipzig.

Hockett, C. (1987). *Refurbishing Our Foundations: Elementary Linguistics from an Advanced Point of View*. Philadelphia, PA: John Benjamins Publishing Company.

Hopper, P., & Traugott, E. (2003[1993]). *Grammaticalization* (2nd ed.). Cambridge, United Kingdom: Cambridge University Press.

Huang, J. (1984). Phrase Structure, Lexical Integrity, and Chinese Compounds. *Journal of the Chinese Language Teachers Association*, *19*(2), 53–78.

Jackendoff, R. (1972). *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Jackendoff, R. (1997). *The Architecture of the Language Faculty*. Cambridge, MA: The MIT Press.

Jackendoff, R. (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution*. New York, NY: Oxford University Press.

Jackendoff, R. (2010). The Ecology of English Noun-Noun Compounds. In R. Jackendoff (Ed.), *Meaning and the Lexicon: The Parallel Architecture 1975-2010* (pp. 413–451). New York, NY: Oxford University Press.

Jackendoff, R., & Wittenberg, E. (2014). What You can Say without Syntax: A Hierarchy of Grammatical Complexity. In F. Newmeyer & L. Preston (Eds.), *Measuring Grammatical Complexity* (pp. 65–82). New York, NY: Oxford University Press.

Jackendoff, R. (2015). The Parallel Architecture and its Place in Cognitive Science. In B. Heine & H. Narrog (Eds.), *The Oxford Handbook of Linguistic Analysis* (2nd ed., Vol. 1, pp. 593–614). New York, NY: Oxford University Press.

Jackendoff, R., & Audring, J. (2016). Morphological Schemas: Theoretical and Psycholinguistic Issues. *The Mental Lexicon*, *11*(3), 467–493.

Jespersen, O. (1922). *Language, Its Nature, Development and Origin*. London: George Allen & Unwin LTD.

Josefsson, G. (1998). *Minimal Words in a Minimal Syntax* (Vol. 19). Philadelphia, PA: John Benjamins Publishing Company.

Kageyama, T. (2001). Word Plus: The Intersection of Words and Phrases. In J. Van de Weijer & N. Nishihara (Eds.), *Issues in Japanese Phonology and Morphology* (pp. 245–276). Berlin: Mouton de Gruyter.

Kendall, M. (1976). *Selected Problems in Yavapai Syntax: The Verde Valley Dialect*. New York, NY: Garland Publishing, Inc.

Kornfeld, L., & Saab, A. L. (2003). Morphology and Syntax: Prepositional Prefixes in Spanish. *Topics in Morphology. Selected Papers from the Third Mediterranean Morphology Meeting (MMM3)*, 227–240.

Kubozono, H. (1995). *Gokeisei-to on-in-kouzo*. Tokyo: Kuroshio.

Lapointe, S. (1985). *A Theory of Grammatical Agreement*. New York & London: Garland Publishing, Inc.

Lefebvre, C., & Muysken, P. (1988). *Mixed Categories*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Lewis, G. L. (1967). *Turkish Grammar*. New York, NY: Oxford University Press.

Lieber, R. (1980). *On the Organization of the Lexicon* (Dissertation). Massachusetts Institute of Technology, Cambridge, MA.

Lieber, R. (1992). *Deconstructing Morphology: Word Formation in Syntactic Theory*. Chicago, IL: The University of Chicago Press.

Lieber, R., & Scalise, S. (2007). The Lexical Integrity Hypothesis in a New Theoretical Universe. *On-Line Proceedings of the Fifth Mediterranean Morphology Meeting (MMM5)*, 1–24.

Marantz, A. (1995). The Minimalist Program. In G. Webelhuth (Ed.), *Government and Binding Theory and the Minimalist Program* (pp. 351–382). London: Blackwell.

Meibauer, J. (2007). How Marginal are Phrasal Compounds? Generalized Insertion, Expressivity, and I/Q- Interaction. *Morphology*, *17*, 233–259.

Mukai, M. (2006). *A Comparative Study of Compound Words in English, Japanese, and Mainland Scandinavian* (Dissertation). University of Newcastle, Callaghan, New South Wales, Australia.

