Qualico2018
International Quantitative Linguistics Conference
July 5-8, Wroclaw, Poland

TOWARDS A SYNERGETIC APPROACH TO WORD ORDER

Gertraud Fenk-Oczlon

Alpen-Adria Universität, Klagenfurt, Austria

Outline

- ***basic word order***
  - prevalence of the subject initial word orders SOV and SVO
  - SOV/SVO variation and some proposed explanations
- ***synergetic linguistics/systemic typology***
- ***cognitive constraints on word order***
  - the constant flow of linguistic information
    - the more predictable first
  - working memory constraints on the length of predicate argument structures
- ***SVO/SOV variation: interactions with word length, morphological type***, etc.

Basic word order

- basic ordering of the subject (S), verb (V), and object (O) in declarative sentences

- possible orders of clause constituents
  - SOV, SVO, VSO, VOS, OVS, OSV

- SVO and SOV are most frequent
- VSO is moderately frequent
- all other word orders are rare

Some proposed explanations for subject first preference & SOV/SVO variation

- Jackendoff (2002) Subject first preference = Agent first preference
- Gibson et al. (2013) The noisy-channel hypothesis predicts a shift from the default SOV order to SVO order for semantically reversible events
- Schouwstra & de Swart (2014) SOV/SVO variation depends on the meaning of the verb (in gesture tasks)
  - intensional verbs like *think lead to SVO*
  - extensional verbs like *throw* lead to SOV

Dependency length minimization

- Dependency lengths concern the distances between syntactically related words. Hypothesized principle: The distance between linguistic heads and dependents is minimized
- Ferrer-i-Cancho (2004) "*Euclidean distance minimization*" hypothesis for the distance between syntactically linked words
- Futrell et al. (2015) provide cross-linguistic evidence (in 37 languages) for a universal syntactic property of languages: dependency lengths are shorter than chance
- Liu et al. (2017) Short dependency distance, may be explained in terms of general cognitive constraint of limited working memory

Principle of predictability maximization

- Ferrer-i-Cancho (2017):
- Principle of predictability maximization/ uncertainty minimization

  "the verb (or the head in general) should be postponed and eventually placed last to maximize its predictability"

- the Principle of dependency length minimization is in conflict with the Principle of predictability maximization

Synergetic Linguistics

- Köhler (1986) The central axiom of synergetic linguistics is that language systems possess self-regulating and self-organizing mechanisms
- “Like other self-organising systems, language is characterised by the presence of cooperative and competitive processes which, together with the external forces of biology, psychology, physics, the social system and others, form the dynamics of the system.” (Köhler 1993)

Systemic Typology:
Interactions between the sub-systems of language

- *Systemic Typology* (Fenk-Oczlon & Fenk 1999): each language goes through self-organizing processes **optimizing the interactions between its subsystems** (phonological, morphological, and syntactical), and the **interaction** with its 'natural' environment, e.g. the **cognitive system**
  - At any point of time the constraints of our cognitive mechanisms are constraints on diachronic development, typological differentiation, and cross-linguistic variation

Cognitive constraints on word order

- hypothesis of a constant flow of linguistic information
- memory constraints

The constant flow of linguistic information

- In an effective and economical communication system, the information transmitted should be **distributed as uniformly** as possible across small time spans, and the average level of information transmitted per time should **not exceed capacity limits** (Fenk & Fenk 1980: 402f)

- Jäger's Uniform Information Density hypothesis also "predicts that language production is affected by a preference **to distribute information uniformly** across the linguistic signal"

  "UID predicts that speakers aim to transmit information uniformly close to, **but not exceeding, the channel capacity"** (Jäger 2010: 26f)

