Lexical integrity is the norm

or, "violating lexical integrity in a specific and limited way"

Jamie Y. Findlay

University of Oslo

LFG23, 24 July 2023

Preliminaries

Lexical integrity hypothesis: **empirical claim** vs. **design choice**

LIH ✔
- Integrity is unmarked
- Apparent exceptions require explanation

LIH ✘
- Violations are unmarked
- Apparent preference for integrity requires explanation

Outline

- Review of the data for and against LI

- Three more general arguments for LI

- Case study: coordination of prefixes

Evidence of LI

1. **Ellipsis in coordination** (Simpson 1991: 51)
    a. John liked the play, and Mary ~~liked~~ the movie.
    b. *John liked the play, and Mary dis~~liked~~ it.

2. **Modification** (Williams 2007: 384)
    a. [How complete] are your results?
    b. *[How complete]-ness do you admire?

3. **Splitting a constituent** (Cappelle 2022: 188)
    a. A masochist is [someone], I believe, [who likes pain].
    b. *Someone who likes pain is a [masoch], I believe, [-ist].

4. **Preposing an element** (Cappelle 2022: 188)
    a. Down$_i$ went the weapons __$_i$ on the ground.
    b. *de$_i$-conflict __$_i$ escalation

Violations of LI

5. **Compounding** (Capelle 2022: 189)
   a. [[over the fence]$_{PP}$ gossip]$_N$
   b. She sniffed and gave Dot a [[why-do-you-do-this-to-me]$_{CP}$ look]$_N$.

6. **Inflection** (Carnie 2000: 91; www)
   a. He [[I-don't-care]$_{IP}$-d]$_V$ his way out of the room.
   b. If you are [Thelma and Louis]$_{NP}$-ing]$_V$ please keep away from the cliffs.

7. **Derivation** (Bruening 2018: 6; Spencer 2005b: 83)
   a. His general [[ok-with-less-than-we-should-aim-for]$_{AdjP}$-ness]$_N$ makes him an undesirable candidate.
   b. a [why-does-it-have-to-be-me]$_{CP}$-ish]$_{Adj}$ expression

Syntax
👀
Morphology

8. **Coordination of prefixes** (Siegel 1974: 147; Spencer 2005b: 82)
   a. [mono- and tri]-syllabic
   b. [pro- and en]-clitics
   c. [socio- and politico]-economic
   d. [pre- and even to some extent post]-war (economics)
   e. [pro- as opposed to anti]-war
   f. [hypo- but not hyper]-glycaemic

Choosing a side

(1–4)

John liked the play, and Mary dis- it.
How completeness do you admire?
Someone who likes pain is a masoch-, I believe, -ist.
de-conflict escalation

- Impressionistically more fundamental.

- 'Core' rather than 'periphery'?

(5–8)

Over the fence gossip
He I-don't-care-d his way out of the room.
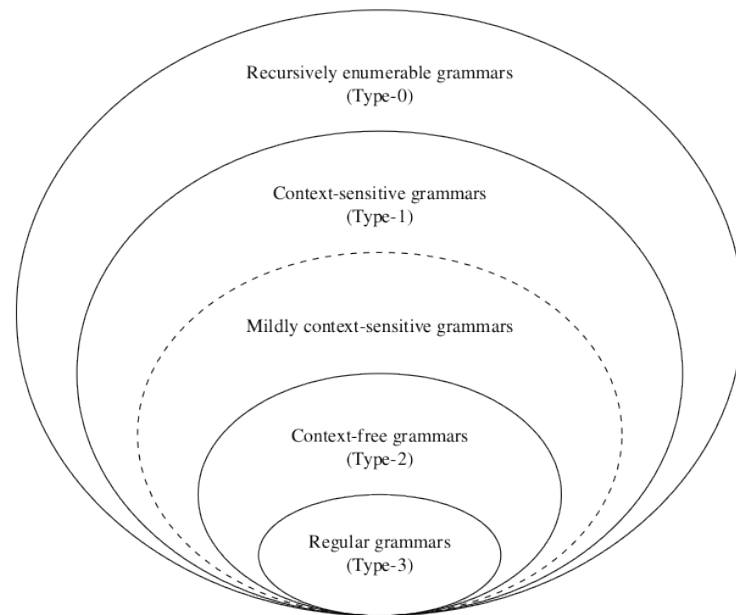*a why-does-it-have-to-be-me-ish expression*
pro- as opposed to anti-war

- Impressionistically more playful.

