L2 Lexical Patterns through the Lens of a Word Association Task

Svetlana Vetchinnikova

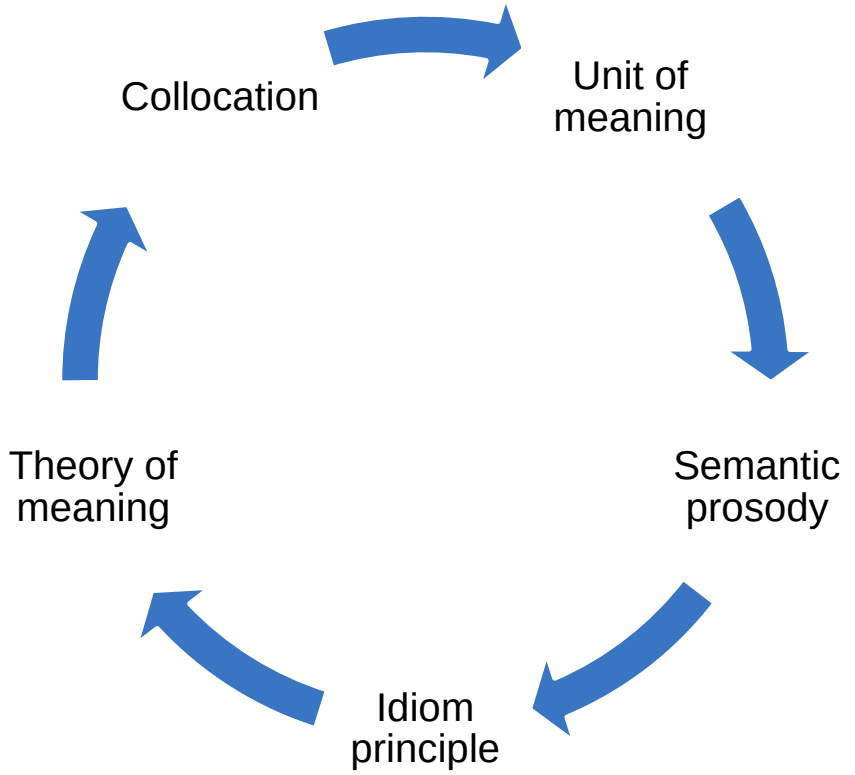
GlobE, 13-14 June 2012, Helsinki

Lexico-grammatical patterns

- chunks, routines, fixed expressions, fossilized forms, lexical phrases, lexicalized sentence stems, irregular phrases, formulaic sequences, collocations, etc.

- Formulaic Language (Wray)
- Pattern Grammar (Hunston & Francis)
- Construction Grammar
- Linear Unit Grammar (Sinclair & Mauranen)
- Lexical Priming (Hoey)
- Sinclair: idiom principle, unit of meaning...

Sinclair: a system of concepts



A unit of meaning is:

- a type of a MWU,
- an **independent** lexical item produced on the idiom prinicple, i.e. as a single choice,
- a meaning-shift unit.

“Core” meaning, meaning shift & a continuum of delexicalisation

- *hand* ⟹
- *on the one/other hand* ⟹
- meaning-shift and delexicalisation
- *hard work → on the one/other hand → kick the bucket*

Independent and dependent uses of a word (Sinclair 1987/2004)

- Besides, his **hands** were trembling.
- ‘I put my **hand** in my pocket to get my cigarettes.’
- His **hands** gripped and whitened.
- But Dot knew he was a doctor because his **hands** smelled clean and soapy.

- ‘You're just in time to give me a **hand** .’
- Professional sculptors on **hand** to help advise.
- We shook **hands** warmly;
- ‘Have any of you ever caught a bear with your bare **hands** ?’

[BNC]

The model of a unit of meaning

VARIABILITY

Unit of meaning (abstracted form-meaning pairing):

- obligatory components
  - *the core*: invariable formal component
  - *semantic prosody*: communicative purpose
- optional components
  - *collocation*: verbatim association
  - *colligation*: association with a gram. class
  - *semantic preference*: association with a sem. set

Example: *naked eye* (Sinclair 1996/2004)

	semantic preference	colligation	collocation			
			co-occurring word	co-occurring word	origin	
It is (not) *It can (not) be*	*visible* *obvious* *discernible* *unnoticed* *seen*	*by* *with* *via*	*the*	*naked*	*eye*	etc.

Semantic prosody of ”difficulty”.

