Two Fundamental Similarities in Biological and Language Evolution and their Fundamental Differences

John A. Hawkins, University of California Davis & Cambridge University

[A] Introduction

Within biology genetic replication creates variants to which adaptive selection can apply. These replications, in both genes and living cells more generally, are not always perfect copies. Some of the genetic variants may then confer certain advantages for survival in successive generations. One thinks of Darwin's (1860) finches and the gradual adaptations in the shapes and sizes of their beeks, which optimized the gathering of different food sources in neighboring islands. The fittest variants survived, thrived and reproduced. Genetic mixing through sex-based reproduction then adds to this pool of variation, with the combinations respecting biological rules of dominance or recessiveness, etc.

Within linguistics there is similarly a pool of variants in many areas of different languages, e.g. alternative word orders, relative clauses with or without gaps, presence or absence of case marking, etc. Principles of preferential selection also apply to this pool, with the result that some may become more frequent and are conventionalized at the expense of others (see Haspelmath 1999). Adding to the pool of variants in linguistics is grammatical mixing from distinct languages as a result of language contact and bilingualism (Trudgill 2011).

Both fields are fundamentally similar in that they see evolution and change emerging from a pool of variants, and they both see the mixing of distinct individuals (with their different genes or grammars) as contributing to this pool along with more internal replications across successive generations. But within these similarities are profound differences in the principles of preferential selection from the pool. First, selection in biology reflects survival and the fitness of the species as a result of food procurement, adapting to the environment (warmth/coolth), physical abilities (use of the hands), bigger brains (being smarter), and so on. In linguistics, the selection does not benefit survival and reproduction, but rather efficiency in communication and in the linguistic code, i.e. using the most minimal processing effort to achieve communicative success in delivering meanings from speaker to hearer as rapidly as possible (Hawkins 2004, 2014). Second, the biological rules of gene combination and reproduction are quite different from the social dynamics of contact and influence under bilingualism and from the psycholinguistics of first or second language learning and equal or unequal bilingualism, all of which lead to different kinds of language change (Hawkins & Filipovic 2023, Trudgill 2011).

[B] The Pool of Variants in Grammars and Language Usage

In the patterns of variation across the world's languages we see some clear preferences. These preferences reveal principles of efficiency and complexity that motivate them. One was suggested to me when working on Greenberg's word order correlations (Hawkins 1983, 1990).

(a) Greenberg's word order correlations

(1) a. vp[went pp[to the movies]] b. [[the movies to]pp went]vp
---------- ----------
c. vp[went [the movies to]pp] d. [pp[to the movies] went]vp
---------- ----------

(2) a. vp[V pp[P NP]] = 161 (41%) b. [[NP P]pp V]vp = 204 (52%)
c. vp[V [NP P]pp] = 18 (5%) d. [pp[P NP] V]vp = 6 (2%)
Preferred (2a)+(b) = 365/389 (94%) [Data from Dryer's 1992 sample]

The adjacency of V and P in (1a) and (1b) guarantees the smallest possible string of words for recognition (/construction) of VP and its two I(mmediate C(onstituent)s (V and PP).

Hypothesis: the recognition and construction of all phrases and their combinatorial relations prefers the smallest possible string of words for processing: more generally Minimize Domains.

(b) Minimize Domains (MiD) [Hawkins 2004]

> The human processor prefers to minimize the connected sequences of linguistic forms and their conventionally associated syntactic and semantic properties in which relations of combination and/or dependency are processed. The degree of this preference is proportional to the number of relations whose domains can be minimized in competing sequences or structures, and to the extent of the minimization difference in each domain.

(c) Preferred selections in performance from the pool in a head-initial language

Linear orderings are preferred that reduce the number of words needed to recognize and construct phrases and their combinatorial relations, making phrase structure processing faster.

(3) a. The man vp[waited pp1[for his son] pp2[in the cold but not unpleasant wind]]
1 2 3 4 5
---

b. The man vp[waited pp2[in the cold but not unpleasant wind] pp1[for his son]]
1 2 3 4 5 6 7 8 9

The three items, V, PP1, PP2 (ICs of the VP) can be recognized and constructed on the basis of five words in (3a), compared with nine in (3b), assuming that (head) categories such as P immediately project to mother nodes such as PP, enabling the parser to construct them on-line.

