

Yang (2000):  
“Internal and external forces in language change”

Phillip Potamites

March 23, 2008

## 1 my questions

- Is the argument circular?
  - re. frequency
  - re. heterogeneity and innovation
- Does the argument only go half as far as it should?
  - re. parameters
  - re. competing/coexisting grammars
  - re. biological contribution to language (a.k.a. UG)
- Does the argument pin itself to untenable claims?
  - i.e. ‘biology doesn’t change’.
  - i.e. ‘change requires heterogeneity’.
  - i.e. ‘ascendent grammars must triumph’.

## 2 some quotes

(p. 233)

- “However, this discrete view of language learning is at odds with the facts of language development.”
- “In general, child language development is gradual, thus providing no support for the on/off view of parameter setting”.

(p. 237)

- “If we assume, as there is no reason not to, that the biological endowment of UG is held constant from generation to generation, we may conclude that the source for the discrepancy between two generations of speakers lies in the linguistic evidence: generation  $n$  and  $n+1$  are exposed to sufficiently different linguistic evidence and thus form different knowledge of language as a result.”
- “What makes the linguistic evidence for generation  $n+1$  different from that of the previous generation? There are many possibilities. Migration of foreign speakers may introduce novel expressions that were previously unseen; linguistic innovation may modify the linguistic evidence for the next generation; or social and cultural factors may influence the distributional patterns of linguistic expressions used in a population. These are interesting topics of research and are an integral part of a complete explanation of language change. However, they are not directly relevant to a formal model of language change. We are chiefly concerned with the predictive consequences of such changes: what happens to language learners when the linguistic evidence is altered, and how does it affect the composition of the linguistic population as a result? This is much like the population genetic theory of evolution, which concerns the predictable changes in the population once some new genotypes are introduced. The precise manner in which new genes arise, which could be mutation, migration, and so on, is a separate question that often contains too much contingency to give a firm answer.”

(p. 241)

- “We concluded that heterogeneity in the linguistic evidence, however introduced, is a prerequisite for language change. Once the homogeneity is punctured, language learners form internal representations of coexisting grammars.”

### 3 some formulas

(cf. Bush and Mosteller, 1951, 1958: “linear reward-penalty scheme”)

$$\begin{aligned} \text{if } G_i \rightarrow s \text{ then } & \begin{cases} p'_i = p_i + \gamma(1 - p_i) \\ p'_j = (1 - \gamma)p_j \end{cases} \\ \text{if } G_i \nrightarrow s \text{ then } & \begin{cases} p'_i = (1 - \gamma)p_j \\ p'_j = \frac{\gamma}{N-1} + (1 - \gamma)p_j \end{cases} \end{aligned}$$

$$\begin{aligned} p_i + \gamma(1 - p_i) &= p_i + \sum_j \gamma p_j, \quad \sum_j \gamma p_j = \gamma(1 - p_i) \\ \frac{\gamma}{N-1} + (1 - \gamma)p_j &= p_j + (\frac{\gamma}{N-1} - \gamma p_j), \quad \Delta_j = \frac{\gamma}{N-1} - \gamma p_j \\ \sum_j \Delta_j &= \gamma p_i = \gamma(1 - \sum_j p_j) = \sum_j (\frac{\gamma}{N-1} - \gamma p_j) \end{aligned}$$

penalty probability:

$$c_i = Pr(G_i \leftrightarrow s | s \in E) \tag{1}$$

asymptotic distributions (cf. Narendra and Thathacher 1989, pp. 162-165):

p.235: "For simplicity but without loss of generality, suppose that there are two grammars in the population..."

$$\lim_{t \rightarrow \infty} p_1(t) = \frac{c_2}{c_1 + c_2}, \quad \lim_{t \rightarrow \infty} p_2(t) = \frac{c_1}{c_1 + c_2} \quad \left( \frac{c_2 + c_3}{c_1 + c_2 + c_3} ? \right) \tag{2}$$

P. 238: "Suppose a proportion  $\alpha$  of  $G_1$  expressions are incompatible with  $G_2$  and a proportion  $\beta$  of  $G_2$  expressions are incompatible with  $G_1$  ... a proportion  $p$  of expressions can be viewed as generated by  $G_1$  and a proportion  $q$  of expressions can be viewed as generated by  $G_2$ ".

$$E_{G_1, G_2} = pG_1 + qG_2 \quad (\leftarrow \text{not real/scalar math?}), \quad c_1 = \beta q, \quad c_2 = \alpha p \tag{3}$$

$$p' = \frac{\alpha p}{\alpha p + \beta q}, \quad q' = \frac{\beta q}{\alpha p + \beta q}, \quad \frac{p'}{q'} = \frac{\alpha p}{\beta q} \quad (\text{i.e.} = \frac{c_2}{c_1}) \tag{4}$$

