WIREs COGNITIVE SCIENCE

Article type: Focus Article

Article title: Cognitive processing, language typology, and variation

Kaius Sinnemäki, Helsinki Collegium for Advanced Studies, University of Helsinki;
kaius.sinnemaki@helsinki.fi

Abstract

Linguistic typological preferences have often been linked to cognitive processing preferences but often without recourse to typologically relevant experiments on cognitive processing. This paper reviews experimental work on the possible parallels between preferences in cognitive processing and language typology. I summarize the main theoretical accounts of the processing-typology connection and show that typological distributions arise diachronically from preferred paths of language change, which may be affected by the degree to which alternative structures are preferred (e.g. easier) in acquisition or usage. The surveyed experimental evidence shows that considerable support exists for many linguistic universals to reflect preferences in cognitive processing. Artificial language learning experiments emerge as a promising method for researching the processing-typology connection, as long as its limitations are taken into account. I further show that social and cultural differences in cognition may have an effect on typological distributions and that to account for this variation a multidisciplinary approach to the processing-typology connection has to be developed. Lastly, since the body of experimental research does not adequately represent the linguistic diversity of the world's languages, it remains as an urgent task for the field to better account for this diversity in future work.

The past 50 years of typological research has accumulated several generalizations on cross-language distributions, including the much celebrated word order correlations.[1, 2] A central goal in typology is

to determine, by using data from a wide range of languages, the extent of cross-linguistic variation and the interrelationships among linguistic patterns.[3] A typological generalization, once established, is typically hypothesized to reflect biases in cognitive processing via conventionalizing structures which are preferred (e.g. easy) in language acquisition or use.[4–8] However, it is still unclear to what extent such generalizations may be grounded in cognitive biases, owing to a large gap between the fields of typology and psycholinguistics.[9]

In this review, I first summarize the main theoretical accounts of the connection between typology and processing preferences. I then review to what extent typological generalizations may or may not reflect constraints on language that could be grounded in cognitive processing. I limit myself to morphosyntactic patterns, especially linear ordering and argument marking, since those are the most thoroughly studied in typology. Lastly I focus on the challenges that cultural and linguistic diversity (among others) present to the processing-typology connection.[10,11]

DIACHRONIC AND COGNITIVE ROOTS OF TYPOLOGICAL GENERALIZATIONS

Typological generalizations fall into two main types, absolute universals, which are true of all languages (e.g. all languages have CV syllables) , and statistical universals, which are statistical tendencies (e.g. most languages have subject agreement).[1] Typologists are mainly interested in the latter, because half a century of empirical research has unearthed only a few absolute universals.[3,12] Evidence from computational models further shows that absolute universals may be practically impossible to detect from the available data.[13]

There are two main sampling approaches that typologists use for researching statistical universals. First, the effects of shared inheritance and horizontal transmission (language contact) may be controlled via genealogically balanced sampling, by controlling how many languages are chosen from each language family,[2,14,15] and the remaining effect is then attributed to cognitive bias.[16-18] However, genealogically balanced sampling produces mere synchronic snapshots of the type-distributions and does not allow making diachronic interpretations from those distributions.[17] The second approach tries to solve this problem by taking an essentially diachronic approach to universals, whereby universal are understood as structural pressure that changes languages over time (see References 16-21 and references there). Languages are then sampled both across and within families to assess the probability of a language shifting from one type to another.

Linguistic patterns arise through diachronic change and, therefore, the typological distributions of linguistic patterns are the product of preferred paths of language change.[3,6–8,17] More precisely, preferred structures are developed frequently in language change and maintained more easily compared to dispreferred structures, which do not develop as frequently, and when they do,

they are more susceptible to marginality or loss.[17,19,20] For instance, in transitive sentence across languages the subject (S) tends to occur before the object (O) (1a), while the reverse order (1b) is very rarely dominant in a language and when it is, there is pressure to shift to S-O order.[17]

1. (a) John kissed Mary.
   (b) Mary John kissed.

However, disagreement exists over the main driving force of such changes, whether by language acquisition or language use, and over the extent to which cognitive processing is assumed to affect those changes (see Fig. 1).

