LETTERS TO *LANGUAGE*

Language accepts letters from readers that briefly and succinctly respond to or comment upon either material published previously in the journal or issues deemed of importance to the field. The editor reserves the right to edit letters as needed. Brief replies from relevant parties are included as warranted.

Length of review process

August 19, 2004

To the Editor:

I recently had the surprising and annoying experience of having to wait for twelve and a half months before I received a first decision after submitting a paper to a high-prestige linguistics journal. If I had known that the reviewing process can take that long for that journal, I would probably have submitted my paper elsewhere. So knowing the average reviewing time of a journal is important information. In *Language*, it is sometimes mentioned in the Editor's Department (and was recently said to be about four months). It would be good if other journals were similarly explicit about their reviewing speed. It would be even better, however, if readers didn't have to believe the Editor's average figures, but if they could see for themselves, that is, if every published article contained the first-decision date, in addition to the received and accepted dates. This would require very little additional work and almost no additional space, and if a leading journal like *Language* adopted this policy, it could become widespread practice. So please consider making this slight change to *Language*'s policies. Reviewing speed is an essential ingredient of journal quality. To assess journal quality in an objective way, readers need the full details.

Sincerely,

MARTIN HASPELMATH
[haspelmath@eva.mpg.de]

Editor's reply: The length of the review process and length of time to ultimate publication (which depends in part on the review process and in part on an author's ability to make revisions in a timely manner) are both of great concern to me. This suggestion seems an eminently reasonable one, and the added scrutiny that it affords into the process, as well as the very useful function of alerting readers to the timeliness of the published material, is certainly welcome. Readers should note that we have adopted this practice, beginning with this issue.

A nomenclatural desideratum

October 20, 2004

To the Editor:

In my book notice of Darya Kavitskaya's *Compensatory lengthening* (*Language* 80.3.626, 2004), I pointed out that Farsi and Persian are two names for the same language. In the course of my remarks, I also mentioned two language families: Gur and Voltaic. Not known to me then (and also unknown to Kavitskaya) was the fact that these two are also names for the same language family. Someone should publish a definitive list of all the languages/dialects/language phyla/families with alternate names, since no one linguist could possibly keep abreast of this. SIL International's *Ethnologue* (www.ethnologue.com) is the best source known to me that deals with this problem in nomenclature, but in looking up *Afroasiatic*, for example, we find only the older hyphenated designation *Afro-Asiatic*. The alternate names for this language family are not listed: *Hamito-Semitic*, *Semito-Hamitic*, *Lisramic*, *Erythraic* or *Erythrean*, *Afrasian*, and *Lamekhite* (in this connection, see Alan S. Kaye and Peter T. Daniels, 'Comparative Afroasiatic and general genetic linguistics', *Word* 43.3.429–58, 1992, especially pp. 430–32). Although

https://doi.org/10.1353/lan.2005.0021 Published online by Cambridge University Press

Ethnologue is certainly a fine beginning, what I am suggesting the field needs is a definitive, cross-referenced dictionary of languages/dialects and language families/phyla for all of the alternate names and the historical circumstances of their usages.

ALAN S. KAYE
[akaye@exchange.fullerton.edu]

Editor's reply: Sounds like an interesting idea to me; any volunteers?

Reply to Eska and Ringe

October 27, 2004

To the Editor:

The September 2004 issue of *Language* included an extended commentary by Joseph Eska and Don Ringe ('Recent work in computational linguistic phylogeny', *Language* 80.569–82) on a recent article by Peter Forster and Alfred Toth ('Toward a phylogenetic chronology of ancient Gaulish, Celtic, and Indo-European', *Proceedings of the National Academy of Sciences* 100.9079–84, 2003). Since we were not invited to reply to that commentary, we felt it appropriate to request that the Editor publish this short comment, since some basic clarifications on our article may be useful to the potential readers.

First, it is important to emphasize that our raw data table for Gaulish does not present our own translations, but Latin translations written by ancient scribes 2,000 years ago. We were careful to point out that the ancient scribes 'may not have been perfectly bilingual and may conceivably have distorted some of the language items' (supplementary website, www.pnas.org/cgi/content/full/1331158100/DC1/1). The reader should therefore not misinterpret our raw data table as an error-free dictionary.