- 'Periphery' rather than 'core'?

Three more general arguments for LI

Formal language theory

> If morphology has the full power of syntax, why are there no clear morphological equivalents of unbounded or nested dependencies? [...] Similarly, why do we fail to find reduplication in the syntax, if there is no important formal distinction between morphology and syntax?
>
> (Asudeh et al. 2013: 5)

- Morphology almost entirely finite-state

- Syntax almost entirely context-free

Recursively enumerable grammars
(Type-0)

Context-sensitive grammars
(Type-1)

Mildly context-sensitive grammars

Context-free grammars
(Type-2)

Regular grammars
(Type-3)

Free word order ≠ free morpheme order

mil-es   coqu-um laud-at.
soldier-NOM *cook*-ACC *praise-3SG.PRES.INDIC*
'The soldier praises the cook.'

miles laudat coquum.

miles coquum laudat.

laudat miles coquum.

laudat coquum miles.

coquum miles laudat.

coquum laudat miles.

mil-es coqu-um laud-at.

*es-mil coqu-um laud-at

*mil-es um-coqu laud-at.

*mil-es coqu-um at-laud.

etc.

coquum laudat miles bonum

*coqu- laudat miles -um

LI is the norm

Any randomly selected stretch of discourse is likely to prove that morphologically complex words stay together as undivided units and that they tend not to include any above-word-level components.

(Cappelle 2022: 204)

LI violations as % of morphologically complex words
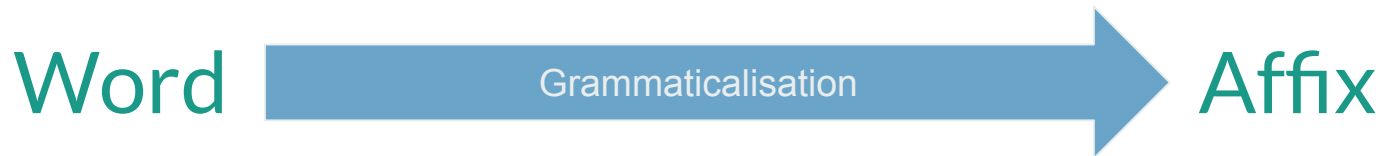

4%

LI violations as % of all words

1%

So how do we explain the apparent exceptions?

Non-projecting categories

Word $\xrightarrow{\text{Grammaticalisation}}$ Affix

X^0 $\qquad\qquad$ $\widehat{X}$ $\qquad\qquad$ —

$X^0 \to X^0, \widehat{Y}$

$\widehat{X} \to \widehat{X}, \widehat{Y}$
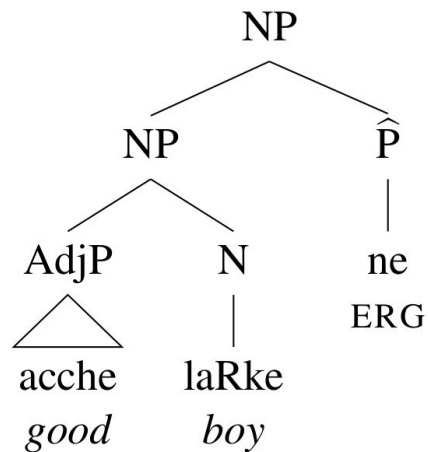
$XP \to XP, \widehat{Y}$

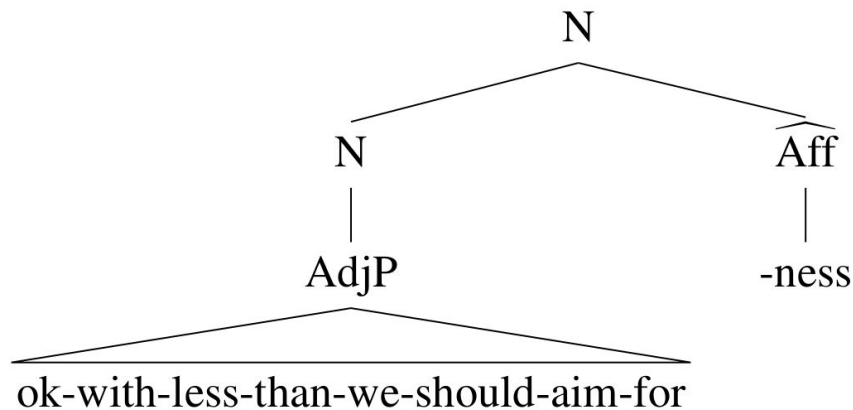(Toivonen 2003) $\qquad$ (Arnold & Sadler 2013) $\qquad$ (Spencer 2005a)