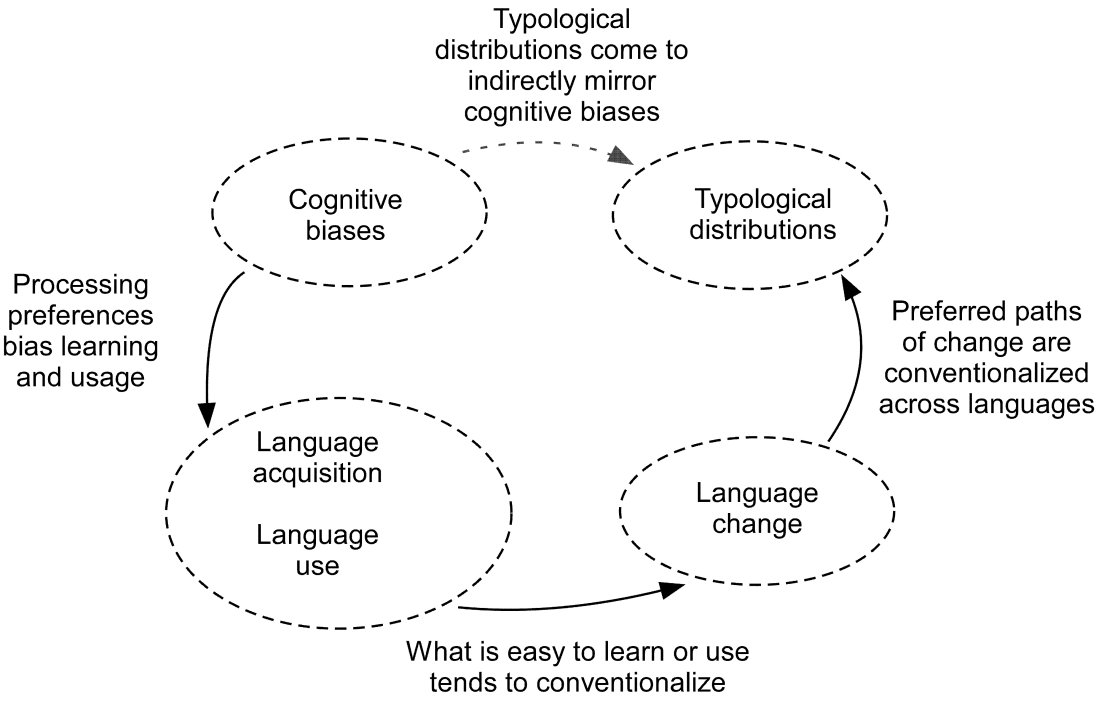

Figure 1. Typological distributions reflect cognitive biases indirectly. In the generative theory, cognitive biases affect language acquisition but in usage-based theory they primarily affect language use. By hypothesis, what is preferred (e.g. easier) in acquisition or usage tends to become conventionalized across languages and what is dispreferred conventionalizes less often or tends to wither away. Statistical universals arise when linguistic structures across languages change through such preferred paths of language change.

Theoretical accounts of the processing-typology connection

There are two main theories for accounting how typological distributions end up reflecting cognitively grounded preferences in language change. In generative grammar language change is assumed to occur mainly in language acquisition: children are assumed to enter the acquisition process with genetically endowed constraints or learning biases, which delimit the set of possible grammars to be acquired and thus facilitate the acquisition process.[8,22-25] Language change is then widely assumed to result from imperfect learning of linguistic patterns, which are retained through the acquisition process until adulthood.[26] A growing body of research in this tradition interprets the constraints as probabilistic cognitive learning biases and searches for converging evidence for typological distributions from artificial learning experiments (see Reference 8 for a review).

It has long been observed that children are prone to innovate and over-generalize linguistic structures. Recent learning experiments support this observation and suggest that children are even more prone to regularize and systematize inconsistent input structures than adult learners are,[27] providing potential evidence for the role of children in language change. However, the child's propensity to regularization cannot be the sole reason for language change, since children are quick to retreat from overgeneralization (see Reference 28 for a review). When faced with inconsistent input structures, children also do not unanimously regularize a particular structure but choose different input structures to regularize.[27] While children may introduce novelties and changes to their input, evidence from acquisition studies suggests that generally these changes do not spread to adult language and they are not maintained into adulthood (see References 12 and 29 and references there).

In usage-based models of grammar, on the other hand, grammar is assumed to be directly shaped by language use, so that easy and frequently used linguistic structures tend to conventionalize in grammars, while more difficult structures tend to disappear.[4,5,12] These models emphasize the important role of frequency in shaping grammars, since repetition leads to conventionalization but also to more economical production (reduction via automatization) and easier processing (greater accessibility due to stronger mental representation). Structural asymmetries are assumed to emerge from a combination of multiple factors in language use, such as processing preferences, language acquisition and conversational interaction, whereby language universals may reflect a multitude of factors.

Evidence for the usage-based approach comes from several sources. Evidence from large corpora suggest that typological distributions tend to match in direct proportion the preferences found in performance.[4] Evidence from language acquisition emphasizes the important role of frequency in acquisition, and frequency also plays a central role in the process of grammaticalization

(see References 7 and 12 and references there). Overall, usage-based accounts of language change are empirically well-supported.

Although frequency seems important in shaping grammars, recent experimental and corpus data suggests that efficient distribution of information in a sentence (that is, communicative efficiency) may provide a more general account of preferences in performance.[9,30] This account may also better explain why rare and difficult patterns can be conventionalized in particular contexts instead of disappearing. For instance, while structures in which the object precedes the subject are generally dispreferred in German, evidence from corpora and experiments shows that in whole-part constructions OVS structures (2b) are more acceptable and easier to process than the corresponding SVO structures (2c) (Reference 31, p. 7). Rare patterns can, therefore, become preferred in marked contexts to express particular functions.

2. (a) Peter hat den Wagen gewaschen.
   Peter has the.ACC car washed
   'Peter has washed the car.'
   (b) Den Außenspiegel hat er ausgelassen.
   The.ACC side.mirror has he.NOM left.out
   'The side mirror, he left out.'
   (c) Er hat den Außenspiegel ausgelassen.
   He.NOM has The.ACC side.mirror left.out
   'He left out the side mirror.'

Detecting biases in processing and learning

In order for the theoretical accounts to succeed in modeling the relationship between cognitive processing and typological distributions, cognitive biases need to be well-defined and empirically assessable.[9] A basic step towards this direction is to separate cognitive processing from the complexity of linguistic structures to enable the empirical assessment of their relationship.[4,16,32] The next step then is to provide a feasible definition for cognitive bias.