The constant flow of linguistic information

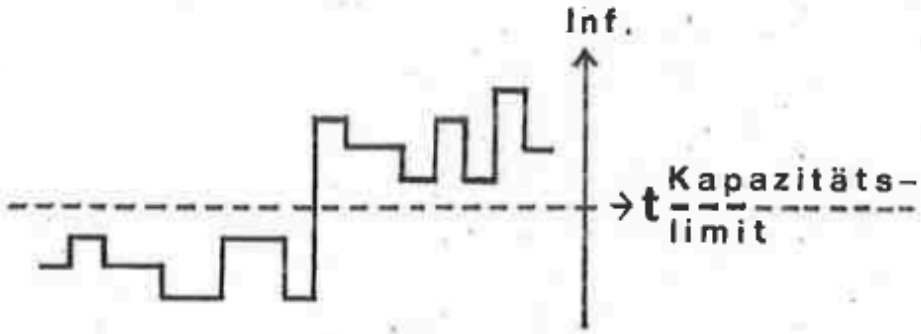

Abb. 1

Schema eines kapazitätsüberfordernden und unökonomischen Informationsflusses

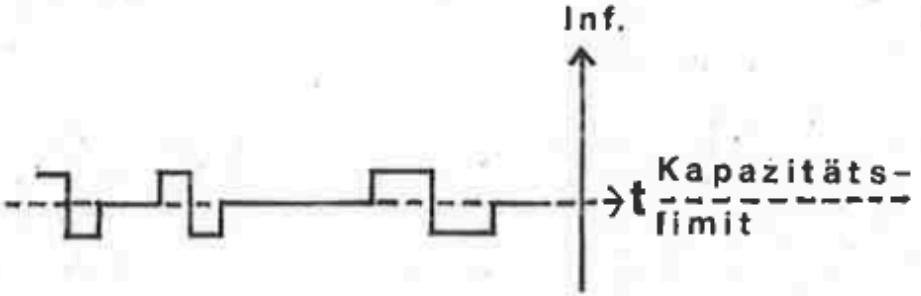

Abb. 2

Schema eines ökonomischen und der Kapazität besser angepaßten Informationsflusses

The constant flow of linguistic information

We propose two main mechanisms that contribute to a rather constant flow of linguistic information

1. the more predictable, the shorter

   our 1980 study revealed a proportionality function between word information (in bits) and word length (in number of syllables) in nine languages

2. the more predictable first

The constant flow of linguistic information
"The more predictable first"

- the first positions of sequences are associated with the lowest predictability or highest information (Shannon 1951)

  $\rightarrow$ **to place elements conveying a high amount of information at the beginning of a sequence, would produce peaks of cognitive overload**

- an appropriate strategy to avoid such informational peaks is to begin a sentence with those words having a higher predictability in this context, e.g. with (group) of words referring to (group) of words of the preceding sentence, and with terms coding concepts activated by this preceding sentence

Fenk-Oczlon (1983, 1989, 2001)

“the more predictable first”
old before new

- what has already appeared in the preceding discourse, i.e. what is *old* and *familiar* in the textual or situational context carries
- less subjective information than a new element in the same context. In this context it is more expectable, its analysis requires fewer cognitive costs

Fenk-Oczlon (1983, 1989, 2001)

“the more predictable first ”
topic before comment

“The distinction between **old** and **new** information is the principal phenomenon which underlies discussions of what have been called **topic** and **comment**, or theme and rheme” (Chafe 1970)

subject first word orders
"the more predictable first"

a possible explanation for the prevalence of SOV and SVO word orders across languages:

- subjects are highly discourse prominent and most frequently **topic (old information)** and therefore **more predictable** than verbs or objects which are prototypically **comments** and **new information**

Fenk-Oczlon 1983a

The constant information flow and SOV/SVO variation

- It was argued that SOV word order is preferred in agglutinative languages, because of their tendency to have very long verb forms containing much grammatical information (Fenk-Oczlon 1983). Extremely long verb forms are also characteristic for polysynthetic languages (many native American languages) which also tend to SOV
- All prototypical **agglutinative** and/or **polysynthetic** languages in our sample are **SOV** (Basque, Hopi, Japanese, Korean, Marrangu, Navaho, Telugu, Turkish)
- Languages with **isolating** or **fusional** morphology and therefore **shorter verb forms** tend to **SVO** or **VSO** word order. (e.g. Thai, Mandarin, English, Hawaiian)