Müller, S. (2006). Phrasal or Lexical Constructions. *Language*, *82*(4), 850–883.

Müller, S. (2016). *Grammatical Theory: From Transformational Grammar to Constraint-Based Approaches*. Berlin: Language Science Press.

Müller, S. (2018). The End of Lexicalism as We Know It? *Language*, *94*(1), e54–e66.

Müller, S., & Wechsler, S. (2014). Lexical Approaches to Argument Structure. *Theoretical Linguistics*, *40*, 1–76.

Nespor, M. (1985). The Phonological Word in Italian. In H. van der Hulst & N. Smith (Eds.), *Advances in Nonlinear Phonology* (pp. 193–204). Dordrecht, The Netherlands: Foris Publications.

Nordlinger, R., & Sadler, L. (2016). Morphology in LFG and HPSG. To appear in J. Audring & F. Masini (Eds.), *Oxford Handbook of Morphological Theory*. New York, NY: Oxford University Press.

Pollard, C., & Sag, I. (1987). *Information-Based Syntax and Semantics* (Vol. 1). Stanford, CA: CSLI Publications.

Pollard, C., & Sag, I. (1994). *Head-Driven Phrase Structure Grammar*. Chicago, IL: University of Chicago Press.

Postal, P. (1969). Anaphoric Islands. In R. Binnick, A. Davison, G. Green, & J. Morgan (Eds.), *Papers from the Fifth Regional Meeting of the Chicago Linguistic Society* (pp. 205–239). Chicago, IL: Chicago Linguistic Society.

Rainer, F., & Varela, S. (1992). Compounding in Spanish. In S. Scalise (Ed.), special issue of *Rivista di Linguistica* (pp. 117–140).

Sadock, J. (1991). *Autolexical Syntax: A Theory of Parallel Grammatical Representations*. Chicago, IL: University of Chicago Press.

Sag, I. (2007). Remarks on Locality. In S. Müller (Ed.), *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar* (pp. 394–414). Stanford: CSLI Publications.

Sag, I. (2012). Sign-Based Construction Grammar: An Informal Synopsis. In H. C. Boas & I. A. Sag (Eds.), *Sign-Based Construction Grammar* (pp. 61–197). Stanford: CSLI Publications.

Sag, I., Kaplan, R., Karttunen, L., Kay, M., Pollard, C., Shieber, S., & Zaenen, A. (1986). Unification and Grammatical Theory. *Proceedings of the West Coast Conference on Formal Linguistics*, *5*, 238–254.

Sapir, E. (1921). *Language: An Introduction to the Study of Speech*. New York, NY: Harcourt, Brace.

Savini, M. (1983). *Phrasal Compounds in Afrikaans: A Generative Analysis* (Masters). University of Stellenbosch, Stellenbosh, Republic of South Africa.

Scalise, S. (1986). *Generative Morphology*. Dordrecht, The Netherlands: Foris Publications Holland.

Schachter, P., & Otanes, F. (1972). *Tagalog Reference Grammar*. Berkeley and Los Angeles: University of California Press.

Selkirk, E. (1982). *The Syntax of Words*. Cambridge, MA: The MIT Press.

Selkirk, E. (1984). *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: MIT Press.

Shibatani, M., & Kageyama, T. (1988). Word Formation in a Modular Theory of Grammar: Postsyntactic Compounds in Japanese. *Language*, *64*, 451–484.

Shieber, S. M. (1986). *An Introduction to Unification-Based Approaches to Grammar*. Brookline, MA: Microtome Publishing.

Shieber, S. (1992). *Constraint-Based Grammar Formalisms: Parsing and Type Inference for Natural and Computer Languages*. Cambridge, MA: The MIT Press.

Simpson, J. (1983). *Aspects of Warlpiri Morphology and Syntax* (Dissertation). Massachusetts Institute of Technology, Cambridge, MA.

Spencer, A. (1991). *Morphological Theory: An Introduction to Word Structure in Generative Grammar*. Malden: Blackwell Publishers Inc.