By cognitive biases I refer here to processing complexity, which can be measured by experimental methods such as reaction times and brain activity,[9,32] and behavioral preferences in language learning, which can be interpreted as evidence for cognitive biases in learning.[8,24] Two main theoretical accounts have been proposed for processing complexity (see References 9 and 33 for recent reviews). Memory-based theories emphasize the limits of human memory capacity, which engender difficulty when overloaded in processing (e.g. long-distance dependencies).[4,32,34,35] In

expectation-based theories, on the other hand, processing difficulty is tied to the probability of a linguistic structure in a given context, whereby different linguistic structures reduce uncertainty of the rest of the sentence to varying degrees (e.g. in English objet-initial sentences are unexpected and therefore difficult to process).[30,36,37] While in some situations these theories make virtually opposite predictions concerning processing difficulty, in the light of recent findings predictions of both theories are supported but at different sentence loci, a result which calls for a more integrated model for sentence processing.[33]

When reviewing the experimental literature, my working hypothesis is that cognitive biases are reflected in the experiments in terms of the length of reaction or processing times, the degree of brain activity or the degree to which alternative linguistic structures are (dis)preferred in behavioral tasks. I include experiments on both language learning and language use, assuming generally that what is preferred in usage tends to be preferred in learning as well.[4,5]

EXPERIMENTAL EVIDENCE FOR THE PROCESSING-TYPOLOGY CONNECTION

There is great variation in the degree of attention given to different typological patterns in experimental studies. Two patterns in particular have been carefully investigated, namely, the word order of subject and object in the sentence and the relativization of subjects and objects. Other patterns have received much less attention and have been researched by fewer methods.

Word order and relativization of subject and object

The first of Greenberg's[1] famous universals states that languages tend to position the subject of declarative main clause before the object (1a) as opposed to positioning the object before the subject (1b). This generalization has been investigated in a number of studies and is strongly supported by typological evidence (see Reference 17 and references there). It is also strongly supported by experimental evidence from a wide range of languages (see References 38, 39 and references there). Regardless of the experimental method used (reading times, eye-tracking, event-related brain potentials (ERP) or functional magnetic resonance imaging (fMRI), among others) object-initial sentences are harder to process than subject-initial sentences.

However, there are some complicating issues. A number of factors may alleviate or even neutralize the processing load of an object-initial sentence, such as case marking of the object,[40] supportive discourse context[41] and the semantic-pragmatic prominence of the arguments (e.g. animacy).[39] The prominence factors may even outweigh any ordering preferences of subject and object. For instance, Japanese speakers preferred to position animate NPs before inanimate NPs in a sentence recall task, and independently of the grammatical role of the argument.[42] Data from

German further shows that initial inanimate nominative subjects are more difficult to process than dative-marked objects (3).

3. Dann wurde der Mantel dem Arzt gestohlen.
   then was the.NOM coat the.DAT doctor stolen
   'Then the coat was stolen from the doctor.' (Reference 43, p. 1397)

These results seem to cast doubt on the universality of the subject preference and emphasize the role of semantic-pragmatic factors in sentence production and comprehensions. Nevertheless, further evidence suggests that an initial argument in such constructions is first preferentially interpreted as subject but this interpretation can be modulated at a later stage in processing through discourse context, animacy and pragmatic factors.[44,45] These findings have been explained in the cross-linguistically motivated extended Argument Dependency Model (see Reference 39 and 45 for details). Since this model accounts for language comprehension only, it is unclear whether it would account for the preferences in the production of languages like Japanese as well.

One further challenge to the universality of the subject preference is that almost all the languages experimented have had a preferred subject-object word order in their grammar. The risk here is that the observed preference may be simply due to the typological properties of the languages studied so far. In fact, recent experiments using sentence plausibility judgments in Kaqchikel Maya, a language with basic verb-object-subject order (VOS), shows that object-initial orders are easier to process regardless of the animacy of the object.[46] However, since the verb occurs in the beginning of the sentence, it provides early access to argument structure making both arguments more accessible. Such results, therefore, do not necessarily provide a challenge to the subject-preference as a universal cognitive preference but they do call for more research on verb-initial and object-initial languages.

The subject preference has also been demonstrated in relative clauses. According to typological data, all languages can relativize the subject of the relative clause (4a) but not necessarily the object (4b).[47] The possibility of relativizing the object thus implies the possibility of relativizing the subject but not vice versa. Experimental evidence from a number of languages provide support for this typological generalization showing that subject relative clauses are easier to process than object relative clauses (see References 38, 39, 48 and references there).

4. (a) The man that $\_\__{SUBJ}$ drove the car.
   (b) The car that the man drove $\_\__{OBJ}$.

However, there is also evidence against the subject-preference in relative clauses. Evidence from self-paced reading assignments in Russian show that no overall processing cost exists for object relative clauses compared to subject relative clauses.[33] Experimental evidence from Basque using self-paced reading tasks and ERP has further shown that object relative clauses are easier to process than subject relative clauses[48] despite the greater difficulty of object-initial order in declarative main clauses.[38] It is thus unclear whether the subject preference in relative clauses is a universal cognitive preference or a processing bias in particular types of languages. More research in typologically diverse languages is needed to solve the issue.