(3a) VP: IC-to-word ratio = 3/5 or 60% (3b) VP: IC-to-word ratio = 3/9 or 33%

MiD prefers short before long ICs in head-initial phrases (Hawkins 1994, 2004).

Structures like (3) were examined (Hawkins 2000) in which the two PPs were permutable with truth-conditional equivalence (i.e. the speaker had a choice). Only 15% (58/394) had long before short (most with strong semantic dependencies between V and the P in the long PP and a preferred adjacency between them for semantic processing). Among those with at least a one-word weight difference, 82% had short before long, and there was a gradual reduction in the dispreferred long before short orders the bigger the weight difference (PPS = shorter PP, PPL = longer PP):

(4) n = 323	PPL > PPS by 1 word	by 2-4	by 5-6	by 7+
[V PPS PPL]	60% (58)	86% (108)	94% (31)	99% (68)
[V PPL PPS]	40% (38)	14% (17)	6% (2)	1% (1)

Numerous other structures reveal the same gradient short-long preference, e.g. English Heavy NP Shift, [V NP PP] => [V PP NP], Extraposition, Extraposition from NP, Particle Movement, see e.g. Hawkins (1994, 2004), Wasow (2002), Stallings (1998), Lohse et al (2004).

(d) <u>Preferred selections in performance from the pool in a head-final language</u>

Long before short orders provide minimal domains in head-final languages, in which constructing categories (V, P, Comp, case particles, etc) are on the right, e.g. in Japanese:

(5) a. Mary ga [[kinoo John ga kekkonsi-ta to]s it-ta]vp
Mary NOM yesterday John NOM married that said, i.e.
Mary said that John got married yesterday.
b. [kinoo John ga kekkonsi-ta to]s Mary ga [it-ta]vp

(6) a. (Tanaka ga) [[Hanako kara]pp [sono hon o]np katta]vp
Tanaka NOM Hanako from that book ACC bought
'Tanako bought that book from Hanako'
b. (Tanaka ga) [[sono hon o]np [Hanako kara]pp katta]vp

(7) <u>Relative ordering of NPo and PPm in Japanese structures like (6)</u>

	ICL>ICS by 1-2 words	by 3-4	by 5-8	by 9+
[ICS ICL V]	34% (30)	28% (8)	17% (4)	9% (1)
[ICL ICS V]	66% (59)	72% (21)	83% (20)	91% (10)

The pattern of gradient long-short preferences in (7) is the exact mirror image of English (4).

(e) Conventionalizing the preferred word orders in the history of Greek

(8) Performance-Grammar Correspondence Hypothesis (PGCH)
Grammars have conventionalized syntactic structures in proportion to their degree of preference in performance, as evidenced by patterns of selection in corpora and by ease of processing in psycholinguistic experiments.

Seržant & Rafiyenko (2021) capture the timeline of grammatical conventionalization for head ordering in the diachrony of Greek from an earlier very free word order in Archaic Greek comprising both head-initial and head-final orders. Head-initial selections from the pool became gradually and simultaneously more frequent, culminating in the basic head-initial grammar of Early Byzantine and Modern Greek, see (9):

(9) **Word order variation and change in the history of Greek** (Seržant & Rafiyenko 2021)
Adp = Adposition (Prep or Postp); Gen = Genitive/Poss. phrase modifying N; {O,V} = Object of Verb; UPPER CASE letters = basic orders, lower case = minority or non-basic variants (cf. Hawkins 1983, Dryer 2005)

(i) Archaic Greek (750-450 BC): nadp/adpn (both postpositions and prepositions)
genn/ngen (both prenominal and postnominal genitives)
ov/vo (both object-verb and verb-object)

(ii) Classical Greek (450-315 BC), Hellenistic (340-0 BC), Roman Period (50-250 AD):
Gradually increasing head-initiality: nadp/ADPN (mainly prepositions)
genn/NGEN (mainly postnominal genitives)
ov/VO (mainly verb-object)

