On the Linearity of Information Structure

Jon Scott Stevens

The Information-structural (IS) notions of Focus, Ground and Givenness determine constituent order in a variety of languages, and there is an apparent preference for Given or Grounded elements to precede Focused elements, except perhaps in cases of contrastive Focus. Using the terminology of Kiss (1998, 2007), "information focus" tends to occur at the right edge of a clause. Distinct syntactic operations are responsible for this linear order effect, and thus a single elegant syntactic explanation for these cross-linguistic facts is elusive. This paper suggests that this phenomenon is a long-term product of pressure to choose syntactic variants that have a particular linear order effect. Such a pressure could arise from processing and planning effects, though other factors are likely involved as well. I argue for the plausibility of this account by defending three premises: 1) right-edge Focus can facilitate processing and planning in cases where the Focus of a sentence is new information, 2) right-edge Focus variants will become preferred in usage, 3) usage preferences can affect language change. But first, I review some data which illustrate the problem.

First, it has been argued[13,15] that scrambling often has the effect of placing Focus at the right edge. Focused elements themselves typically only scramble when they are contrastively interpreted. This is shown in Dutch in (1). Also, Yiddish object shift has been analyzed as a form of scrambling[21] and shows the same Given-before-Focus effect, shown in (2).

(1) a. dat [$_{DP}$ alleen DIT boek] Jan Marie t_{DP} geeft.
that only THIS book John Mary gives.
b. dat [$_{DP}$ zo'n boek] alleen JAN Marie t_{DP} geeft.
that such-a book only JOHN Mary gives.
c. *dat [$_{DP}$ het boek] Jan Marie t_{DP} geeft.
that the book John Mary gives.
(Neeleman and van de Koot 2008, p.140)

(2) a. Ikh hob gekoyft dem bikhl. (canonical order)
I have bought the book
b. Ikh hob dem bikhl gekoyft. (*bikhl* must be topical)
I have the book bought
(Wallenberg 2010, p.29)

Similar effects have been reported for Hindi[18] and Finnish[9] as well. Apart from scrambling, long-distance movement to the left periphery pushes Focus to the right in Russian[1] and Finnish[20], where a special left-edge position is occupied by topical/Given elements, as in (3).

(3) Q: What do you think of the new president?

a. ??Narod ljubit novogo prezidenta. (canonical order)
people love new president

b. Novogo prezidenta narod ljubit.
new president people love

(Bailyn 2003, p.16)

Italian demonstrates a third syntactic phenomenon. In examples like the following, the subject is extraposed when it is narrowly Focused, i.e. in the answer to 'who spoke?'.

(4) Ha parlato [GIANNI]$_F$.
has spoken Gianni
Gianni has spoken.
(Belletti 2004, p.5)

Interestingly, the canonical SV order is only possible when the subject is contrastively interpreted[2]. So we see at least three behaviorally distinct syntactic operations responsible for the same linear order effect. As Büring (2011) points out, purely syntactic theories of this phenomenon are not sufficiently constrained in their possibilities. Büring argues that prosodic explanations[22] are promising, but points out that there is counter-evidence[12,14] to the claim that IS-correlated movement always affects prosody. Alternatively, one may take the "mapping" approach[19], where a distinct IS component of grammar imposes linear order preferences. I suggest this is not necessary. Instead, I suggest that these linear order effects are due to principles underlying choice from among syntactic variants.

First, I argue that there are cases where right-edge Focus facilitates processing. Consider examples of vagueness where certain elements (double underlined below) need context to supply a precise denotation. When these elements are discourse-new, placing them later in the sentence allows intra-sentential context to supply the necessary information to resolve the denotation immediately. When discourse-old, on the other hand, linear order makes no difference, because by definition the preceding context must have already supplied the full denotation[17]. Semantic processing is incremental and predictive[6,7,9] and hearers are quite sensitive to IS/prosodic notions like Givenness[7]. This should create a processing advantage for discourse-new context-dependent elements at the right edge of a clause. Such configurations correlate with the increased acceptability, as in the following.