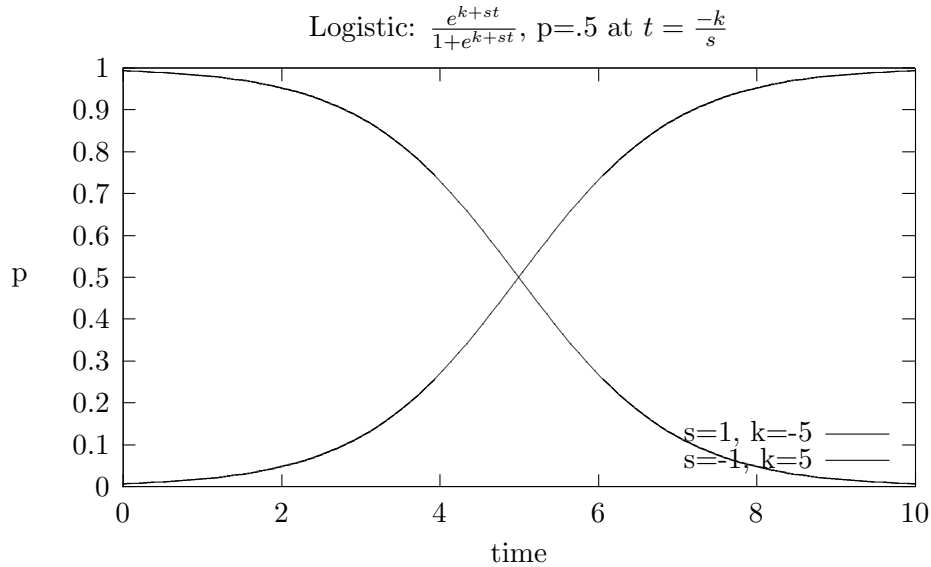
(did we just replace limits with iterations? because each generation is assumed to reach that limit?)

(p. 239, ex. 8,9)

- "The fundamental theorem of language change  
 $G_2$  overtakes  $G_1$  if  $\beta > \alpha$ : the advantage of  $G_2$  is greater than that of  $G_1$ "
- "Corollary: Once a grammar is on the rise, it is unstoppable."

p. 240: "One grammar ... replacing another ... over time.": (cf. Kroch or Frisch (1997))

(but if  $\alpha, \beta$  are constant, shouldn't the functions be linear with constant slopes  $\alpha/\beta$  and  $\beta/\alpha$ ?)



## 4 french and english V2 loss

(p. 242)

- “In languages such as German, the V2 constraint is very strongly manifested. V>2 patterns are restricted to *if...then* clauses, left dislocation, a small number of adverbs, and so on and are quite rare in frequency (cf. Hawkins, 1985).”<sup>1</sup>

(p. 241)

- “Advantage of V2 grammar over SVO grammar  
V2 → s but SVO → s: VS (XVSO, OVS)”
- “Advantage of SVO grammar over V2 grammar  
SVO → s but V2 → s: V>2 (SXVO, XSVO)”

(modern)	
german, etc. VS	≈30%
english V>2	<10%

### 4.1 french

(stats from Roberts, 1993)

	VS	V>2
Froissart, Chroniques (c. 1390)	18%	12%
15 Joyes (14esme Joye) (c. 1400)	5%	15%
Chartier, Quadrilogue (1422)	7%	11%

	SV	VS	pro
15th century	48%	10%	42%
16th century	77%	3%	15%

<sup>1</sup>ex. 13: “Rare V>2 patterns in modern German

...denn Johann hat gestern das Buch gelesen  
...so Johann had yesterday the book read

Notice in section 4.2 how early ME’s pronoun exception tends to fail with exactly this kind of phrase starter. Odd. Here’s an early MF example of V>2 with a similar temporal adverb introduction (ex. 16):

Lors la royne fist Santré appeller  
then the queen made Santré to-call

(p. 243)

- “Our analysis of the loss of V2 in French crucially relies on the fact that null subjects were lost after V2 was lost.”

## 4.2 english

(stats from Kroch and Taylor, 1997 or Kroch, et al., 2000)

### 4.2.1 southern dialect (early ME V2)

	% verbs inverted (over the subject):	
preposed XP	NP subject	pronoun subject
NP complement	93 (50/54)	5 (4/88)
PP complement	75 (12/16)	0 (0/12)
adj complement	95 (20/21)	33 (7/21)
<i>þa/then</i>	95 (37/39)	72 (26/36)
<i>now</i>	92 (12/13)	27 (8/30)
PP adjuncts	75 (56/75)	2 (2/101)
adverbs	57 (79/138)	1 (1/182)

(p. 246)

- “When subject pronouns could no longer be analyzed as clitics but only as noun phrases, the SVO grammar gained a greater advantage than the V2 grammar and eventually rose to dominance.”