(iii) Early Byzantine Greek (500-700 AD): ADPN (prepositions)
NGEN (postnominal genitives)
VO (verb-object)

(f) Interim conclusion re biology and linguistics

The selection preferences driven by communicative efficiency principles such as MiD ([Bb]) in language usage, historical change and conventionalization are quite different from the selection preferences that apply to genetic variants in terms of species fitness, survival, etc. Communicative efficiency involves speed and minimal processing effort in relaying a message successfully from Speaker to Hearer (cf. Levshina 2022). This is a language- and communication-specific principle of preference very different from the adaptations of species in Darwin (1860) and in Thomas Halliday's (2023) survivals of species that could adapt to a changing environment in the evolving periods of life on earth versus extinctions for those that could not. Both fields, biology and linguistics, see a pool of variants as underlying evolution and change, therefore, but the selection pressures on their respective pools are quite different, as are the pools themselves.

[C] Adding to the Pool of Variants through Grammatical Mixing in Bilingualism

In biology sex-based reproduction adds to the pool of variants through the mixing and combination of genes from distinct individuals. In linguistics a mixing and combination of grammars from distinct individuals speaking different languages takes place when there is bilingualism and language contact (see Trudgill 2011 for extensive discussion and examples).

(a) Maximizing Common Ground under bilingualism

Hawkins & Filipovic (2023) survey the bilingualism and language change literature and propose:

(10) **Maximize Common Ground** (MCG, Filipović 2019: 60)
Bilingual learners and speakers maximize common grammatical and lexical representations and their associated processing mechanisms in their two languages, L1a and L1b.

(10a) If L1a and L1b share a given construction, grammatical rule or word meaning, and associated processing mechanisms, then these shared entities will be used more frequently in both languages. These entities may be the preferred or majority pattern in one language and a minority or dispreferred in the other, but they will still be the pattern of choice in the bilingual's use of both languages.

(10b) If L1a and L1b do not share a given construction, grammatical rule or word meaning, and associated processing mechanisms, then common ground will be created by either introducing entities and rules from one language into the other, or removing them from both. New shared entities will be introduced wherever possible *within* the constraints of current grammatical and usage conventions for the relevant language.

(10c) Violations of a grammatical or usage convention in L1a or L1b that occur when maximizing common ground (i.e. potential language change) will be in proportion to the strength of the environmental and psycholinguistic factors enumerated in Filipović (2019) and Filipović & Hawkins (2019).

(b) Exemplifying MCG effects

With respect to **(10a)** a minority word order in one language, e.g. adjective before noun (AdjN) coexisting with a majority noun before adjective NAdj, as in French and Spanish, will gain in frequency when speakers of these languages are in bilingual contact with a language that has only AdjN, like English (references in Filipovic & Hawkins 2019).

An example of **(10b)** involves evidentiality, i.e. the obligatory grammatical expression of the speaker's source of information for the proposition being expressed, whether as direct and witnessed or instead more indirect. Turkish is a language with obligatory grammatical marking of evidentiality, and among Turkish and English bilinguals this leads to the more frequent expression in English of paraphrasing constructions with evidential meaning (*it*

appears that …, *I am informed that …*, etc), i.e. through a significant increase in the use of optional grammatical and lexical means that are not routinely and obligatorily selected in English but that remain within its grammatical constraints. These optional means maximize the common ground between two languages without actually changing any grammatical conventions in the receiving language (references in Filipovic & Hawkins 2019).

Examples of **(10c),** in which grammatical conventions do change when maximizing common ground, include basic word order and even the head ordering typology of languages in bilingual contact with one another. This has been documented across the globe, with productive shifts both from head-initial to head-final and from head-final to head-initial, reflecting largely sociolinguistic and demographic relations between speakers of the two languages (see below).