Explaining apparent LI violations:

- **Compounds** (Lowe 2015)
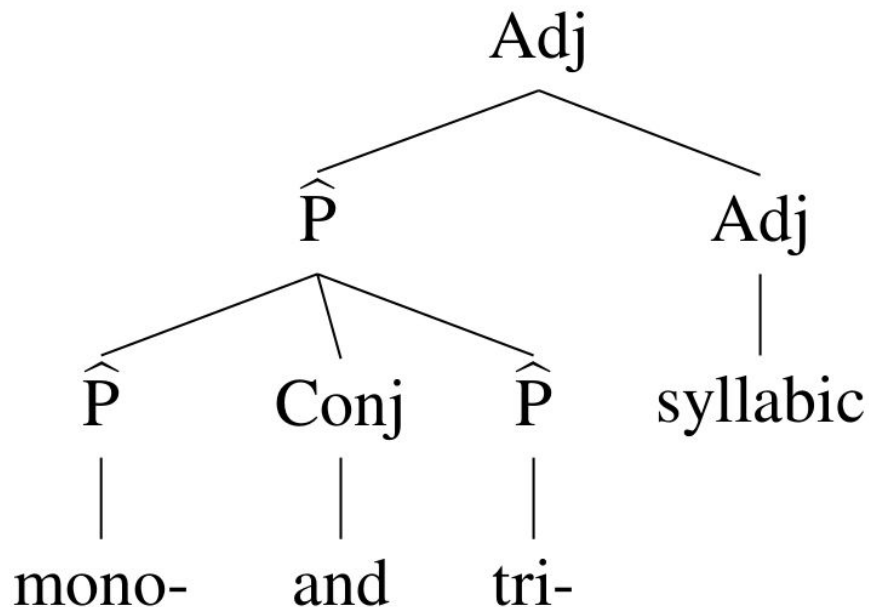- **'Phrasal affixes'** (Spencer 2005a)

Phrasal affixes

(Spencer 2005b)

Coordination of prefixes in English

Simple cases

9. **Coordination of prefixes**
   a. [mono- and tri]-syllabic
   b. [pro- and en]-clitics
   c. [socio- and politico]-economic
   d. [pre- and post]-war (economics)
   e. [pro- as opposed to anti]-war
   f. [hypo- but not hyper]-glycaemic

   Spencer (2005b): "**prefixoids**"

#NotAllAffixes

10. **Not all prefixes**
    a. *[un- or re]-tie
    b. *[im- or ex]-port
    c. *[ex- and se]-cretions

11. **No suffixes** (Strauss 1982: 43)
    a. *fear-[some and -less]
    b. *thought-[ful and -less]
    c. *interest-[ed and -ing]

Genuine affixes, not $\widehat{Xs}$.

☐ need not attach to lexical roots

Class I vs. Class II prefixes (Siegel 1974: 115ff.)
- Class I = more affix-like
- Class II = more word-like

☐ compounds?

12. **Cf. -ish** (Norde 2010)
    a. It's easy-ish.
    b. *A:* Is everyone excited? *B:* Ish.
    c. Tomorrow's an easy day (ish): graduation audit, voice lesson, CS lab …