Evidence from artificial language learning experiments

Artificial language learning experiments have been used in a number of recent studies to test the extent to which learning biases match typological distributions (see Reference 8 for a review). In the experiments participants are taught artificial mini-languages usually via presentation software.[24,49] Typical stimuli involves pictures and animations with simultaneous audio descriptions of the objects and events. Learning biases may then be assessed by testing the learners for generalizing the learned patterns to new data or by testing to what extent their behavior shifts towards or away from the input which contains a mixture of patterns.[8,49]

One domain in which these methods have been used is the suffixing preference in morphology. This preference concerns the linear position of bound morphemes in relation to the word stem, namely, that there is a strong typological preference for suffixes (e.g. *-er* in *player*) over prefixes (e.g. *re-* in *replay*).[1,50] Evidence from artificial learning experiments provide robust support for this preference.[8,49] A plausible explanation comes from a domain-general preference to express salient information in the beginning of a sequence; because word roots are more informative than affixes, prefixes violate this preference while suffixes conform to it.[49] However, these issues require further study, since the suffixing preference seems to correlate with the degree of affixation in a language: languages with less affixation tend to use either prefixing or suffixing, while those with a lot of affixation tend to use both.[6]

Another domain investigated by these experiments is the interaction between case marking and animacy in languages with so-called differential object marking. In these languages only untypical objects (e.g. animate) are case-marked (5a), while typical objects (inanimate) are unmarked (5b), as in Hindi (Reference 51, p. 79).

5. (a) Ilaa-ne ek bacce-ko ʊʈʰaaya.

Ila-ERG one child-ACC lift.PERF

'Ila lifted a child.'

(b) Ilaa-ne ek haar uʈʰaaya.

Ila-ERG one necklace lift.PERF

'Ila lifted a necklace.'

There is wide agreement in typology that differential object marking depends on animacy in this way (see Reference 52 and references there, although see Reference 53). A recent artificial learning experiment shows that when optional case marking in the training data was unconnected to animacy, the participants tended to change the languages so that only animate objects were case-marked.[54] Since animacy of the referent increases the ease of its retrieval from memory, case marking of animate objects might be related to conceptual accessibility.[10] As for language comprehension, case marking of animate objects may be explained by the principles of distinctness and minimality in the framework of the extended Argument Dependency Model (see Reference 39 and references there). According to these principles the arguments should be kept distinct from one another via case marking, word order or the semantic properties of the referents but avoiding unnecessary distinctness at the same time.

Experimental evidence supports another well-known typological pattern that case-marking languages tend to use free word order while languages with no case marking tend to fixate their word order for argument discrimination.[16] An experiment using mixed pattern input showed that participants trained in a case-less language with relatively free word order regularized the word order for argument discrimination while those trained in a case-marking language did not regularize word order.[55]

One further group of typological generalizations which has been studied in this paradigm is the word order correlations between the phrasal head and its modifiers in various types of phrases. According to cross-language data, if the verb precedes the object, languages prefer to use prepositions and to position the head noun before the genitive (head-first order, e.g. *behind John's house*), but if the verb follows the object, languages prefer postpositions and to position the head noun after the genitive (head-last order, e.g. *house John's behind* in cryptic English).[1,2] An experiment using miniature languages with consistent or inconsistent head-ordering for different participant groups showed that the participants preferred the consistent head-ordering.[56] A related typological dispreference exists against the order adjective-noun-numeral (e.g. *beautiful flowers three* in cryptic English), while no such dispreference exists against the head-terminal orders (adjective/numeral-noun and noun-adjective/numeral) or the order numeral-noun-adjective in the

noun phrase. An experiment using mixed-pattern input showed that the adjective-noun-numeral order is dispreferred also by learners.[24]

It is clear even from this brief review that artificial language learning experiments provide promising possibilities for researching the processing-typology connection. To the extent the learning biases in the experiments parallel typological distributions, it is possible that similar cognitive biases may underlie both.[8] One clear advantage of these experiments is that they allow the controlling of interfering factors, which is impossible in studies of classroom second language learning.[57] Despite this advantage it is necessary to emphasize that greater learnability may not always suffice to show a learning bias which could produce universals. A recent experiment used iterated learning of artificial languages to study the learnability and transmission of vowel harmony and showed that while the preferences of the first participants in the learning chains agreed with typological preferences, this bias was not replicated in subsequent participants in the chain (the results of the first participants' forced choice task served as the input to the next participants in the chains).[58] Another concern is interference from the participants' previous language experience, which may affect the learners preferences in the experiments.[59] Future improvements of artificial learning experiments should involve iterated experiments and using participants from different linguistic backgrounds to better control for the effect of familiarity.

Evidence from gesture production tasks

Gesture production is another behavioral experimental paradigm that has recently taken typological generalizations to task. In these experiments, participants first see video clips of events and then gesture the event meanings, which are recorded for later analysis. These experiments have focused on two typological generalizations concerning the interaction of case marking and word order. Languages with dominant SOV order are typologically prevalent compared to other logically possible orders of subject, object and verb,[1,14] but they are also more prone to developing case marking of subject or object compared to languages with some other dominant order, especially SVO (see References 20 and 21 and references there).

Evidence from gesture production tasks provides evidence for these distributions. In these experiments participants tend to either avoid SOV order or invent gestures for case marking when the subject and object are reversible, that is, equally plausible candidates for being the subject (e.g. man pulling a horse), but when the subject and object are non-reversible (e.g. man pulling a chair), participants strongly prefer SOV order.[60,61] In addition, gestures for case marking are less common in SVO orders than in other orders.[61] Since the participants' behavior paralleled typological preferences

it is possible that similar cognitive biases may underlie the participants' production and ultimately the typological distributions as well.