Spencer, A. (2005). Word Formation and Syntax. In P. Štekauer & R. Lieber (Eds.), *The Handbook of Word-Formation* (Vol. 64, pp. 73–97). The Netherlands: Springer.

Steedman, M. (2000). *The Syntactic Process*. Cambridge, MA: MIT Press.

Stump, G. (2001). *Inflectional Morphology: A Theory of Paradigm Structure*. New York, NY: Cambridge University Press.

Subramanian, U. (1988). Subcategorization and Derivation: Evidence from Tamil. In D. Brentari, G. Larson, & L. MacLeod (Eds.), *Papers from the Twenty-fourth Annual Regional Meeting of the Chicago Linguistic Society,* (pp. 353–361). Chicago, IL: Chicago Linguistic Society.

Tokizaki, H. (2017). Prosody and Branching Direction of Phrasal Compounds. *Proceedings of the Linguistic Society of America*, *2*, 21:1-10.

Tokizaki, H., & Kuwana, Y. (2008). *Phrasal Compound and Marked Word Orders*. PowerPoint presented at the Morphology-Lexicon Forum 2008, Kobe University. Retrieved from http://toki.nagomix.net/e/files/dl.php?id=51

Toman, J. (1998). Word Syntax. In A. Spencer & A. Zwicky (Eds.), *The Handbook of Morphology* (pp. 306–321). Hoboken, NJ: John Wiley & Sons.

Trips, C. (2012). Empirical and Theoretical Aspects of Phrasal Compounds: Against the "Syntax Explains It All" Attitude. *On-Line Proceedings of the Eighth Mediterranean Morphology Meeting (MMM8)*, 322–346.

Van Valin, Jr., R. D. (2009). *An Overview of Role and Reference Grammar*. Retrieved from http://www.acsu.buffalo.edu/~rrgpage/rrg/RRG_overview.pdf

Van Valin, Jr., R., & LaPolla, R. (1997). *Syntax: Structure, meaning and function*. New York: Cambridge University Press.

Vigario, M. (2003). *The Prosodic Word in European Portuguese*. Berlin: Mouton de Gruyter.

Wennerstrom, A. (1993). Focus on the Prefix: Evidence for Word-Internal Prosodic Words. *Phonology*, *10*, 309–324.

Wiese, R. (1996). Phrasal Compounds and the Theory of Word Syntax. *Linguistic Inquiry*, *27*(1), 183–193.

Williams, E. (1981). On the Notions "Lexically Related" and "Head of a Word." *Linguistic Inquiry*, *12*(2), 245–274.

Yoon, J. H. S. (2017). Lexical Integrity and Suspended Affixation in Two Types of Denominal Predicates in Korean. *Glossa: A Journal of General Linguistics*, *2*(1), 1–45.

7. Appendices

7.1. List of Abbreviations

*	grammatically unacceptable
-	morpheme boundary
.	indicates more than one grammatical category per single morpheme
1	first person
2	second person
3	third person
A	agglutinative
ABS	absolutive case
ACC	accusative case
ADJ	adjective
ADV	adverbial ending
C	concatenative
CCG	Combinatory Categorial Grammar
CM	Construction Morphology
CxG	Construction Grammar
COP	copula
DET	determiner
DM	Distributed Morphology
D-structure	deep structure
ERG	ergative case
F	fusional
FUT	future tense
GB	Government and Binding
GEN	genitive case
GPSG	Generalized Phrase Structure Grammar
HPSG	Head-Driven Phrase Structure Grammar
I	isolating (traditional morphological type)
ILL	illative case

IMPF	imperfect (i.e. past tense, imperfective aspect)
IS	isolating (phonological fusion)
LFG	Lexical Functional Grammar
LI	Lexical Integrity
LOC	locative (case)
LV	light verb
MASC	masculine gender
MP	Minimalist Program
NEG	negation
NEUT	neuter gender
NMZ	nominalizer
NOM	nominative case
NP	noun phrase
OBJ	object
P&P	Principles and Parameters
PA	Parallel Architecture
PAST	past tense
PL	plural number
PRES	present tense
PREVERB	preverb
REV	reversative
RM	Relational Morphology
SBCG	Sign-Based Construction Grammar
SG	singular number
SMH	Split Morphology Hypothesis
S-structure	surface-structure
SUB	subject
SUFF	suffix
VP	verb phrase
VZR	verbalizer
YP	a phrasal category headed by some constituent Y