Cross-linguistic variation

13. **German**
    a. ✔[Im- und Ex]-port

    b. Peter be- und Maria ent-lud   den LKW.
       Peter BE and Maria un-loaded the truck
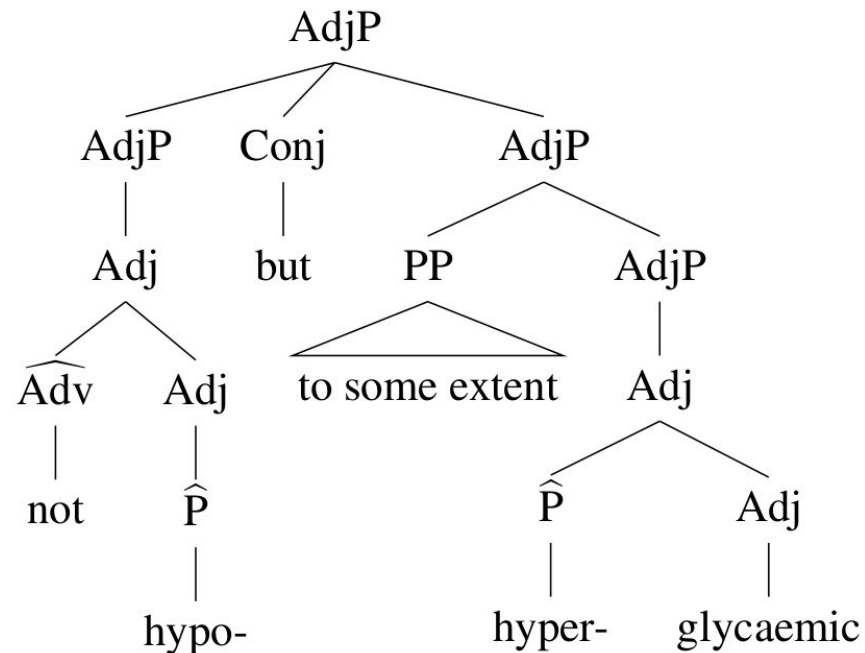   'Peter loaded and Maria unloaded the truck.'

14. **French**
    a. ?*[mono- et tri]-syllabe

    b. [Gesturing to a companion's empty wine glass] Re?

More complex examples

15. **Complex cases**
    a. totally pro- rather than simply not anti-Corbyn
    b. pre- as opposed to merely post-war
    c. not hypo- but to some extent hyper-glycaemic

Full lexicalisation

For some speakers, some prefixes (esp. *pre-/post-* and *pro-/anti-*) have become fully lexicalised

16. **Coordination with full P**
   a. Rival protests held in Croatia <u>for and anti</u>-abortion. ([www](#))

   b. The main storyline is there's an opportunity for a new gas line but there's a lot of local division between those who are <u>pro and against</u> it. ([www](#))

   c. Conservative MP Fiona Bruce asked the prime minister to give reassurance that the party supports the "highest standards of animal welfare" <u>before and post</u>-Brexit. ([www](#))

   d. a light but effective dual purpose oil which can be used as both a <u>pre and after</u> wax treatment ([www](#))
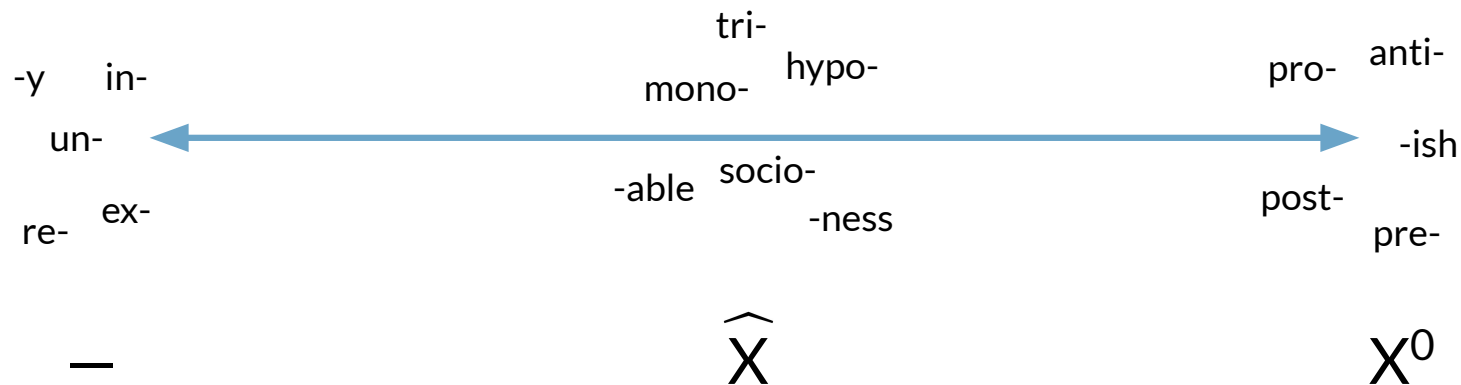
17. **Standalone uses**
   What's your stance on fox hunting?
   a. Oh, I'm totally anti! Barbaric sport!
   b. These days I'm rather pro, actually.