Interim conclusion

In summary, there is a growing body of experimental evidence showing that preferences in learning and processing agree with typological preferences. However, different typological generalizations have received varying degrees of attention in the experimental literature, and more research is needed to solidify the many initial results reviewed here. In addition, more experimental research is needed to study other typological generalizations which have received little or no attention so far. These include, just to name a few, the hierarchies for number and person, the relationship between inflectional synthesis and the suffixing preference, the relationship between definiteness and object case, and the relationship between the relative orders of object and verb and the relative clause and its head noun.

CHALLENGES AND FUTURE DIRECTIONS

We have seen that there is a large body of experimental evidence showing that many well-known statistical universals may reflect cognitive processing preferences. These processing preferences may be linked to typological distributions via models that connect language change with learning and usage.[5,7,8] However, there are at least four challenges that future research needs to address to establish a more principled processing-typology connection.

Challenges from cross-linguistic and cognitive diversity

One of the more pressing challenges to the processing-typology connection comes from cross-linguistic diversity. At the current stage research on language processing and learning does not adequately reflect the degree of cross-linguistic diversity. The majority of experimental data come from the languages of Europe,[10,62] but the greatest linguistic diversity clusters in the Pacific and the Americas. Moreover, the structural-typological profiles of languages in these areas appear to be markedly different from those in Eurasia and Africa.[63] More research is, therefore, desperately needed from languages in the Americas and the Pacific. A good start would be the design of small diversity samples based on social and typological profiles of languages in large typological databases such as the World Atlas of Language Strutures.[62,64]

Another challenge comes from social and cultural differences in cognition. Cognitive research has long overlooked such variation,[18,65] but current research suggests that because

languages are acquired, learned and used in different social and cultural contexts, those contexts may bias language acquisition and use so much so that language structures are adapted to these biases,[5,11] ending up reflected in typological distributions as well.[66,67] Initial findings emphasize the role of social context (but see Ref 15 for the role of cultural context).[11,67-69] In small communities people have a high degree of shared knowledge. This type of social context favors the maintenance and even the development of linguistic structures that are difficult for outsiders and adults to learn (e.g. irregularity and morphological complexity). In large communities, on the contrary, communication takes place more among strangers from different languages or dialects and the more adults there are in the community learning its language, the greater the likelihood is for the community's language to become easier and more regular. A solution to this challenge should involve better integration of typology, cognitive studies, anthropology and sociolinguistics.[10,62,65,69-71]

The existence of social and cultural differences in cognition suggests also that the results of the artificial language learning paradigm need to be recast in this wider view on cognitive variation. While evidently tapping some cognitive biases, artificial learning methods (using adult participants) necessarily simulate adult second language learning whereby the learning context is more reminiscent of communication between strangers than between intimates. As a result, these methods may overemphasize the cognitive bias in learning towards regularization and greater transparency, and, therefore, overestimate the degree to which typological generalizations reflect learning biases. A possible solution could be to develop experiments on communication between intimates perhaps involving artificial languages or language games.

How to better link diachrony with processing and typology

Another set of challenges involves the indirect relationship between cognitive processing and typological generalizations through language change. One challenge is to clarify to what extent and in what way cognitive processing may affect language change. According to research on diachrony, many statistical universals could be artifacts of similar linguistic structures arising from different diachronic origins without a need to recourse to cognitive preferences in acquisition and usage.[72,73] As was mentioned earlier, the relative orders of noun and adposition and the genitive and the head noun correlate typologically. However, this correlation may largely owe to the fact that adpositions often arise from possessive constructions retaining the original order in language change (see Reference 73; but see Reference 26 for criticism). While cognitive preferences in language learning and use may arguably affect language change to some extent,[5,73] it is also possible that those preferences differ from the cognitive processes behind the historical development of linguistic

patterns (see References 29 and 72 and references there). Future work should, therefore, better address the relationship between cognitive processes and language change.

One further challenge involves the development of typological methodology. On the one hand, the majority of typological generalizations are based on genealogically balanced sampling, which does not provide dynamic interpretation for synchronic distributions. Approaching universals as preferred laws of type-change is a recent methodological development in the field and more research is required to test a wider range of hypotheses with these methods. Implementing phylogenetic methods to typological data is a promising further development to model correlated changes within language families (see Reference 65 for a review). On the other hand, the effect of horizontal transmission (areal diffusion) is often seen merely as a nuisance factor in typology. However, if a linguistic pattern spreads quickly from one language to another, this ease of transfer may be evidence for cognitive biases operating in language contact, thus demanding a more nuanced approach to the relationship between horizontal transmission and language universals.[21]

Conclusion

Wide agreement exists in linguistics for assuming that typological distributions are affected by preferences in language learning and processing to some degree. Different theoretical accounts vary in how much weigh they assign to cognitive biases.[8,12] Despite such differences it is clear that what links processing and typological distributions is language change. This connection between diachrony and typology has been entertained since Greenberg's early work, but typologists have been able to implement it only recently to modeling typological distributions.[3,17-19,65]

In this review I have surveyed experimental evidence for several statistical universals. As the body of relevant psycholinguistic work is increasing, the more evidence it seems to provide that typological generalizations may be reflected in cognitive processing preferences. Despite the traditional gap between typology and cognitive science the reviewed works propose fruitful possibilities for researching the processing-typology connection. However, since even the most thoroughly researched preferences (the subject preference) do not yet provide conclusive evidence about this connection, the issue is still somewhat open to what degree typological generalizations are grounded in cognitive biases. Future research on typologically more diverse languages is likely to provide a much needed more detailed picture on the issue.