In (11) we summarize some well-documented cases of VO to OV shifts that were accompanied by cross-categorial shifts to head-final head ordering within other phrases such as adpositional phrases (prepositions yielding to postpositions), noun phrases (noun before adjective and genitive becoming adjective and genitive before noun), etc:

(11) **VO -> OV** in Austronesian languages in coastal New Guinea, e.g. VO **Takia** shifting to OV under the influence of Papuan **Waskia** (Ross 1996);
Yaqui (Southern Uto-Aztecan now with SOV and head-finality) through contact with SOV Hokan and Northern Uto-Aztecan languages (Lindenfeld 1973);
Sri Lanka Portuguese originally an SVO creole, now rigid SOV under Tamil influence (Heine & Kuteva 2005).

In (12) we give examples of the reverse OV to VO shifts, i.e. head-final to head-initial:

(12) **OV -> VO** in Uto-Aztecan **Pipil** through contact and bilingualism with neighboring VO Mayan languages (Campbell 1985);
Papuan **Kuot** (of New Ireland) from Papuan OV to VSO and consistent head-initial orders, surrounded by head-initial Austronesian languages (Lindström 2002);
Eskimo varieties (OV) in bilingual contact with English (Fortescue 1993).

(c) <u>Sociolinguistic and psycholinguistic determinants of grammar mixing and change</u>
E.g. whether a grammaticalized evidentiality morpheme and morpho-syntax are transferred by MCG from an L1a (which has them) into an L1b (which does not) depends on whether L1a is the socially more dominant and prestigious language and on population demographics. Evidentiality *was* transferred from Turkish into Bulgarian, one of the languages of the Ottoman Empire ruled by Turkish speakers (Slobin 2016), but *not* into higher prestige Greek, despite the same Ottoman rule and hence exposure to Turkish (Lindstedt 2016). Conversely, evidentiality *was* transferred into socially dominant Andean Spanish from Quechua and Aymara, despite political subjugation, but through sheer strength of numbers of the bilingual speakers using evidentiality in their L1s (Slobin 2016, Filipović 2019). It is these

social variables that crucially determine whether a property transfer will occur at all, and the direction of transfer among the languages in question.

Psycholinguistic constraints primarily involve the balanced or unbalanced nature of the bilingualism, which impacts the amount and nature of grammatical transfers. In very unbalanced bilingualism involving L2 learners, common ground between L1 and L2 is regularly created through both positive (i.e. grammatical) and negative (ungrammatical) transfers of L1 features into L2 and grammatical simplifications (Trudgill 2011). Definite and indefinite article omission errors by Russian and Japanese learners of English reflect level of proficiency, and these "errors" are progressively reduced as proficiency improves, resulting in less MCG between L1 and L2 and separate and ultimately correct conventions for the two languages (Hawkins & Filipović 2012, Filipović & Hawkins 2013). More balanced bilingualism leads to more complex features being borrowed, rather than imperfectly learned errors (Trudgill 2011).

[D] Fundamental Differences in Preferential Selections from the Pool of Variants in Biology and Linguistics

There is a pool of variants that leads to evolution and change in both disciplines through selectional preferences, in species and their genes on the one hand, and in languages and their grammars on the other. The pool undergoes gradual changes across successive generations, with imperfect copies and preferred selections applying to what went before and changing the composition and distribution of items in the pool. Adding to this pool of variants is then the mixing of genes and of grammars from distinct individuals. But here the similarities between the disciplines stop. The principles of preferential selection, and the items themselves to which they apply, are quite different: communicative efficiency (e.g. Minimize Domains) and Maximize Common Ground under bilingualism (with its social and psychological determinants) on the one hand; and survival of the fittest and dominant vs recessive genes, etc, in sex-based reproduction on the other.