Second, cognate coding is not the only valid approach to reconstructing language trees/networks. While we use cognate coding ourselves where appropriate (P. Forster, A. Toth, and H.-J. Bandelt, 'Evolutionary networks of word lists: Visualising the relationships between Alpine Romance languages', *Journal of Quantitative Linguistics* 5.3.174–87, 1998; P. Forster, T. Polzin, and A. Arne Röhl, 'Evolution of English basic vocabulary within the network of Germanic languages', *Phylogenetic methods and the prehistory of languages*, Cambridge: McDonald Institute, to appear, 2005), it appears unnecessarily dogmatic to declare this the only option, especially where the language records involved are fragmentary, and where the experts are unable to agree among themselves on translations, let alone on etymologies. Language researchers should not be discouraged from experimenting with similarity coding wherever this may appear to be a promising direction.

Third, it may be helpful to summarize how our phylogenetic approach dealt with the raw data. As explained above, in the Gaulish application we had to assume that two substantial sources of error were present in the raw data. Our approach therefore included an error test on the initial data. Specifically, our Basque negative control revealed an error rate of about 5/35 characters for the 490 entries in the table; a proportion of these errors are then eliminated prior to the final analysis via the multistate cutoff. Further errors are resolved via the network construction itself (enhanced error detection is one of the strengths of network approaches over tree approaches). Remaining errors ideally are correctly repositioned by the network; other errors are highlighted as reticulations, and some may remain undetected by the network. Why did we delegate error detection to a late stage rather than to an early stage in the analysis? The answer is that we wished to avoid ascertainment bias in data compilation and coding (for details see the supplementary website noted above).

Despite the substantial difficulties connected with Gaulish records, our resulting

https://doi.org/10.1353/lan.2005.0021 Published online by Cambridge University Press

phylogeny in the published Figure 3 (p. 9083) seems to make sense: the Romance languages group together, the Celtic languages group together, Greek and English are on separate branches, and most impressively, the ancient languages are correctly placed in positions ancestral to their modern counterparts. Another powerful application of phylogenetics to language is the identification of potential cognates: For example, we consciously gave Latin *filia* and Spanish *hija* separate codings, and the resulting network immediately suggested that *hija* derives from *filia*. Those who are interested in why our 'DNA' approach may work for languages are referred to the online tutorial (www.mcdonald.cam.ac.uk/genetics/gaulish_tutorial.pdf) and especially to our ongoing work (see the Forster et al. 2005 reference above) which exploits the new capabilities of the software package Network 4.1. Continuously updated versions are available for free at www.fluxus-engineering.com.

DR. PETER FORSTER
[pf223@cam.ac.uk]

Authors' reply: Forster's response, on behalf of himself and Toth, to our 2004 critique of their work on the phylogeny of Celtic languages fails to address any of our criticisms concerning their methodology; he does, however, misrepresent us on several points. For instance, we do not state that cognation is the only acceptable basis for coding; we do say that it is objective, replicable, and scientific, whereas their judgments are subjective, nonreplicable, and unscientific. He also repeats some claims that we showed to be false; for instance, they rely not only on contemporary Latin translations of ancient Celtic texts, but also on modern interpretations, some of which are demonstrably wrong. We still do not see how their failure to weed errors out of a database of only thirty-five items is defensible; if they are really obsessed with as certainment bias, they should have used every piece of reliably analyzed evidence available, which would have given them a data set many times larger and some hope of making progress. Having failed to get the data right and to analyze them scientifically, they have now also failed to defend their approach, in our view. We do not see how this can lead to any expectation of ultimate success.

JOSEPH F. ESKA
[eska@vt.edu]
DON RINGE
[dringe@unagi.cis.upenn.edu]

'The future of language' (*Science*)*

November 18, 2004

To the Editor:

In a recent Viewpoint article appearing in *Science* ('The future of language', *Science* 303.1329–31, 2004), David Graddol assesses the development and future prospects of Chomsky's linguistic theory as follows:

'When Noam Chomsky published his groundbreaking book on Syntactic Structures in 1957. . . , syntax became regarded as the central problem in linguistic description. But in the future, we may come to appreciate how far the Chomskyan approach has led linguists down a blind alley. Over the last half-century, theories of syntax have lost touch with language as spoken by people in the real world, and have retreated into abstract studies of universal features of human cognition.' (p. 1330)

Although abstraction is hard to quantify, and Graddol's assessment of the last half century cites among Chomsky's voluminous work over this period only the 1957 *Syntactic structures* (The Hague: Mouton), we provisionally agree with Graddol's claim that theories of syntax have become progressively more abstract in the last fifty years. However, far from being a 'retreat' 'down a blind alley', Chomsky's approach continues to

https://doi.org/10.1353/lan.2005.0021 Published online by Cambridge University Press

enable linguists to productively investigate two central formal problems concerning human knowledge of language and its use.