18. **Attachment to phrases**
   a. pre-[founding of Rome]
   b. post-[Alex's graduation]
   c. I'm anti-[Heather getting with Josh].

A spectrum of (de)grammaticalisation

Conclusion

Conclusion

- Good reasons to maintain LIH.

- But if we do, then we need an account of the apparent exceptions.
  - Compounds (Lowe 2015)
  - Coordination of prefixes
  - Derivational affixes on phrases (modification of Spencer 2005a)
  - Inflectional affixes on phrases …?

- In this view, there *are* no exceptions to LI: apparent exceptions are "analysed away".

- Is this question begging?
  - If it walks like a duck and quacks like a duck …

- Given the formal and statistical arguments, the burden of proof is with the opponents of LIH.

Conclusion

- This is not to say that LIH-less approaches must be wrong!

- Rather, they must find ways to ensure the relevant facts are captured in the other direction (e.g. by resisting "aggressive decomposition").

- But this does make them less perspicuous.
    - LI is the norm – exceptions should be explained exceptionally.

- The intermediary status of non-projecting words does this, by describing a syntactic atom which is nonetheless syntactically deficient.

Questions?

References

- Arnold, Doug & Louisa Sadler. 2013. Displaced dependent constructions. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG'13 Conference*, 48–68. CSLI Publications.

- Asudeh, Ash, Mary Dalrymple, and Ida Toivonen. 2013. Constructions with Lexical Integrity. *Journal of Language Modelling* 1(1). 1–54.

- Bruening, Benjamin. 2018. The lexicalist hypothesis: both wrong and superfluous. *Language* 94(1). 1–42.

- Cappelle, Bert. 2022. Lexical Integrity: a mere construct or more a construction? *Yearbook of the German Cognitive Linguistics Association* 10(1). 183–216.

- Carnie, Andrew. 2000. On the definition of X^0 and XP. *Syntax* 3(2). 59–106.

- Lowe, John J. 2015. The syntax of Sanskrit compounds. *Language* 91(3). e71–e115

- Norde, Muriel. 2010. Degrammaticalization: three common controversies. In Katerina Stathi et al. (eds.), *Grammaticalization: current views and issues*, 123–150. John Benjamins.

- Siegel, Dorothy C. 1974. *Topics in English morphology*. Ph.D. thesis, MIT.

- Simpson, Jane. 1991. *Warlpiri morpho-syntax: a lexicalist approach*. Kluwer Academic Publishers.
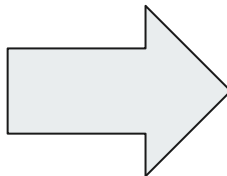
References

- Spencer, Andrew. 2005a. Case in Hindi, in Miriam Butt and Tracy Holloway King, editors, *Proceedings of the LFG'05 Conference*, 429–446. CSLI Publications.

- Spencer, Andrew. 2005b. Word formation and syntax. In Pavol Štekauer & Rochelle Lieber (eds.), *Handbook of word-formation*, 73–97. Springer.

- Strauss, Steven L. 1982. *Lexicalist phonology of English and German*. Foris.

- Toivonen, Ida. 2003. *Non-projecting words: a case study of Swedish particles*. Kluwer Academic Publishers.

- Williams, Edwin. 2007. Dumping lexicalism. In Gillian Ramchand & Charles Reiss (eds.), *The Oxford handbook of linguistic interfaces*, 353–382. Oxford University Press.

More data

9. **Coordination of stems** (cf. Spencer 2005b: 82)
   a. [write- or print]-able
   b. *[irrepair- or irreplace]-able

   c. [mouse- or rat]-like
   d. (both) [cake- and bread]-like

   e. *[milk- or cream]-y
   f. [cat and mouse]-y

Only words are good hosts

More elliptical examples

1. Paradigm shift [from anti-] [to pro-establishment] shocks Bitcoin community. (www)

2. Many critically ill patients [. . .] will undergo transition from [early pro-] to [later anti-inflammatory] phenotypes. (www)