Multi-word units in L2: “errors”, “deviations”

- **make a great job (did a great job);*
- **on the meantime (in);*

Yorio 1989

- **reach an aim (achieve)*
- **make (a) career (have)*

Nesselhauf 2005

errors → operating on the open-choice principle → idiom principle is not available → lack of sensitivity to MWUs

The study: data

1. Corpus data

- 5 students – 5 Master's theses
- longitudinally collected drafts
- drafts form one student = one corpus
- 5 "individual corpora"

2. Word association data

- each student: 3-6 WATs
- 1WAT: 100-120 stimulus words
- + retrospective comments on WA responses

Methods: Concgram & n-gram

N-gram, AntConc 2-5-grams, ≥ 3times	Concgram, Chris Greaves 2-word concgrams Cut offs: T score=2 MI value = 3
deal with (3) process of (3)	undergo/change
In the following I will (10)	following/I (11) following/will (12)
systematic changes (3)	systematic changes (6)

Methods:
Word Association Task

Categorization of responses:

Meaning-based responses:

beginning → end

hard → easy

contribution → add: “if you make a contribution, you add”,

Syntagmatic responses:

each → other

case → study

+ Form-based responses:

principle → principal

Concgrams/n-grams vs. WA responses

Concgrams/n-grams	WA responses
sound/correspondences	correspondences → sound
sake/for	sake → reason
basic/vocabulary	basic → meaning
deal with (3)	dealing → handling

Hypothesis

Meaningfully associated components of units of meaning are also psychologically associated.

Lexical usage patterns displayed by corpus data will be reproduced in word association responses.

Psycholinguistic reality of the model: collocation

- *comparative – method,*
- *orthographic – similarity,*
- *training – data,*
- *trial – excavation,*
- *goods – grave (grave goods),*
- *affairs – foreign (foreign affairs),*
- *decision – maker,*
- *sentinel – survey*
- *errors - sampling (sampling errors),*
- *heavy – metal,*
- *shared – space,*
- *based – on,*
- *derive – from,*
- *caused – by,*
- *concentrate – on,*
- *descend – from,*
- *access – to,*
- *despite – fact (despite the fact),*
- *revolves – around,*
- *arise – from*

Psycholinguistic reality of the model: colligation

difficult – to do

1 Sweden since 1323 (chapter 3.3). It is, however, **difficult *to*** know if Catholic Lutheran or Orthodox
2 of what religious should be. The term is **difficult *to*** define because it concerns things that are
3 dug through the grave. In such cases it has been **difficult *to*** determine what finds belong to the filling and
4 The occurrence of wooden constructions was **difficult *to*** sort out because of the lack of documentation
...
8 This overlapping of burial rituals makes it **difficult *to*** distinguish between the two inhumation burial

concerning – this

1 of the data influence the decisions I took **concerning** the architecture of the implementation. A finite
2 nd archaeological sources give only hints **concerning** the contact and relation between the Sumeriann
3 vided, and since I will not get into details **concerning** the etymology of the words, I will make a

Psycholinguistic reality of the model: semantic preference

lasted – time

1 reference]. The investigations in Luistari have **last**ed between 1969 and 1992 and the excavations
2 d between 1969 and 1992 and the excavations **last**ed for fourteen years. A series of publications of
3 ways [reference]. The excavations **last**ed for 6 weeks and 60 graves were found whereof 52

captures – variation

1 so that the rules that cause the changes can be **captured** and identified. The identified sound changes
2 uch "noise" in the data can be relatively easily **captured** computationally, and that actually is of the
3 logical mappings in the lexicon. The rules that **captures** the stem change of verbs like spy could look
4 in plural nouns. Morphophonological rules can **capture** the stem change of, for example, the verb spy