Compared to all other forms of animal communication (see Marc D. Hauser, Noam Chomsky, and W. Tecumseh Fitch, 'The faculty of language: What is it, who has it, and how did it evolve?', *Science* 298.1569–79, 2002), human language is a highly structured formal combinatorial system and, in addition, the number of discrete well-formed sentences generated by the system is infinite.

This property of discrete infinity characterizes EVERY human language; none consist of a finite set of sentences. The unchanged central goal of linguistic theory over the last fifty years has been and remains to give a precise, formal characterization of this property and then to explain how humans develop (or grow) and use discretely infinite linguistic systems. The first detailed specification of the problem, a framework for inquiry, and a route to a solution were outlined by Chomsky in the mid 1950s, placing this mode of linguistic inquiry in a psycho-biological setting. Humans (barring pathology) come to know the natural language known by members of the community in which they are reared. What is human linguistic knowledge, and how does it develop in the individual? The mature and arguably uniquely human linguistic capacity allows us to produce and comprehend an unbounded (i.e. practically speaking, infinite) number of sentences. Given that each natural language has an infinite number of DIFFERENT kinds of well-formed sentences that manifest an unbounded number of DISTINCT PATTERNS and given that native speakers come to have the capacity to analyze and interpret them all, it appears that they do this by having acquired a finite set of abstract (algebraic) RULES that can recursively apply to yield this infinite set of well-formed sentences with their different patterns. Being cognitively endowed with such rules (coupled with processing mechanisms) you can (effortlessly and unconsciously) analyze and interpret previously unencountered (hence not stored) linguistic stimuli, such as the sentence you are now reading and understanding.

Graddol in fact seems to embrace a similar position in saying that 'The human brain is able to store experience of how words pattern' (p. 1330). A central concern then must be to formally specify what these (presumably abstract) word PATTERNS are, and how linguistic 'experience' is codified by human beings and gives rise to these abstract patterns stored in the brain (to paraphrase Graddol). Once one notes that the number of distinct patterns a language displays is open-ended, the specification of 'how words pattern' can be (and, by hypothesis, is) specified by a finite set of rules (i.e. a formal grammar) able to generate this infinite set of patterns. Indeed, this is what syntactic theory aims to do. Graddol, in fact, seems to embrace a position differing little, if any, from the Chomskyan perspective he disparages in writing: 'It seems that much of what we have expected of grammars can be better explained by focusing on words and the complex way in which they keep each other's company' (ibid.).

Linguists have discovered that this 'complex way of keeping company' that words manifest is indeed far from simple and constitutes a formal and empirical problem forming a key component of any overall account of human knowledge of language.

Guided by this Chomskyan perspective, linguists have discovered, through various forms of experimentation, that the class of possible human grammars is highly constrained (presumably by human biology), and they have, over the last fifty years, formulated increasingly refined, sophisticated, and arguably abstract grammars. Rather than being signs of retreat as Graddol suggests, the abstraction of current

https://doi.org/10.1353/lan.2005.0021 Published online by Cambridge University Press

theorizing has gone hand in hand with increasing the sophistication and subtlety of the questions being asked and the domains susceptible to productive inquiry.

But hasn't this (ongoing) 'abstraction' caused us to 'lo[se] touch with language as spoken by people in the real world' (ibid.), as Graddol claims? Not at all. In fact, so far as we can tell, everyone who studies what people actually do (including Graddol) in fact presupposes that individuals possess these grammatical resources, that is, they know the language. After all, speakers and listeners 'in the real world' can and do produce and understand sentences never before encountered. How if not via rules (or Graddol's unspecified but clearly abstract 'stored patterns') that enable this combinatorial creativity? Speakers and listeners 'in the real world' know that sentences express different thoughts depending on how they are put together (or depending on how put together they are!). How is this possible except via the possession of grammars—that is, specifications of sound-syntax-semantics relations? How speakers and listeners 'in the real world' USE their language is partially a function of the grammar of that language, so the 'abstract study' of this grammar seems to us (as Chomsky has repeatedly stressed) a logically ineliminable aspect of any study of what people do, such as 'speaking in the real world'.