(5) Pat: Fluffy's eating Charlie's eggs.
Mo: Give Fluffy a [small piece of BACON]$_{NEW}$ to distract him.
/ ?Give a [small piece of BACON]$_{NEW}$ to Fluffy to distract him.

(5') Pat: These last few pieces of bacon are pretty small.
Mo: Give a [small piece of bacon]$_{OLD}$ to FLUFFY.
/ Give FLUFFY a [small piece of bacon]$_{OLD}$.

The Focus of a sentence is not necessarily discourse-new, but a non-Focused or Given element is necessarily discourse-old. Any processing effects of this sort will favor right-edge Focus when it conveys new information. Right-edge new-information Focus may also be favored from a language planning perspective. Placing accessible, discourse-old elements first may allow for more efficient planning of the formulation of the context-sensitive meaning. To estimate the frequency of new-information Focus we can look to a corpus study[5] of IS effects on the English ditransitive alternation[3] which finds discourse-new Focus in 37% of cases.

Turning to the second premise, processing and planning advantages should translate to a usage preference. Under Grice's Maxim of Manner ("be clear"), we should expect speakers to avoid garden paths and other difficult-to-process utterances. Furthermore, even a selfish speaker will prefer the cognitively simpler utterance if it requires less planning effort. Because interlocutors tend to align their grammatical choices with each other[16], these cognitive preferences could ultimately yield a strong asymmetry in syntactic usage patterns, e.g. between (5) and (5'). While other factors must be at play, this creates a unidirectional pressure toward right-edge Focus constructions.

Finally, usage preferences have the effect of skewing the input to new generations of language learners. This can have long-term effects on a language. To illustrate, consider a simple learning model. Given a choice between syntactic variants A and B, learners adopt baseline probabilities that directly mirror the usage of the adults in their population. Then, the learners develop context-specific preferences and thereby become the new generation of adults. The new adults prefer one particular variant in a certain context, but in all other contexts they choose A or B based on the baseline probabilities they acquired. Their usage patterns become the input to a new generation, and so on, until a stable state is reached. Fig. 1 shows the results of a simulation that introduces B into an A-only grammar, where adults prefer whichever variant places Focus at the right edge when the Focus is discourse-new (an estimated 30% of the time). The graph plots the proportion of right-edge Focus variants for each generation of speakers.

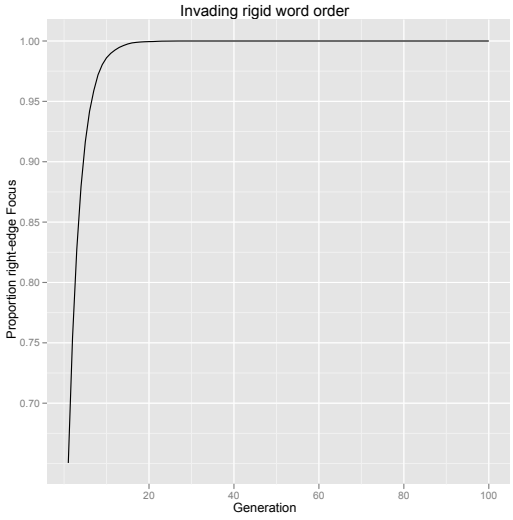

FIG. 1

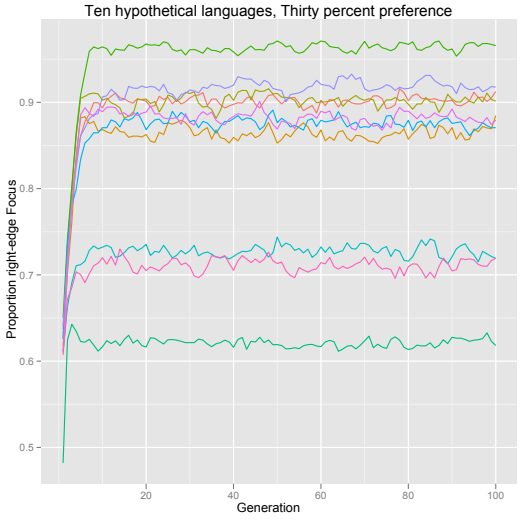

FIG. 2

This is an idealized simulation; of course other pressures will potentially counteract this effect. To simulate unpredictable social pressures, we can add a random (possibly negative) number to the 30% preference parameter. Fig. 2 shows the result of such a simulation, with ten simulated "languages", and with each generation of learners hearing 1000 adult utterances. The result looks like a typological tendency to exhibit the kind of IS-linear order correspondence discussed here.