7.2. Taxonomy of Lexical Integrity Violations

1. Manipulation

a. Movement (into and out of words)		• **Estonian**, **Hungarian** preverbs (Ackerman and Webelhuth 1997:3-4) • **Mandarin Chinese** verb-object compounds (Huang 1984:64) • **Mohawk** noun incorporation (Lieber 1992:140, Baker 1985a) • **Udi** person markers/endoclitics (Harris 2002:3-4) • **Italian** *trasporto latte* constructions (Lieber and Scalise 2007:6-8)
b. Word-part ellipsis †		• **English** (Bruening 2017:11-13, Chaves 2008; Spencer 2005:82-83; Lieber and Scalise 2007:6, Spencer 2005:82) • **Italian** (Bruening 2017:11, Nespor 1985) • **Portuguese** (Booij 2009:88, Vigario 2003:251) • **German** (1998:321) • **Korean** (Yoon 2017:4) • **Turkish** (Lewis 1967, Akkuş 2015)

2. Access ‡

a. Syntax/semantics 'sees' morphology §	VP nominalization and constituent order	• **Quechua** (Ackema and Neeleman 2004:11, Lefebvre and Muysken 1988)
	Anaphoric islands	• **English** (Postal 1969, Lieber 1992:121-122)
	Suspended affixation (*gruppeninflection*) †	• **English** (Bruening 2017:11-13, Chaves 2008; Spencer 2005:82-83; Lieber and Scalise 2007:6, Spencer 2005:82) • **Italian** (Bruening 2017:11, Nespor 1985) • **Portuguese** (Booij 2009:88, Vigario 2003:251) • **German** (1998:321) • **Korean** (Yoon 2017:4) • **Turkish** (Lewis 1967, Akkuş 2015)
	Sublexical indexing	• **Yavapai** switch reference (Lieber 1992:124-125, Kendall 1976)
	Semantic scope phenomena	• **English (**Lieber and Scalise 2007:11-12) • **Spanish** (Lieber and Scalise 2007:11, Rainer & Varela 1992, Felíu & Fábregas 2003, Kornfeld & Saab 2003)
	Focus targeting sublexical units	• **English** (Bruening 2017:13, Selkirk 1984:271, Wennerstrom 1993, Artstein 2004)
	Sublexical access for case assignment	• **Georgian** (Booij 2009:88-89, Harris 2006)
b. Morphology 'sees' syntax/semantics §	Construction dependent morphology	• **Dutch** (Booij 2009:90-91)
	Agreement and verb position	• **Dutch** (East Netherlandic varieties) (Ackema and Neeleman 2004:11, Haeringen 1958)

3. Order

a. Word-level phenomena	Order of inflectional and derivational affixes	• **Sanskrit, Georgian** (Beard 1998)
b. Phrase-level phenomena ‡	Phrasal compounding	• **English** (Lieber 1992:11) • **German** (Meibauer 2007:236) • **Afrikaans** (Botha 1981:74-75) • **Dutch** (Ackema and Neeleman 2004:152) • **Swedish** (Mukai 2006:67) • **Japanese** (Shibatani and Kageyama 1988:455, Tokizaki 2017:2) • **Mandarin Chinese** (Wiese 1996:185)
	Words zero-derived from phrases	• **English** (Carney 2000:91, Bruening 2017:4-5)
	Words derived from constructions	• **German** nominalization and adjectival passive (Bruening 2017:10, Müller 2006; Müller and Wechsler 2014)
	Derivational morphology on phrases	• **English** (Spencer 2004:12; Bruening 2017:5-6) • **Dutch** nominalization (Ackema and Neeleman 2004:149) • **Tagalog** verbal derivation (Lieber 1992:16-17, Schachter and Otanes 1971) • **Tamil** nominalization (Lieber 1992:17-18, Subramanian 1988) • **Estonian** deverbal adjectives and nouns (Ackerman and Webelhuth 1997:3-4) • **Italian** phrasal derivation (Lieber and Scalise 2007:9) • **Japanese** *word plus* (W+) (Lieber and Scalise 2007:10-11, Kageyama 2001)
	Inflectional morphology on phrases	• **English** possessive markers (Lieber 1992:14) • **Warlpiri** case markers (Lieber 1992:14-15, Simpson 1983)
	Bracketing paradoxes	• **English** (Lieber 1992:147-148) • **Dutch** (Booij 2009:93-95)