As Elaine McNulty (p.c.) observes, one sympathetic to Graddol's overall anti-abstractionism could correspondingly argue, oddly, that contemporary theories of the abstract representations generated by the human visual system have lost touch with what is (actually) seen by people in the real world. Yet any account of what people actually see (their friend Bob, shrubbery, a skyscraper) will be couched in terms of a person USING their human visual system, as the system is specified within an (abstract) general (shrubbery-independent) theory of human vision.

Despite humans being part of the physical world, in human inquiry into human properties there is a disturbing tendency to eschew the methods that have proven themselves so successful in the physical sciences. Chief among these is what Steven Weinberg ('The forces of nature', *Bulletin of the American Academy of Arts and Sciences* 29.4.28–29, 1976) has called the 'Galilean style' wherein nature is investigated by making abstract models, the investigation of whose properties illuminates the structures of 'the real world'. Graddol seems to find this sort of enterprise suspect (although he implicitly embraces it, as noted). But if the history of science is any guide, it would seem that the willingness to abstract is necessary if one is to go beyond superficial observations and get to what Weinberg described as 'the roots of explanation of the natural world'.

The scientific approach, including 'abstraction' (championed in the psycho-biological linguistic domain by Chomsky) is ineliminable from productive 'language study'. As Chomsky ('Three factors in language design', *Linguistic Inquiry*, to appear, 2005) notes, even those who explicitly and strenuously reject it, routinely adopt it implicitly (Graddol being an example, as we note). Indeed, we should expect our theories to become more 'abstract' as they attain greater explanatory depth.

SAM EPSTEIN
[sepstein@umich.edu]
NORBERT HORNSTEIN
[nhornste@umd.edu]

Notes

*We wish to acknowledge the following people for invaluable discussion of the issues discussed here: Noam Chomsky, Elaine McNulty, Daniel Seely, Richard Lewis, Acrisio Pires, and Pam Beddor. We also thank Brian Joseph for his interest in providing a forum for expressing our perspective. None of them necessarily subscribes wholly or in part to any of the views expressed here. An earlier version of this letter

https://doi.org/10.1353/lan.2005.0021 Published online by Cambridge University Press

was submitted to *Science* Letters to the Editor, in the hope of responding to Graddol's article in the very journal in which it appeared. Our letter was rejected by *Science*, with the explanation that they receive many more letters than they have space to publish.

'*Most*'—A correction

December 15, 2004

To the Editor:

I would like to correct a minor point in my article '*Most*' (*Language* 80.658 – 706, 2004). The ambiguity analysis of *most* that I attributed to Philip L. Peterson ('On the logic of ''few'', ''many'', and ''most'' ', *Notre Dame Journal of Formal Logic* 20.155–79, 1979) was based on a misunderstanding, as Laurence Horn (p.c.) has pointed out to me. My arguments against an ambiguity analysis between a lower-bounded-only meaning and a bilateral, 'all'-excluding meaning for *most* should be taken instead as arguments against a possible theoretical stand on scalar quantifiers (such was Sir William Hamilton's proposal for *some* (*Lectures on logic*, Edinburgh: Blackwood and Sons, 1860)). But whereas I unjustifiably criticized Peterson on an ambiguity he did not posit, I should explicitly criticize him for the untenable ambiguity analysis he did propose. According to Peterson, *most* is semantically ambiguous between the standard lower-bounded-only meaning and a meaning whereby the majority not just exceeds the complement minority, but GREATLY exceeds it. In effect, my analysis of *most* as implicating a noteworthy amount shows that this interpretation must be implicated rather than entailed. First, a noteworthy majority need not be so large objectively (see my ex. 38, p. 691, about a 50.6% noteworthy majority). Second, even when subjects interpreted *most* as a large majority (specifically, larger than *more than half*—see ex. 37, p. 690), they nonetheless did not view this interpretation as entailed, as is evident from their truth-value judgments in case the 'large majority' interpretation turned out to be false (see the discussion in §6.6, p. 697ff.).

MIRA ARIEL
[mariel@post.tau.ac.il]

https://doi.org/10.1353/lan.2005.0021 Published online by Cambridge University Press