Obituary

Ivan A. Sag

Emily M. Bender
University of Washington

Ivan Sag died on September 10, 2013, after a long illness. He is survived by his wife, Stanford sociolinguistics professor Penny Eckert. In a career spanning four decades, he published over 100 articles and 10 books, centered on the theme of developing precise, implementable, testable, psychologically plausible, and scalable models of natural language, especially syntax and semantics.

Sag earned a B.A. from the University of Rochester (1971), an M.A. from the University of Pennsylvania (1973) and a Ph.D. from MIT (1976), all in Linguistics. He held teaching positions at the University of Pennsylvania (1976–1979) and Stanford University (1979–2013) where he was the Sadie Dunham Patek Professor in Humanities since 2008. In addition, he taught at ten LSA Linguistic Institutes (most recently as the Edward Sapir Professor at the Linguistic Institute at the University of Colorado, Boulder, in 2011), at five ESSLLIs, as well as in summer schools or other visiting positions at NTNU (Trondheim), Université de Paris 7, Rijksuniversiteit Utrecht, the University of Rochester, the University of Chicago, and Harvard University. At Stanford he was a founding member of and active participant in the Center for the Study of Language and Information (CSLI), which housed the HPSG and LinGO projects. He was also a key member of the group of faculty that developed and launched the Symbolic Systems program (in 1985) and was director of Symbolic Systems from 2000–2001 and 2005–2009. Among other honors, he was elected to the American Academy of Arts and Sciences in 2007 and named a Fellow of the Linguistic Society of America in 2008.

Sag's advisor was Noam Chomsky; throughout his career, he saw himself as furthering what he understood to be the original Chomskyan enterprise. However, in the late 1970s, he broke with the Chomskyan mainstream because he felt it had abandoned central aspects of that original enterprise. Here's how Sag told the story to Ta! magazine in 1993:

> Well, it has always been hard for me to reconcile the original Chomskyan research goals and methods with most of the modern work that goes on. Though Chomsky has denied this. Current work in so-called Government and Binding Theory is basically formulating ideas in ways that are so loose that they do not add up to precisely constructing hypotheses about the nature of language.
>
> All too often, people throw around formalisms that have no precise interpretation and the consequences of particular proposals are absolutely impossible to assess. In my opinion that is just not the way to do science. I think that the original goals of generative grammar do constitute a set of desiderata for the science of language that one can try to execute with much greater success than current work in GB has achieved or is likely to, given the directions it seems to be going in.

The result of Sag holding to the initial goals of generative grammar, even when mainstream syntax did not, has been an enormous boon for the field of computational linguistics. Whereas much mainstream work in theoretical syntax is neither explicitly formalized nor concerned with broad coverage, the frameworks that Sag was instrumental in helping to create (Generalized Phrase Structure Grammar, Head-driven Phrase Structure Grammar, and Sign-Based Construction Grammar [Boas and Sag 2012], but especially HPSG) are implementable and in fact implemented, and demonstrably scalable.

Sag first encountered the community working on what would come to be called Generalized Phrase Structure Grammar (GPSG, a term coined by Sag), and in particular Gerald Gazdar and Geoff Pullum, at the 1978 LSA Linguistic Institute. Gazdar and colleagues set out to show that English (and other natural languages) could in fact be described with context-free models, as Pullum and Gazdar (1982) had debunked all previous arguments against that claim. But more importantly Sag and his colleagues developing GPSG strove to be formally precise, in order to support valid scientific investigation.[1] The GPSG book (Gazdar et al. 1985) begins by throwing down the gauntlet:

> This book contains a fairly complete exposition of a general theory of grammar that we have worked out in detail over the past four years. Unlike much theoretical linguistics, it lays considerable stress on detailed specifications both of the theory and of the descriptions of parts of English grammar that we use to illustrate the theory. We do not believe that the working out of such details can be dismissed as 'a matter of execution', to be left to lab assistants. In serious work, one cannot 'assume some version of the X-bar theory' or conjecture that a 'suitable' set of interpretive rules will do something as desired, any more than one can evade the entire enterprise of generative grammar by announcing: 'We assume some recursive function that assigns to each grammatical and meaningful sentence of English an appropriate structure and interpretation.' One must set about constructing such a function, or one is not in the business of theoretical linguistics. (p. ix)