7.3. Typological Survey Tables

7.3.1. Full Typological Survey of Languages Exhibiting LI Violations

Language	Language Family	Traditional Morphological Type(s)	Average Semantic Density	Semantic Density (Verb)	Fusion Type(s) (Verb)	Semantic Density (Noun)	Fusion Type(s) (Noun)
Afrikaans	Indo-European (Germanic)	A/F/I	5	2-3	C	2-3	C
Dutch	Indo-European (Germanic)	A/F/I	7	4-5	C	2-3	C
English	Indo-European (Germanic)	A/F/I	5	2-3	C	2-3	C
Estonian	Uralic	A/F	9	6-7	C	2-3	C
Georgian	Kartvelian	A	11	8-9	C	2-3	C
German	Indo-European (Germanic)	A/F/I	7	2-3	C	4-5	C
Hungarian	Uralic	A/F	7	4-5	C	2-3	C
Italian	Indo-European (Romance)	F	7	4-5	C	2-3	C
Japanese	Japonic	A	5	4-5	C	0-1	IS/C
Korean	Koreanic	A	7	6-7	C	0-1	IS/C
Mandarin Chinese	Sino-Tibetan (Chinese)	I	1	0-1	IS/C	0-1	IS/C
Mohawk (Akwesasne)	Iroquoian (Northern Iroquoian)	A	11	6-7	C	4-5	C
Portuguese	Indo-European (Romance)	F	7	4-5	C	2-3	C
Quechua (Huallaga)	Quechuan	A	13	8-9	C	4-5	C
Sanskrit	Indo-European (Indo-Iranian)	A/F	7	4-5	C	2-3	C
Spanish	Indo-European (Romance)	F	7	4-5	C	2-3	C
Swedish	Indo-European (Germanic)	A/F/I	5	2-3	C	2-3	C
Tagalog	Austronesian (Malayo-Polynesian)	A/I	3	2-3	C	0-1	IS
Tamil	Dravidian (Southern)	A	7	4-5	C	2-3	C
Turkish	Turkic (Southern)	A	11	6-7	C	4-5	C
Udi	North (East) Caucasian (Lezgic)	A/F	7	4-5	C	2-3	C
Walapai	Chochimí-Yuman (Yuman)	A	15	10-11	C	4-5	C
Warlpiri	Australian (Pama-Nyungan)	A	3	2-3	C	0-1	C

7.3.2. Inflectional Semantic Density of the Verb and LI Violation Types

Semantic Density (Verb)	1. Manipulation		2. Access‡		3. Order	
	Movement	Word-part ellipsis†	Syntax/semantics 'sees' morphology†,§	Morphology 'sees' syntax/semantics§	Word-level	Phrase-level‡
0-1	X					X
2-3		X	X			X
4-5	X	X	X	X	X	X
6-7	X	X	X			X
8-9			X		X	X
10-11			X			X

7.3.3. Inflectional Semantic Density of the Noun and LI Violation Types

Semantic Density (Noun)	1. Manipulation		2. Access‡		3. Order	
	Movement	Word-part ellipsis†	Syntax/semantics 'sees' morphology†,§	Morphology 'sees' syntax/semantics§	Word-level	Phrase-level‡
0-1	X	X	X			X
2-3	X	X	X	X	X	X
4-5	X	X	X			X