- Agglutinative languages place the verbal modifiers e.g. for negation, causation, and reflexive or reciprocal action after the verb root which results in very long verb forms (Lehmann 1973)
- Nominal modifiers such as relative, adjectival, and genitival expressions precede nouns in SOV languages

The constant information flow and SOV/SVO variation

- Placing long and **less predictable verbs** late, conforms to the constant flow of linguistic information hypothesis and to Ferrer-i-Cancho's (2017) principle of predictability maximization. It is, moreover, in line with the strategy of putting **heavy constituents late** (Behagel 1909; Arnold et al., 2000), since less predictable units tend to have longer forms (Zipf 1935; Fenk-Oczlon 2001)

The constant information flow and SVO/SOV variation in sign languages

- the unmarked word order in sign languages seems to be SOV (e.g. Goldin-Meadow et al. 2008)
- in sign languages verbs tend to be longer than nouns on average (e.g. Hunger 2006; Johnston 2012)

  $\rightarrow$ SOV should be the preferred word order
- open questions:
  - **do more predictable and shorter verbs lead to SVO in sign languages?**
  - also in improvised gesture tasks?

How long are predicate-argument structures?

- **Prototypical distance between heads and dependends** shows in simple declarative sentences encoding one predicate-argument structure (S, O, V)

  - simple declarative sentences seem to be universal from a syntactic perspective (quite a number of languages use almost exclusivly a series of minimal-predications instead of more complex sentences (Sasse 1991; Heeschen 1994)

Memory constraints: How long are predicate-argument structures?

- **General assumption**:

  Natural languages have developed in adaptation to cognitive functions and constraints, e.g. working memory limitations
  - temporal span : Baddeley's (1986) phonological loop ~2sec; Fraisse's (1957)
  - limited capacities: 7 ± 2 (Miller 1956); 4 ± 1 (Cowan 2001)

- **Hypothesis**: (Fenk-Oczlon 1983)

  The number of syllables per simple clause (predicate-argument structure) will vary within the range of Miller's magical number seven plus or minus two

.