Acknowledgements

The author thanks Balthasar Bickel for helpful comments when preparing the manuscript

References

1. Greenberg, JH. Some universals of grammar with particular reference to the order of meaningful elements. In Greenberg, JH, ed. Universals of Language, 2nd edition. MIT Press, Cambridge, MA, 1966, 73-113.
2. Dryer, MS. The Greenbergian word order correlations. Language 1992, 68:81-138. http://www.jstor.org/stable/416370.
3. Bickel, B. Typology in the 21st century: major current developments. Linguistic Typology 2007, 11:239-251. doi:10.1515/LINGTY.2007.018.
4. Hawkins, JA. Efficiency and Complexity in Grammars. Oxford University Press, Oxford, 2004.
5. Christiansen, MH, Chater, N. Language as shaped by the brain. Behavioral and Brain Sciences 2008, 31:489-558. doi:10.1017/S0140525X08004998.
6. Enrique-Arias, A. Accounting for the position of verbal agreement morphology with psycholinguistic and diachronic explanatory factors. Studies in Language 2002, 26:1-31. doi:10.1075/sl.26.1.02enr.
7. Bybee, J. Frequency of Use and the Organization of Language. Oxford University Press, Oxford, 2007.
8. Culbertson, J. Typological universals as reflections of biased learning: evidence from artificial language learning. Language and Linguistics Compass 2012, 6:310-329. doi:10.1002/lnc3.338.
9. Jaeger, TF, Tily, H. On language 'utility': processing complexity and communicative efficiency. WIREs Cognitive Science 2011, 2:323-335. doi:10.1002/wcs.126.
10. Jaeger, TF, Norcliffe, EJ. The cross-linguistic study of sentence production. Language and Linguistics Compass 2009, 3:866-887. doi:10.1111/j.1749-818x.2009.00147.x.
11. Trudgill, P. Sociolinguistic Typology: Social Determinants of Linguistic Complexity. Oxford University Press, Oxford, 2011.
12. Bybee, J. Language universals and usage-based theory. In Christiansen, MH, Collins, C, Edelman, S, eds. Language Universals. Oxford University Press, Oxford, 2009, 17-40. doi:10.1093/acprof:oso/9780195305432.003.0002.
13. Piantadosi, ST, Gibson, E. Quantitative standards for absolute linguistic universals. Cognitive Science 2013. doi:10.1111/cogs.12088.
14. Dryer, MS. Large linguistic areas and language sampling. Studies in Language 1989, 13:257-292.
15. Perkins, RD. Deixis, Grammar, and Culture. John Benjamins, Amsterdam, 1992.
16. Sinnemäki, K. Language Universals and Linguistic Complexity: Three Case Studies in Core Argument Marking. PhD dissertation, University of Helsinki; 2011.
17. Bickel, B. Distributional biases in language families. In Bickel, B, Grenoble, LA, Peterson, DA, Timberlake, A, eds. Language Typology and Historical Contingency: In Honor of Johanna Nichols. John Benjamins, Amsterdam and Philadelphia, 2013, 437-466.
18. Levinson, SC, Greenhill, SJ, Gray, RD, Dunn, M. Universal typological dependencies should be detectable in the history of language families. Linguistic Typology 2011, 15:509-534. doi:10.1515/LITY.2011.034.
19. Maslova, E. A dynamic approach to the verification of distributional universals. Linguistic Typology 2000, 4:307-333.

20. Sinnemäki, K. Word order in zero-marking languages. Studies in Language 2010, 34:869-912. doi:10.1075/sl.34.4.04sin.
21. Bickel, B. Distributional typology: statistical inquiries into the dynamics of linguistic diversity. In Heine, B, Narrog, H, eds. The Oxford Handbook of Linguistic Analysis, 2nd edition. Oxford University Press, Oxford, in press.
22. Chomsky, N. Aspects of the Theory of Syntax. MIT Press, Cambridge, MA, 1965.
23. Newmeyer, FJ. Possible and Probable Languages: A Generative Perspective on Linguistic Typology. Oxford University Press, Oxford, 2005.
24. Culbertson, J, Smolensky, P, Wilson, C. Cognitive biases, linguistic universals, and constraint-based grammar learning. Topics in Cognitive Science 2013, 5:392-424. doi:10.1111/tops.12027.
25. Lightfoot, D. The Development of Language: Acquisition, Change, and Evolution. Blackwell, Oxford, 1999.
26. Kiparsky, P. Universals constrain change; change results in typological generalizations. In Good, J, ed. Linguistic Universals and Language Change. Oxford University Press, Oxford, 2008, 23-53.
27. Reali, F, Griffiths, TL. The evolution of frequency distributions: relating regularization to inductive biases through iterated learning. Cognition 2009, 111:317-328. doi:10.1016/j.cognition.2009.02.012.
28. Ambridge, B, Pine, JM, Rowland, CF, Chang, F, Bidgood, A. The retreat from overgeneralization in child language acquisition: word learning, morphology, and verb argument structure. WIREs Cognitive Science 2013, 4:47-62. doi:10.1002/wcs.1207.
29. Diessel, H. New perspectives, theories and methods: diachronic change and language acquisition. In Bergs, A, Brinton, L, eds. Historical Linguistics of English: An International Handbook. Mouton de Gruyter, Berlin, 2012, 1599-1613.
30. Piantadosi, ST, Tily, H, Gibson, E. Word lengths are optimized for efficient communication. PNAS 2011, 108:3526-3529. doi:10.1073/pnas.1012551108.
31. Weskott, T, Hörnig, R, Fanselow, G, Kliegl, R. Contextual licensing of marked OVS word order in German. Linguistische Berichte 2011, 225:3-18.
32. Gibson, E. Linguistic complexity: locality of syntactic dependencies. Cognition 1998, 68:1-76.
33. Levy, R, Fedorenko, E, Gibson, E. The syntactic complexity of Russian relative clauses. Journal of Memory and Language 2013, 69:461-495. doi:10.1016/j.jml.2012.10.005.
34. Hudson Kam, CL, Newport, EL. Getting it right by getting it wrong: when learners change languages. Cognitive Psychology 2009, 59:30-66. doi:10.1016/j.cogpsych.2009.01.001.
35. Lewis, RL, Vasishth, S. An activation-based model of sentence processing as skilled memory retrieval. Cognitive Science 2005, 29: 1-45. doi:10.1207/s15516709cog0000_25.
36. Hale, J. The information conveyed by words in sentences. Journal of Psycholinguistic Research 2003, 32:101-123. doi:10.1023/A:1022492123056.
37. Levy, R. Expectation-based syntactic comprehension. Cognition 2008, 106:1126-1177. doi:doi:10.1016/j.cognition.2007.05.006.
38. Erdocia, K, Laka, I, Mestres-Missé, A, Rodriguez-Fornells, A. Syntactic complexity and ambiguity resolution in a free word order language: behavioural and electrophysiological evidences from Basque. Brain and Language 2009, 109:1-17. doi:10.1016/j.bandl.2008.12.003.