References:

Campbell, L. (1985 The Pipil Language of El Salvador, Mouton Publishers, Berlin.
Darwin, C. (1860) On the Origin of Species, John Murray, London.
Dryer, M.S. (1992) 'The Greenbergian word order correlations', Language 68: 81-138.
Dryer, M.S. (2005) 'Determining dominant word order'. In M. Haspelmath, M.S. Dryer, D. Gil, & B. Comrie (eds.), The World Atlas of Language Structures, OUP, Oxford, 371.
Filipović, L. (2019) Bilingualism in Action: Theory and Practice, CUP, Cambridge.
Filipović, L. & J.A. Hawkins (2013) 'Multiple factors in second language acquisition: The CASP model', Linguistics 51(1): 145-176.
Filipović, L. & J.A. Hawkins (2019) 'The CASP model of bilingualism: Language interactions within and across bilingual minds. International Journal of Bilingualism 23(6): 1223-1248.
Fortescue, M. (1993) 'Eskimo word order variation and its contact-induced perturbation', Journal of Linguistics 29(2): 267-28
Halliday, T. (2023) Otherlands: A World in the Making, Penguin Random House, UK.
Haspelmath, M. (1999) 'Optimality and diachronic adaptation', Zeitschrift für Sprachwissenschaft 18: 180-205.
Hawkins, J.A. (1983) Word Order Universals, Academic Press, New York.

Hawkins, J.A. (1990) 'A parsing theory of word order universals', Linguistic Inquiry 21: 223-261.
Hawkins, J.A. (1994) A Performance Theory of Order and Constituency, CUP, Cambridge.
Hawkins, J.A. (2000) 'The relative ordering of prepositional phrases in English: Going beyond manner-place-time', Language Variation and Change 11: 231-266.
Hawkins, J.A. (2004) Efficiency and Complexity in Grammars, OUP, Oxford.
Hawkins, J.A. (2014) Cross-linguistic Variation and Efficiency, OUP, Oxford.
Hawkins, J.A. & L. Filipović (2012) Criterial Features in L2 English: Specifying the Reference Levels of the Common European Framework, CUP, Cambridge.
Hawkins, J.A. & L. Filipovic (2023) 'Bilingualism-induced language change: What can change, when and why?', Linguistics Vanguard 10(s2): 115-124. https://doi.org/10.1515/lingvan-2023-0089
Heine, B. & T. Kuteva (2005) Language Contact and Grammatical Change, CUP, Cambridge.
Levshina, N. (2022) Communicative Efficiency: Language Structure and Use, OUP, Oxford.
Lindenfeld, J. (1973) Yaqui Syntax, University of California Press, Berkeley.
Lindstedt, J. (2016) 'Multilingualism in the Central Balkans in late Ottoman times'. In M. Makartsev & M. Wahlström (eds.), In Search of the Center and Periphery: Linguistic Attitudes, Minorities, and Landscapes in the Central Balkans [Slavica Helsingiensia 49], University of Helsinki, Helsinki, 51-67.
Lindström, E. (2002) Topics in the Grammar of Kuot, a Non-Austronesian Language of New Ireland, Papua New Guinea, PhD dissertation, Stockholm University, Stockholm.
Lohse, B., Hawkins, J.A. & Wasow, T. (2004) 'Domain Minimization in English verb-particle constructions', Language 80: 238-261.
Ross, M.D. (1996) 'Contact-induced change and the comparative method: Cases from Papua New Guinea'. In M. Durie & M.D. Ross (eds.), The Comparative Method Reviewed: Regularity and Irregularity in Language Change, OUP, Oxford, 180-217.
Seržant, I.A. & D. Rafiyenko (2021) 'Diachronic evidence against source-oriented explanation in typology: Evolution of prepositional phrases in Ancient Greek', Language Dynamics and Change 11: 167-210.
Slobin, D.I. (2016) 'Thinking for speaking and the construction of evidentiality in language contact'. In M. Güven, D. Akar, B. Öztürk & M. Kelepir (eds.), Exploring the Turkish Linguistic Landscape: Essays in Honor of Eser Erguvanli-Taylan, John Benjamins, Amsterdam, 105-120.
Stallings, L. M. (1998) 'Evaluating Heaviness: Relative Weight in the Spoken Production of Heavy-NP Shift', Ph.D. dissertation, University of Southern California.
Wasow, T. (2002) Postverbal Behavior, CSLI Publications, Stanford University, Stanford.
Trudgill, P. (2011) Sociolinguistic Typology, OUP, Oxford.