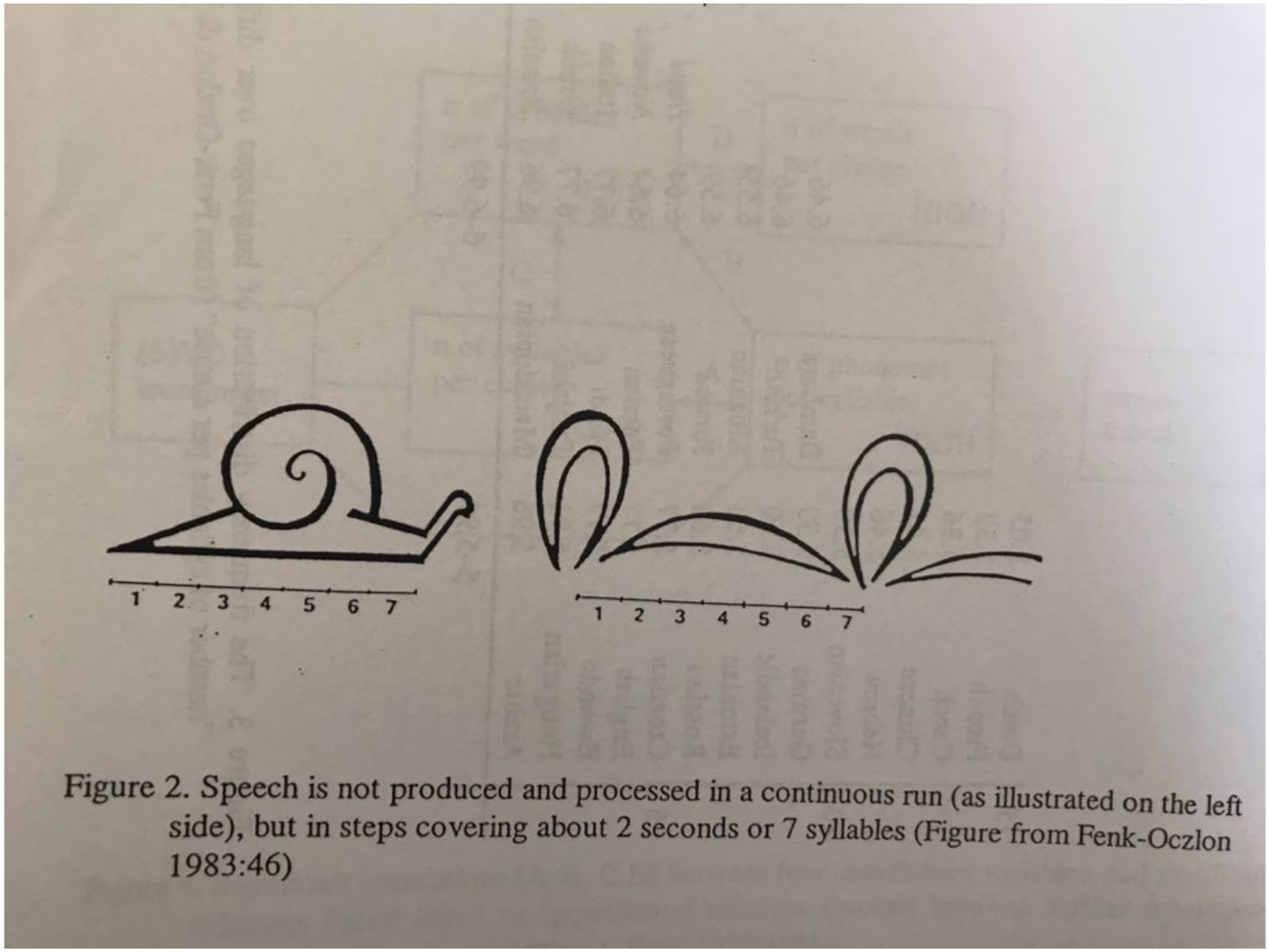

Figure 2. Speech is not produced and processed in a continuous run (as illustrated on the left side), but in steps covering about 2 seconds or 7 syllables (Figure from Fenk-Oczlon 1983:46)

Method

- Native speakers of 51 languages from all continents (19 Indo-European, 32 Non-Indo-European) were asked to translate a matched set of 22 simple declarative sentences encoding one predicate-argument structure
- Subjects were furthermore instructed to read their translations in normal speech and to count the number of syllables and the number of words

Results

In our sample of 51 languages

- **the mean number of syllables per simple clause is 7.02**, ranging from 4.64 in Thai up to 10.96 in Telugu. The mean number of phonemes per syllable is 2.24, ranging from 2.79 in German to 1.76 in Hawaiian

- The mean number of **words** is **3.62**, ranging from 2.5 in Arabic to 4.4 in English

The mean number of syllables per clause in 51 languages

		Dutch	5.05										
		Czech	5.36										
		Mandarin	5.46										
		Sloven.	5.50	Canton.	6.05								
		German	5.50	Cham	6.32								
		Iceland.	5.59	Bambara	6.46								
		French	5.64	Turkish	6.46								
		Estonian	5.68	Albanian	6.55								
		Russian	5.68	Yoruba	6.59								
		Croatian	5.77	Portug.	6.64								
		English	5.77	Persian	6.64								
		Ewondo	5.77	Finnish	6.73	Hopi	7.12	Kadazan	8.00	Kirundi	9.05		
		Hungar.	5.91	Fukien	6.77	Navaho	7.41	Korean	8.18	Chiquit.	9.14	Hawaiian	10.05
Thai	4.64	Hebrew	5.96	Hindi	6.77	Italian	7.50	Annang	8.23	Javanese	9.18	Roviana	10.05
Vietn.	4.91	Arabic	5.96	Panjabi	6.77	Greek	7.55	Basque	8.27	Indones.	9.46	Japanese	10.23
Lamso	4.96	Polish	5.96	Macedon.	6.96	Spanish	7.96	Tagalog	8.95	Marranju	9.86	Telugu	10.96
4 – 4.99		5 – 5.99		6 – 6.99		7 – 7.99		8 – 8.99		9 – 9.99		10 – 10.99	