The relationship between usage patterns and WA responses

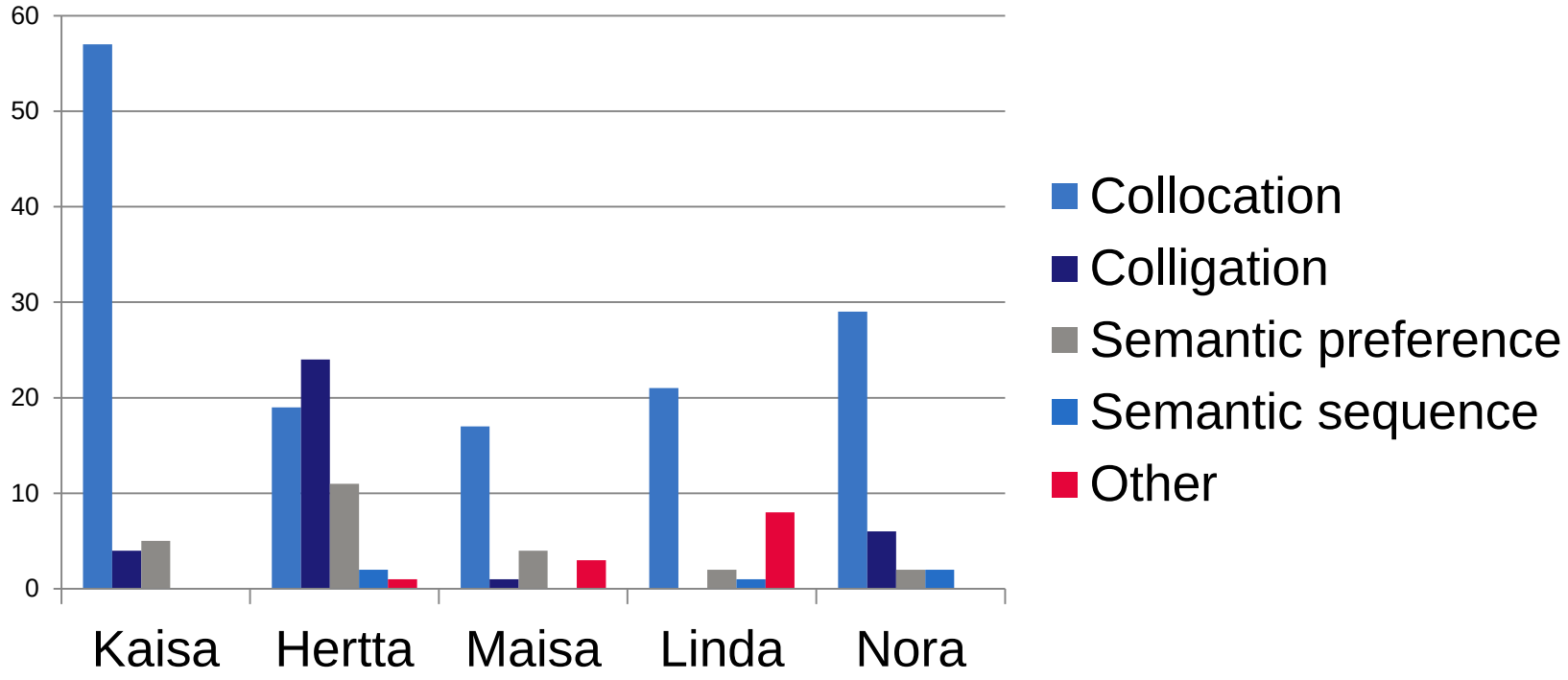
2x2 contingency table

Corpus / WA	Dependent	Independent	Total
Dependent (S)	44	0	44
Independent (M)	31	24	55
Total	75	24	99