The computational benefits of that precision were quickly apparent. In 1981, Sag taught a course on GPSG at Stanford with Gazdar and Pullum. One of the students attending that course, Anne Paulson, was working at Hewlett-Packard Labs and saw the potential for using GPSG as the basis of a question answering system (with a database back-end). Paulson arranged a meeting between her boss, Egon Loebner, and Sag, Gazdar, Pullum, and Tom Wasow, which led to a nearly decade-long project implementing a grammar for English and processing tools to work with it. The project included HP staff as well as Sag, Pullum, and Wasow as consultants, and Stanford and UC Berkeley students, including Mark Gawron, Carl Pollard, and Dan Flickinger.

The work initially set out to implement GPSG (Gawron et al. 1982), but in the context of developing and implementing analyses, Sag and colleagues added innovations to the underlying theory until its formal basis was so different it warranted a new name. The new theory, laid out in Pollard and Sag 1987 and Pollard and Sag 1994, among others, came to be called Head-driven Phrase Structure Grammar (HPSG).

---

1 The formal precision had much more lasting impact than the attempt at a context-free model: Eventually Bresnan et al. (1982), Culy (1985), and Shieber (1985) found more solid counterexamples to the claim that natural languages can be modeled with CF-PSGs.

HPSG synthesizes ideas from GPSG, Dependency Grammar (Hudson 1984), Categorial Grammar (Wood 1993), Lexical Functional Grammar (Bresnan and Kaplan 1982), and even Government and Binding Theory (Chomsky 1981).

Importantly, rather than encoding theoretical results as constraints on the formalism, HPSG defines a flexible formalism (typed feature structures) in which different theories can be defined. This flexibility facilitates testing and synthesis of theoretical ideas developed in other frameworks. The stability of the formalism has been critical to the success of HPSG in computational linguistics, as it has allowed for the development of a variety of processing engines that interpret the formalism and thus can apply grammars to the tasks of parsing and generation (Uszkoreit et al. 1994; Carpenter and Penn 1994; Makino et al. 1998; Copestake 2002; Callmeier 2002; Penn 2004; Crysmann and Packard 2012; Slayden 2012, *inter alios*).

HPSG-based parsing was deployed in Verbmobil (Wahlster 2000), a large-scale, multi-site machine translation project funded by the German government, for which Sag headed up the English Grammar Project (1994–2000), later redubbed LinGO (Linguistic Grammars Online), at CSLI. The English grammar developed in that project (beginning actually in 1993) came to be known as the English Resource Grammar (Flickinger 2000, 2011).[2] The resource grammar idea builds on insights articulated in Gawron et al. 1982, namely, that detailed syntactic and semantic analysis are a crucial component of natural language understanding (in addition to discourse and world knowledge processing) and that the grammar which does that analysis can and should be portable across domains.

The English Resource Grammar has been refined and extended as it has been developed in the context of applications ranging from machine translation of dialogues regarding travel and appointment scheduling (Verbmobil; Wahlster 2000), automated customer service response (YY Technologies), machine translation of Norwegian hiking brochures (LOGON; Oepen et al. 2007), and grammar checking as part of a language arts instructional application (EPGY/Redbird; Suppes et al. 2012). The work of building a grammar such as this involves identifying phenomena in sentences from the domain of interest that the grammar does not yet account for, delimiting the phenomena, and developing and implementing analyses. Throughout the mid and late 1990s and into the early 2000s, the LinGO project at CSLI featured weekly project meetings led by Flickinger, who would bring phenomena in need of analysis for discussion by the group, including Sag and Wasow as well as Ann Copestake, Rob Malouf, Stanford linguistics graduate students, and visitors to CSLI.