This paper does not intend to provide a full account of how Information Structure and syntax interact. Rather, it intends to provide a plausible usage-based account of the phenomenon. More broadly, it intends to underline that we are not bound to look to the grammatical system itself to explain cross-linguistic patterns.

References

1. Bailyn, John F. Does Russian scrambling exist? In S. Kirimi (ed.), Word Order and Scrambling. Oxford: Blackwell. / 2. Belletti, Adriana. 2004. Aspects of the low IP area. In L. Rizzi (ed.), The Structure of CP and IP. Oxford: Oxford University Press. / 3. Bock, J. Kathryn and David Irwin. 1980. Syntactic effects of information availability in sentence production. Journal of Verbal Learning and Verbal Behavior 19, 467-484. / 4. Büring, Daniel. 2011. Syntax and Prosody, Syntax and Meaning. In M. Den Dikken (ed.), Cambridge Handbook of Generative Linguistics. / 5. Collins, Peter. 1995. The indirect object construction in English: an informational approach. Linguistics 33, 35-49. / 6. Cooper, Roger M. 1974. The control of eye fixation by the meaning of spoken language. Cognitive Psychology 6, 84-107. / 7. Dahan, Delphine, Michael Tanenhaus, and Craig Chambers. 2002. Accent and reference resolution in spoken-language comprehension. Journal of Memory and Language 47, 292-314. / 8. Grice, H.P. 1975. Logic and Conversation. In P. Cole and J. Morgan (eds.), Syntax and Semantics, vol. 3: Speech Acts. New York, NY: Academic Press. / 9. Kaiser, Elsi. 2002. The syntax-pragmatics interface of Finnish ditransitive verbs. In M. van Koppen, E. Thrift, E. van der Torre, and M. Zimmermann (eds.), Proceedings of ConSOLE IX. / 10. Kiss, Katalin. 1998. Identificational Focus versus Information Focus. Language 74, 245-273. / 11. Kiss, Katalin. 2007. Topic and focus: Two structural positions associated with logical functions in the left periphery of the Hungarian sentence. In C. Féry, G. Fanselow, and M. Krifka (eds.), Interdisciplinary Studies in Information Structure 6, 69-81. / 12. Kučerová, Ivona. 2007. The Syntax of Givenness. Ph.D thesis, MIT. / 13. Lenerz, Jürgen. 1977. Zur Abfolge nominaler Satzglieder im Deutschen. Tübingen: Gunter Narr Verlag. / 14. López, Luis. 2009. A Derivational Syntax for Information Structure. Oxford Studies in Theoretical Linguistics. Oxford University Press. / 15. Neeleman, Ad and Hans van de Koot. 2008. Dutch scrambling and the nature of discourse templates. Journal of Comparative Germanic Linguistics 11:2, 137-189. / 16. Pickering, Martin and Simon Garrod. 2004. Toward a mechanistic psychology of dialogue. Behavioral and Brain Sciences 27, 169-226. / 17. Stevens, Jon. 2012. Information Structure as Parallel Tree Building. In S. Müller (ed.), Proceedings of HPSG 2011. Stanford: CSLI. / 18. Tickoo, Asha. 1992. On preposing and word order rigidity. Pragmatics 2:4, 467-486. / 19. Vallduví, Enric. 1990. The Informational Component. Ph.D thesis, University of Pennsylvania. / 20. Vilkuna, Maria. 1995. Discourse Configurationality in Finnish. In K. Kiss (ed.), Discourse Configurational Languages. Oxford: Oxford University Press. / 21. Wallenberg, Joel. 2010. Antisymmetry and the Conservation of C-command: Scrambling and Phrase Structure in Synchronic and Diachronic Perspective. Ph.D thesis, University of Pennsylvania. / 22. Zubizarreta, Maria Luisa. 1998. Prosody, Focus, and Word Order. Cambridge, MA: MIT Press.