### 4.2.2 northern dialect (germanic V2)

(p. 247)

- “For the northern dialect, the initial contact with the southern dialect was crucial in the loss of V2.” [“in the period of ME”]

	verb inversion
pre-contact	93%
post-contact	68% (208/305)

% verbs inverted (over the subject),  
*after language contact,*  
 in the northern manuscript (Thorton) of the Mirror of St. Edmund:

preposed XP	NP subject	pronoun subject
NP complement	100 (8/8)	64 (16/25)
PP complement	88 (21/24)	70 (48/69)
adj complement	100 (10/10)	25 (2/8)
<i>then</i>	86 (6/7)	51 (24/47)
<i>now</i>	100 (4/4)	82 (14/17)
adverbs	80 (20/25)	57 (35/61)

(p. 247)

- a. “That is, a V2 language similar to the northern dialect would not lose V2 without language contact, which would introduce a substantial amount of V>2 patterns for the learner, even if its morphological case were lost. Northern Germanic languages such as Swedish, Danish, and Norwegian, with impoverished morphological case systems but nevertheless strongly V2, fall into this category.”
  
- b. “The loss of morphological case resulted in the loss of clitics, which further favored the SVO grammar and eventually culminated in its dominance.” (cf. Van Kemenade (1987) or Kiparsky (1997))

## 5 dynamic detour

(cf. Yang (2003): 30 binary parameters =  $2^{30}$  possible grammars > 1 billion)

- $V^N$ 
  - $p(\{t,t\})=.25$
  - $p(\{t,f\})=.25$
  - $p(\{f,t\})=.25$
  - $p(\{f,f\})=.25$
  
- N
  - $\{p(t)=.5, p(t)=.5\}$

## 6 his conclusions

(p. 248)

- “First, the deviation of child language from adult language is not simply noise or imperfection; it is the reflection of actual grammatical hypotheses.”
- “Second, language use is inherently variable, and there is evidence of multiple grammars in mature speakers during the course of language change.”

## 7 my summary

3 explanations from V2 to SVO:

- a. null subjects (French)
- b. pronoun decliticization (southern English)
- c. language contact (northern English)

## 8 a problem?

In French, null subjects faded much faster than V2.

Lightfoot (1994) citing Catharine Dolan (p.c.) in his review of Roberts (1993) in Language:

	9-11TH C.	12TH C.	13TH C.	14TH C.	15TH C.	16TH C.
NULL SUBJECTS	913	1205	1051	867	765	478
% of all subjects	57%	56%	48%	41%	38%	24%
V2	1153	1791	1854	1597	1489	1204
% of main clauses	72%	84%	85%	76%	75%	60%
SUBJ + VERB	305	417	529	463	471	658
% of main clauses	19%	20%	24%	22%	24%	33%
NON-SUBJ + VERB	848	1374	1325	1134	1018	546
% of main clauses	53%	64%	61%	54%	51%	28%

TABLE 1. Loss of null subjects and verb second.

## 9 my conclusions

- a. Given the variety of flexible exceptions which would need to be incorporated into any adequate description of the V2 phenomena, it appears idealistic to presume some sort of innately specified V2 parameter.
- b. Given the independence, contingency, and gradability of language properties, it is not suitable to speak of “competing grammars”. Heterogeneity needs to be analyzed and measured with regard to specific properties rather than assumed. Languages are more like fluids than rocks.

## References

- Ellegård, Alvar (1953) *The Auxiliary Do*, Almqvist & Wiksell (Göteborg), Stockholm.
- Frisch, Stefan (1997) “The change in negation in Middle English,” *Lingua*, 101, 21–64.
- Han, Chung-Hye (????a) “Cross-linguistic Variation in the Compatibility of Negation and Imperatives,” .
- Han, Chung-Hye (????b) “The Evolution of Do-Support In English Imperatives,” .
- Han, Chung-Hye and Kroch, Anthony (????) “The rise of do-support in English: implications for clause structure,” .
- Kroch, A. and Taylor, A. (1997) “Verb movement in Old and Middle English: dialect variation and language contact,” .
- Kroch, Anthony (????) “Syntactic Change,” .
- Kroch, Anthony and Taylor, Ann (2000) “Verb-Object Order in Early Middle English,” .
- Pintzuk, Susan (1999) *Phrase Structures in Competition*, Garland, New York.
- Yang, C.D. (1999) “A selectionist theory of language development,” in *Proceedings of 37th Meeting of the Association for Computational Linguistics*, pp. 429–435, Association for Computational Linguistics, East Stroudsburg, PA.
- Yang, Charles D. (????) “A Formal Theory of Language Development,” .
- Yang, Charles D. (2000) “Internal and external forces in language change,” *Language Variation and Change*, 12(3), 231–250.
- Yang, Charles D. (2003) *Knowledge and Learning in Natural Language*, Oxford, USA.