7.3.4. Phonological Fusion of Verbal Inflection and LI Violation Types

Fusion Type(s)	1. Manipulation		2. Access‡		3. Order	
	Movement	Word-part ellipsis†	Syntax/semantics 'sees' morphology†,§	Morphology 'sees' syntax/semantics§	Word-level	Phrase-level‡
IS/C	X					X
C	X	X	X	X	X	X

7.3.5. Phonological Fusion of Nominal Inflection and LI Violation Types

Fusion Type(s)	1. Manipulation		2. Access‡		3. Order	
	Movement	Word-part ellipsis†	Syntax/semantics 'sees' morphology†,§	Morphology 'sees' syntax/semantics§	Word-level	Phrase-level‡
IS						X
IS/C	X	X	X			X
C	X	X	X	X	X	X

7.4. Typological Survey Sources

Bickel, B., & Nichols, J. 2013a. Fusion of Selected Inflectional Formatives. In M. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved from http://wals.info/chapter/20

Bickel, B., & Nichols, J. 2013b. Inflectional Synthesis of the Verb. In M. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved from http://wals.info/chapter/22

Bonvillian, N. (1973). *A Grammar of Akwesasne Mohawk*. Ottawa: National Museum of Man and National Museums of Canada.

Booij, G. (2002). *The Morphology of Dutch*. New York, NY: Oxford University Press.

Bradley, P., & Mackenzie, I. (2004). *Spanish: An Essential Grammar*. New York, NY: Routledge.

Chang, S.-J. (1996). *Korean*. Philadelphia, PA: John Benjamins Publishing Company.

Donaldson, B. (1993). *A Grammar of Afrikaans*. New York, NY: Mouton de Gruyter.

Fox, A. (1990). *The Structure of German* (2nd ed.). New York, NY: Oxford University Press.

Harris, A. (2002). *Endoclitics and the Origins of Udi Morphosyntax*. New York, NY: Oxford University Press.

Hewitt, B. G. (1995). *Georgian: A Structural Reference Grammar*. New York, NY: John Benjamins Publishing Company.

Hinds, J. (1986). *Japanese: Descriptive Grammar*. New York, NY: Routledge.

Holmes, P., & Hinchliffe, I. (2013). *Swedish: A Comprehensive Grammar* (3rd ed.). New York, NY: Routledge.

Hutchinson, A., & Lloyd, J. (2003). *Portuguese: An Essential Grammar* (2nd ed.). New York, NY: Routledge.

Lehman, T. (1993). *A Grammar of Modern Tamil*. Pondicherry: Pondicherry Institute of Linguistics and Culture.

Lewis, G. L. (1967). *Turkish Grammar*. New York, NY: Oxford University Press.

Li, C., & Thompson, S. (1981). *Mandarin Chinese: A Functional Reference Grammar*. Berkeley and Los Angeles, CA: University of California Press.

Maiden, M., & Robustelli, C. (2013). *A Reference Grammar of Modern Italian* (2nd ed.). New York, NY: Routledge.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. New York, NY: Longman Inc.

Redden, J. E. (1966). Walapai II: Morphology. *International Journal of American Linguistics*, 32(2), 141–163.

Rounds, C. (2001). *Hungarian: An Essential Grammar*. New York, NY: Routledge.

Schachter, P., & Otanes, F. (1972). *Tagalog Reference Grammar*. Berkeley and Los Angeles, CA: University of California Press.

Shibatani, M. (1990). *The Languages of Japan*. New York, NY: Cambridge University Press.

Simpson, J. (1991). *Warlpiri Morpho-Syntax: A Lexicalist Approach*. Dordrecht: Springer.

Viitso, T.-R. (2007). Structure of the Estonian Language: Phonology, Morphology and Word Formation. In M. Erelt (Ed.), *Estonian Language* (2nd ed.). Tallinn: Estonian Academic Publishers.

Weber, D. J. (1947). *A Grammar of Huallaga (Huánuco) Quechua*. Berkeley and Los Angeles, CA: University of California Press.

Whitney, W. D. (1896). *A Sanskrit Grammar: Including Both the Classical Language, and the Older Dialects, of Veda and Brahmana* (3rd ed.). London: Kegan Paul, Trench, Trübner & Co Ltd.