39. Bornkessel-Schlesewsky, I, Schlesewsky, M. The role of prominence information in the real-time comprehension of transitive constructions: a cross-linguistic approach. Language and Linguistics Compass 2009, 3:19-58. doi:10.1111/j.1749-818x.2008.00099.x.
40. Hyönä, J, Hujanen, J. Effects of case marking and word order on sentence parsing in Finnish: an eye fixation analysis. Quarterly Journal of Experimental Psychology 1997, 50A:841-858.
41. Kaiser, E, Trueswell, JC. The role of discourse context in the processing of a flexible word-order language. Cognition 2004, 94:113-147. doi:10.1016/j.cognition.2004.01.002.
42. Tanaka, MN, Branigan, HP, McLean, JF, Pickering, MJ. Conceptual influences on word order and voice in sentence production: Evidence from Japanese. Journal of Memory and Language, 2011, 65:318-330. doi:10.1016/j.jml.2011.04.009.
43. Grewe, T, Bornkessel, I, Zysset, S, Wiese, R, von Cramon, DY, Schlesewsky, M. Linguistic prominence and Broca's area: the influence of animacy as a linearization principle. Neuroimage 2006, 32:1395-1402. doi:10.1016/j.neuroimage.2006.04.213.
44. Wang, L, Schlesewsky, M, Philipp, M, Bornkessel-Schlesewsky, I. The role of animacy in online argument interpretation in Mandarin Chinese. In Lamers, M, de Swart, P, eds. Case, Word Order and Prominence: Interacting Cues in Language Production and Comprehension. Springer, Berlin, 2012, 91-119.
45. Bisang, W, Wang, L, Bornkessel-Schlesewsky, I. Subjecthood in Chinese: neurolinguistics meets typology. In Jing-Schmidt, Z, ed. Increased Empiricism: Recent Advances in Chinese Linguistics. John Benjamins, Amsterdam, 2013, 23-48.
46. Kiyama, S, Tamaoka, K, Kim, J, Koizumi, M. Effect of animacy on word order processing in Kaqchikel Maya. Open Journal of Modern Linguistics 2013, 3:203-207. doi:10.4236/ojml.2013.33027.
47. Keenan, EL, Comrie, B. Noun phrase accessibility and universal grammar. Linguistic Inquiry 1977, 8:63-99.
48. Carreiras, M, Duñabeitia, JA, Vergara, M, de la Cruz-Pavía, I, Laka, I. Subject relative clauses are not universally easier to process: evidence from Basque. Cognition 2010, 115:79-92. doi:10.1016/j.cognition.2009.11.012.
49. Hupp, JM, Sloutsky, VM, Culicover, PW. Evidence for a domain-general mechanism underlying the suffixation preference in language. Language and Cognitive Processes 2009, 24:876-909. doi:10.1080/01690960902719267.
50. Dryer, MS. Prefixing vs. suffixing in inflectional morphology. In Dryer, MS, Haspelmath, M, eds. The World Atlas of Language Structures Online. Max Planck Institute for Evolutionary Anthropology, Leipzig, 2013 (http://wals.info/chapter/26, accessed February 11, 2014).
51. Mohanan, T. Argument Structure in Hindi. CSLI Publications, Stanford, CA, 1994.
52. Aissen J. Differential object marking: iconicity vs. economy. Natural Languages and Linguistic Theory 2003, 21:435-483. doi:10.1023/A:1024109008573.
53. Bickel, B, Witzlack-Makarevich, A, Zakharko, T. Typological evidence against universal effects of referential scales on case alignment. In Bornkessel-Schlesewsky, I, Malchukov, AL, Richards, M, eds. Scales: A Cross-disciplinary Perspective on Referential Hierarchies. Mouton de Gruyter, Berlin, forthcoming.
54. Fedzechkina, M, Jaeger, TF, Newport, EL. Language learners restructure their input to facilitate efficient communication. PNAS 2012, 109:17897-17902. doi:10.1073/pnas.1215776109.