Fisher's exact test: $p < 0.0001$

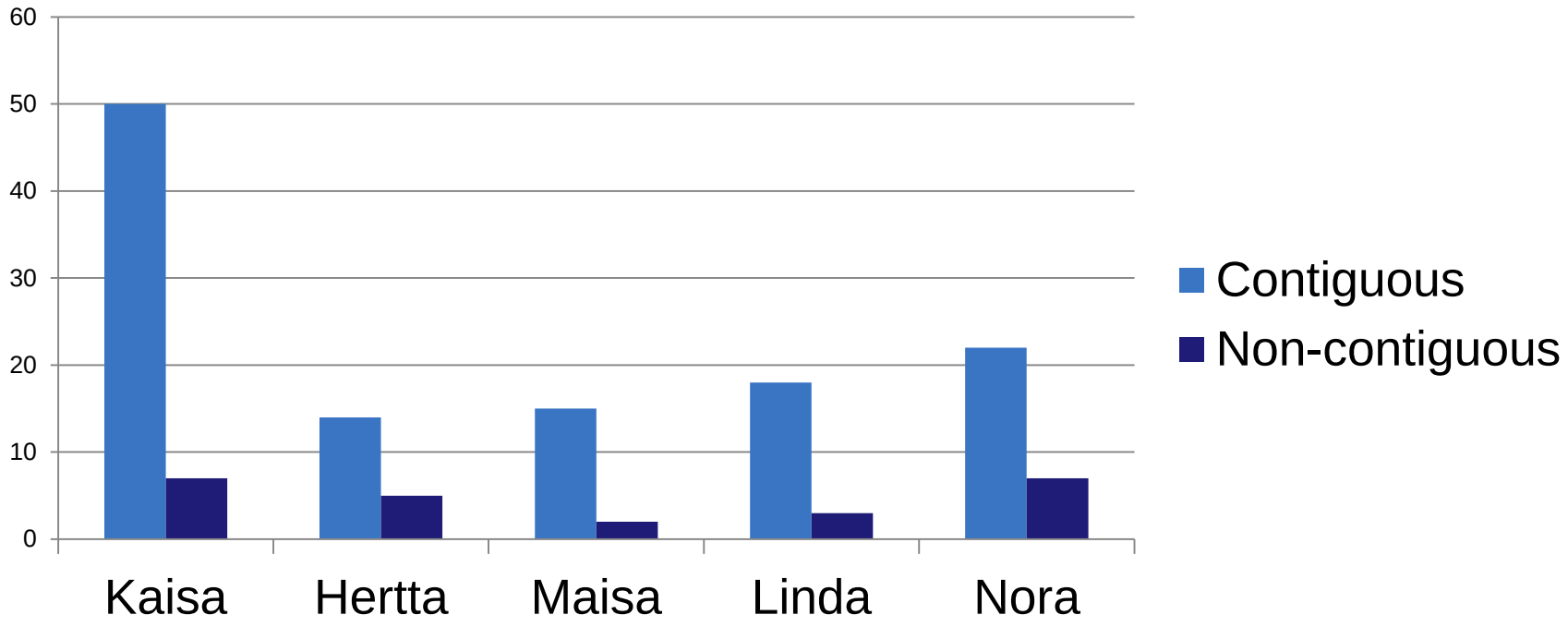
5 data sets: for 4, $p < 0.0001$, for 1, $p = 0.0008$

Collocational associaiton is stronger
(verbatim association between 2 or more words)



Types of syntagmatic responses in the Matching MWU category.

Contiguity and the strength of representation



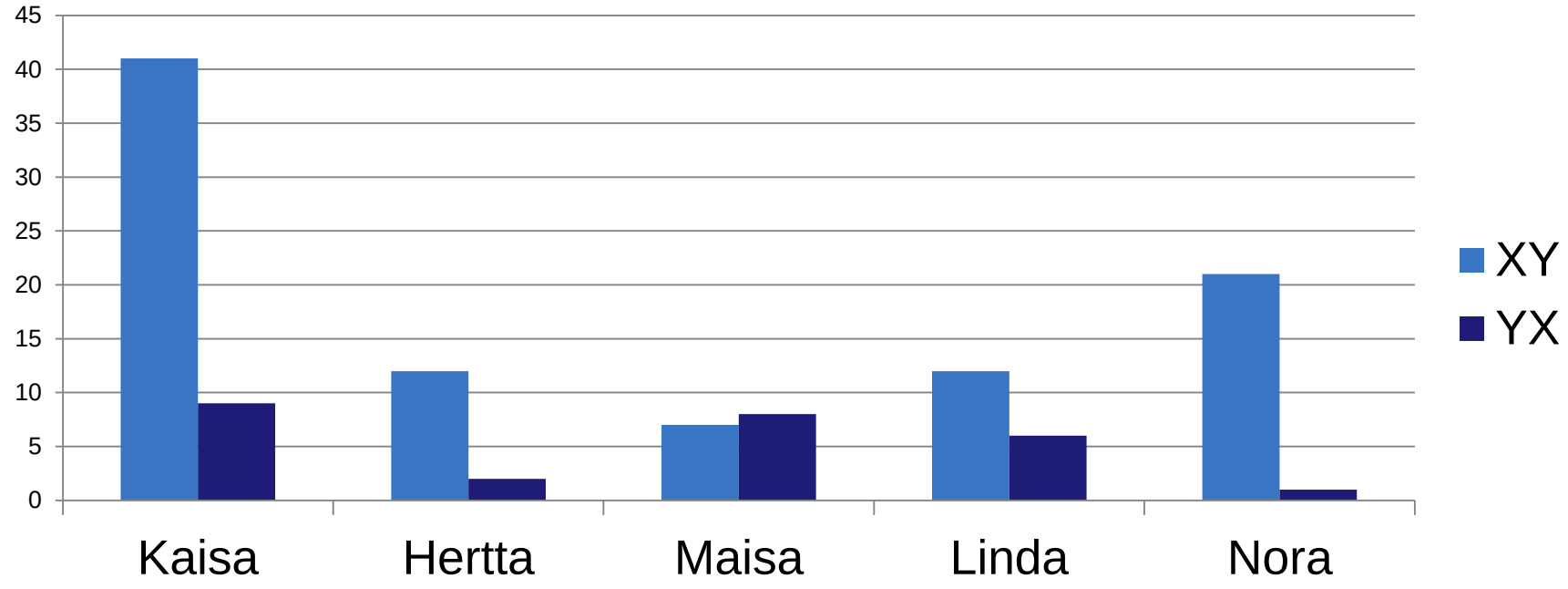
The proportions of contiguous and non-contiguous collocations in collocational Matching MWU S-responses.