Figure 1: The frequency distribution of 51 languages over seven classes of the parameter mean n of syllables per clause

from: Fenk-Oczlon & Fenk 2010:1538

The more syllables per clause, the fewer phonemes per syllable?

Hypothesis:

A negative correlation between the number of phonemes per syllable and the number of syllables per clause.

The result:

r = – 0.73 (sign. p < 0. 01)

→ all relevant working memory constraints discussed in the literature, show in the length of predicate-argument structures: a mean of 7 syllables (Miller‘s 7 ±2), a mean of 4 words (Cowan‘s 4 ±1) and a duration of ~2 sec (10 short syllables or 5 complex syllables correspond to about 2 sec)

Fenk-Oczlon & Fenk (1985; 2001; 2009)

Further significant crosslinguistic correlations

- The more syllables per word, the fewer phonemes per syllable
- The more syllables per clause, the more syllables per word.
- The more words per clause, the fewer syllables per word.
- Low syllable complexity is significantly associated with **SOV order**

Fenk & Fenk-Oczlon, 1993;
Fenk-Oczlon & Fenk 1999

Trade-offs between number of syllables and number of phonemes – self-organizing processes

- **time limits on clause length – time limits regarding memory constraints as well as the breath-cycle – force a trade-off between the length of syllables in number of phonemes and the length of clauses in the number of syllables**

- thus, the segmentation of natural languages and especially the limited size of clauses can be viewed as the result of **self-organizing processes** accounting for such physiological and cognitive constraints

Associations between SVO/SOV word order and other linguistic features

in 34 languages

SVO	SOV
low number of syllables per word	high number of syllables per word
high number of phonemes per syllable	low number of phonemes per syllable
low number of syllables per clause	high number of syllables per clause
high number of words per clause	low number of words per clause
low number of morphological cases	high number of morphological cases
isolating or fusional morphology	agglutinative morphology
prepositions	postpositions
cumulative case exponents	separatist case exponents
stress-timed	syllable-timed

Adapted from Fenk-Oczlon & Fenk, 2005

Conclusion

- we presented typological data showing that word order interacts with other linguistic subsystems in a systematic way
- we could demonstrate that languages have adapted to general cognitive constraints
  - the hypothesis of a *constant flow of linguistic information may* explain the cross-linguistic prevalence of subject first word orders as well as SOV/SVO variation → predictability maximization
  - all relevant working memory constraints discussed in the literature, show in the length of simple predicate-argument structures → dependency length minimization

→ languages tend to keep the **size of clauses and the information flow within these clauses rather constant**

Conclusion

- we propose that a systemic or synergetic approach to word order will advance our understanding of word order evolution and word order variation

Thank you!