55. Fedzechkina, M, Jaeger, TF, Newport, EL. Functional biases in language learning: evidence from word order and case-marking interaction. In Carlson, L, Hoeschler, C, Shipley, TF, eds. Proceedings of the 33rd Annual Conference of the Cognitive Science Society. Cognitive Science Society, Austin, TX, 2011, 318-323.
56. Christiansen, MH. Using artificial language learning to study language evolution: exploring the emergence of word order universals. In Dessalles, JL, Ghadakpour, L, eds. The Evolution of Language: 3rd International Conference. Ecole Nationale Superieure des Telecommunications, Paris, 2000, 45-48.
57. Hulstijn, JH. Second language acquisition research in the laboratory: possibilities and limitations. Studies in Second Language Acquisition 1997, 19:131-143.
58. Rafferty, AN, Griffiths, TL, Ettlinger, M. Greater learnability is not sufficient to produce cultural universals. Cognition 2013, 129:70-87. doi:10.1016/j.cognition.2013.05.003.
59. Goldberg, AE. Substantive learning bias or an effect of familiarity? Comment on Culbertson, Smolensky, and Legendre (2012). Cognition 2013, 127:420-426. doi:10.1016/j.cognition.2013.02.017.
60. Hall, ML, Mayberry, RI, Ferreira, VS. Cognitive constraints on constituent order: evidence from elicited pantomime. Cognition 2013, 129:1-17. doi:10.1016/j.cognition.2013.05.004.
61. Gibson, E, Piantadosi, ST, Brink, K, Bergen, L, Lim, E, Saxe, R. A noisy-channel account of crosslinguistic word order variation. Psychological Science 2013, 24:1079-1088. doi: 10.1177/0956797612463705.
62. Stoll, S, Bickel, B. Capturing diversity in language acquisition research. In Bickel, B, Grenoble, LA, Peterson, DA, Timberlake, A, eds. Language Typology and Historical Contingency: In Honor of Johanna Nichols. John Benjamins, Amsterdam, 2013, 195-216.
63. Bickel, B, Nichols, J. Oceania, the Pacific Rim and the theory of linguistic areas. In Proceedings of the 32nd Annual Meeting of the Berkeley Linguistics Society. Berkeley Linguistics Society, Berkeley, 2006.
64. Dryer, MS, Haspelmath, M, eds. The World Atlas of Language Structures Online. Max Planck Institute for Evolutionary Anthropology, Leipzig, 2013 (http://wals.info, accessed February 11, 2014).
65. Levinson, SC, Gray, RD. Tools from evolutionary biology shed new light on the diversification of languages. Trends in Cognitive Sciences 2012, 16:167-173. doi:10.1016/j.tics.2012.01.007.
66. Sinnemäki, K. Complexity in core argument marking and population size. In Sampson, G, Gil, D, Trudgill, P, eds. Language Complexity as an Evolving Variable. Oxford University Press, Oxford, 2009, 125-140.
67. Lupyan, G, Dale, R. Language structure is partly determined by social structure. PLoS One 2010, 5: e8559. doi:10.1371/journal.pone.0008559.
68. Kusters, W. Linguistic Complexity: The Influence of Social Change on Verbal Inflection. LOT, Utrecht, 2003.
69. Bentz, C, Winter, B. Languages with more second language learners tend to lose nominal case. Language Dynamics and Change 2013, 3:1-27. doi:10.1163/22105832-13030105.
70. Roberts, S, Winters, J. Social structure and language structure: the new nomothetic approach. Psychology of Language and Communication 2012, 16:89-112. doi:10.2478/v10057-012-0008-6.

71. Evans, N, Levinson, SC. The myth of language universals: language diversity and its importance for cognitive science. Behavioral and Brain Sciences 2009, 32:429-492. doi:10.1017/S0140525X0999094X.
72. Cristofaro, S. Cognitive explanations, distributional evidence, and diachrony. Studies in Language 2012, 36:645-670. doi:10.1075/sl.36.3.07cri.
73. Bybee, J. The diachronic dimension in explanation. In Hawkins, JA, ed. Explaining Language Universals. Basil Blackwell, Oxford, 1988, 350-379.

Further Reading/Resources

For readers interested in more detailed accounts of how typological generalizations may be connected with processing preferences, I recommend the following articles.

The processing-typology model of John A. Hawkins: Hawkins, JA. Processing typology and why psychologists need to know about it. *New Ideas in Psychology* 2007, 25:87-107. doi:10.1016/j.newideapsych.2007.02.003.

A more advanced example of how to link typology and language processing. Bornkessel-Schlesewsky, I, Choudhary, KK, Witzlack-Makarevich, A, Bickel, B. Bridging the gap between processing preferences and typological distributions: initial evidence from the online comprehension of control constructions in Hindi. In Richards, M, Malchukov, AL, eds. *Scales*. University of Leipzig, Leipzig, 2008, 397-436.

An overview of how to connect language universals to diachrony and language processing. Moravcsik, E. Explaining language universals. In Song, JJ, ed. *The Oxford Handbook of Linguistic Typology*. Oxford University Press, Oxford, 2011, 69-89.