Direction of association: XY vsYX

- sentinel – survey (27)
- methods – section (1)
- spectrum – software (3)
- census – data (8)

- survey – questions (N96, Matching MWU S, other)
- section - a part (N100, Non-matching, M)
- software – computers (N51, Non-matching, M)
- data – figures (N27, excluded: N4 demographic – figures)

Direction of association



The direction of syntagmatic association in contiguous collocations.

Core meaning effect

number
– digit

1 ations in the original Wiki concept, TWiki adds a **number** of features that make it suitable for e.g.
2 and allows for rapid comparison of a large **number** of languages. Bouchard developed a
3 several other languages, and there have been a **number** of prevailing theories that have first generally
4 non-Assyrologists have come up with a growing **number** of Sumerian etymologies and found more
5 n languages is to systematically present a large **number** of words found in the languages' basic

a + adj + number of + NP pl.:

collocation: *a number of*
colligation: with NP pl, adj.
sem. preference: for an adj. evaluating quanity

Core meaning effect

- *addition → subtraction: in addition to*
- *around → surround: around B.C.*
- *bears → forest: BEAR in mind*
- *hand → leg: on the one/other hand*
- *hard → easy: It BE hard/harder to-inf*
- *order → realization: in order to*
- *find → catch: find out*
- *latest → new: at the latest*
- *rest → here: the rest of*

Does syntagmatic association develop only inside a unit of meaning?

- X and Y:
  - *reliability and validity*
  - *reputation and image*
  - *ideology and issue (positions)*
  - *operative and strategic (communication planning)*
- Semantic sequences (Hunston)
  - *supports – idea* (“somebody supports an idea so I wrote an idea”)
  - *comment - say* (“and then you say something”)

not as an entertainer. The next **comment** illustrates this: “Her mind is that ‘I do my work

Does syntagmatic association develop only inside a unit of meaning?

- *survey - questions* (“survey includes questions”)

households (9 804 women and 3 915 men). The **survey** includes ***question***s about every child born to a

- *private - intimate* (“It was actually a sentence that I deleted yesterday like maybe half an hour before I had to have it in, I don’t like that sentence, well, I had to delete something and deleted a sentence where a guy said that private information is also very intimate information I was like I don’t really know what you mean so I deleted, so I guess it is stuck in my head, that sentence”)

An image provides detailed and personal information, and is therefore very intimate.

Conclusions

It seems that:

- The different kinds of associative links inside the unit are psycholinguistically real.
- Sinclair's hypothesis about the effect of the core meaning is supported by word association data.
- Syntagmatic association operates *inside* a unit of meaning.
- Spreading activation may depend on the proximity and the direction of the association.

References

Anthony, Laurence. 2007. *AntConc(Version 3.2.1w)* [Software]. Available from http://www.antlab.sci.waseda.ac.jp/software.html

Cheng, Winnie, Chris Greaves, John M. Sinclair, & Martin Warren. 2008. Uncovering the extent of the phraseological tendency: Towards a systematic analysis of concgrams. *Applied Linguistics 30/2*. 236–252.

Greaves, Chris. 2009 *ConcGram 1.0: A phraseological search engine*. Amsterdam: John Benjamins.

Sinclair, John M. 1987. Collocation: a progress report. In Ross Steele & Terry Treadgold (eds.), *Language topics: Essays in honour of Michael Halliday,* 319-331. Amsterdam: John Benjamins.

Sinclair, John M. 2004. *Trust the text*. London: Routledge.

Thank you!

(svetlana.vetchinnikova@helsinki.fi)