References

- Arnold, J., Wasow, T., Losongco, A., & Ginstrom, R. (2000). Heaviness vs. newness: the effects of complexity and information structure on constituent ordering. *Language, 76*, 28-55
- Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press
- Behaghel, O. (1909). Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern. *Indogermanische Forschungen*, *25*,110–42
- Chafe, W. (1970) *Meaning and the Structure of Language*. Chicago: Chicago University Press
- Cowan, N. (2001). The magical number 4 in short-term memory. A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87-114
- Fenk, A. & Fenk, G. (1980). Konstanz im Kurzzeitgedächtnis – Konstanz im sprachlichen Informationsfluß? *Zeitschrift für experimentelle und angewandte Psychologie* 27, 400-414
- Fenk, A. & Fenk-Oczlon, G. (1993). Menzerath's Law and the Constant Flow of Linguistic Information. In R.Köhler & B.Rieger (eds.) *Contributions to Quantitative Linguistics,* Springer: Dordrecht
- Fenk-Oczlon, G. (1983). Ist die SVO-Wortfolge die „natürlichste"? *Papiere zur Linguistik* 29, 23-32
- Fenk-Oczlon, G. (1983a). *Bedeutungseinheiten und sprachliche Segmentierung. Eine sprachvergleichende Untersuchung über kognitive Determinanten der Kernsatzlänge*. Tübingen: Narr
- Fenk-Oczlon, G. (1989). Word frequency and word order in freezes. *Linguistics* 27, 517-556
- Fenk-Oczlon, G. (2001). Familiarity, information flow, and linguistic form. In J. Bybee & P. Hopper (eds.) *Frequency and the emergence of linguistic structure*, Amsterdam/Philadelphia: Benjamins, 431-448
- Fenk-Oczlon, G. & Fenk, A. (2001). What language tells us about immediate memory span. In K.W. Kallus, N. Posthumus & P. Jimenez (eds.) Current psychological research in Austria. Proceedings of the 4th scientific conference of the Austrian Psychological Society. Graz: Akademische Druck-u. Verlagsanstalt , 175-179

- Fenk-Oczlon, G. & Fenk, A. (2005). Crosslinguistic correlations between size of syllables, number of cases, and adposition order. In G. Fenk-Oczlon & C. Winkler (eds.), *Sprache und Natürlichkeit. Gedenkband für Willi Mayerthaler*. Tübingen: Narr, 75-86
- Fenk-Oczlon, G. & Fenk, A. (2009). Some parallels between language and music from a cognitive and evolutionary perspective. *Musicae Scientiae* 13, 201-226
- Fenk-Oczlon, G. & Fenk, A. (2010). Measuring basic tempo across languages and some implications for speech rhythm. *Proceedings of the 11th Annual Conference of the International Speech Communication Association* (INTERSPEECH 2010), Makuhari, Japan, 1537-1540
- Ferrer-i-Cancho, R. (2004). Euclidean distance between syntactically linked words. *Physical Review E 70, 056135*
- Ferrer-i-Cancho, R. (2017). The placement of the head that maximizes predictability. An information theoretic approach. *Glottometrics, 39*, 2017, 38-71
- Fraisse, P. (1957). *Psychologie du temps*. Paris: Presses Universitaires de France
- R. Futrell, R. Mahowald, K & Gibson, E. (2015) Large-scale evidence of dependency length minimization in 37 languages, *PNAS* 112 (33) (10336–10341
- Gibson, E., Piantadosi, S. T., Brink, K., Bergen, L., Lim, E., & Saxe, R. (2013). A noisy-channel account of crosslinguistic word-order variation. *Psychological Science,* 24 (7)
- Goldin-Meadow, S., So, W. C., Ozyurek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *PNAS* 105 (27), 9163–9168
- Hunger, B. (2006). Noun/Verb Pairs in Austrian Sign Language (ÖGS). *Sign Language & Linguistics, 9,* 71-94

- Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford University Press
- Johnston, T. (2012). Lexical frequency in sign languages. *Journal of Deaf Studies and Deaf Education* 17(2), 163–193
- Köhler, R. (1986). *Zur linguistischen Synergetik. Struktur und Dynamik der Lexik*. Bochum: Brockmeyer
- Köhler, R. (1993). Synergetic Linguistics In R. Köhler & B.B. Rieger (eds) *Contributions to Quantitative Linguistics,* Springer: Dordrecht
- Lehmann, W. P. (1973). A Structural Principle of Language and its Implications. *Language*, 49, 47-66
- Liu, H., Xu, C. & Liang, J. (2017). Dependency distance: a new perspective on syntactic patterns in natural languages, *Physics of Life Reviews 21*, 171-193
- Schouwstra, M. & de Swart, H. (2014). The semantic origins of word order. *Cognition* 131, 431–436