In these always lively discussions, Sag could be counted on to share his encyclopedic knowledge of theoretical literature pertaining to the phenomenon in question and key examples that had been identified and analyzed in that literature, to suggest analyses, as well as to invent on the spot further examples to illustrate differences in predictions of competing candidate analyses. Supporting his ability to do this was an unsurpassed command of the theory and the workings of the grammar written in it.[3] These meetings not only fed the development of the English Resource Grammar, but

---

2 The initial grammar was developed by Rob Malouf, then a grad student at Stanford. Since 1994 Dan Flickinger has been its lead developer.

3 GPSG and HPSG were developed on the premise that precise formalization is critical to the testing of linguistic hypotheses: Without the formalization, it is not possible to say for certain what the predictions of a theory are. In fact, formalization alone isn't enough: Grammars of the scale supported by the HPSG framework are too complex for humans to reliably do those calculations without the aid of a machine. The one possible exception to this generalization was Ivan Sag.

also formed an important point of contact between computational and theoretical work in HPSG, such that the "pen and paper" theory remained responsive to computational concerns.

Another key result of the LinGO project during the Verbmobil days was the development of Minimal Recursion Semantics (MRS) (Copestake et al. 1995, 2005). Sag and colleagues designed MRS to meet the competing demands of expressive adequacy, grammatical compatibility, computational tractability, and underspecifiability. In other words, it is a computational semantic formalism that allows grammars like the English Resource Grammar to make explicit exactly as much information about semantic predicate argument structure and quantifier and operator scope as is determined by sentence structure, leaving further ambiguity represented via underspecification rather than enumeration of, for example, the full set of possible quantifier scopings for each item.

The experience of the Verbmobil project highlighted the value of collaborative work on natural language processing between grammarians and those working on the software required to develop and deploy grammars in practical applications. In the late 1990s, Sag and others, notably Dan Flickinger, Stephan Oepen, Jun'ichi Tsujii, and Hans Uszkoreit, began exploring ways to continue the collaboration past the end of the Verbmobil project. The result was the DELPH-IN consortium,[4] which has continued to facilitate further exchange of ideas and joint development of software and grammars. As of 2013, DELPH-IN spans four continents and has developed opensource grammars (including some large scale ones) for many languages, all of which are interoperable with the same set of open-source processing (analysis and generation) tools.

A hallmark of Sag's work has been his openness to collaboration across frameworks, subfields, and disciplines. In addition to the interactions he facilitated between linguistics and computational linguistics, Sag also built bridges between formal linguistics and psycholinguistics. Towards the end of his career, together with graduate students at Stanford, Sag turned his attention to developing a clearer understanding of the roles of processing and grammar in determining (un)acceptability. Specifically, he argued that the unacceptability associated with phenomena such as "island constraints" derives not from ungrammaticality but in fact from more general processing pressures (Staum and Sag 2008; Hofmeister and Sag 2010; Staum Casasanto, Hofmeister, and Sag 2010; Hofmeister et al. 2013). More generally, Sag's work on grammatical theory has highlighted the importance of "performance-plausible competence grammar," namely, models of linguistic knowledge that can in fact be embedded in models of human language processing (Sag 1992; Sag and Fodor 1994; Sag, Wasow, and Bender 2003; Sag and Wasow 2011). A key feature of HPSG that makes it "performance plausible" is representation of grammatical information in a form that allows it to be combined flexibly with other grammatical information as well as extragrammatical information relevant to processing (e.g., general world knowledge, or information about the present situation perceived by the hearer) as that information becomes available. This form of grammatical description has been very influential in computational work on incremental processing models which explicitly account for the generation of upcoming grammatical structure (Jurafsky 1996; Hale 2001; Levy 2008).

---

4 Deep Linguistic Processing in HPSG Initiative; `http://www.delph-in.net`.

Where many saw opposition between stochastic and symbolic methods, Sag saw opportunity for hybridization, as early as 1993. When asked by Ta! magazine (in the same interview cited above) about connectionism, he replied:

> If you think of grammar as a transformational grammar, and you look at the connectionist models that have evolved, then it is like looking at apples and oranges. However, if you look at a theory of typed feature structures and a connectionist model, it is more like apples and applesauce....
>
> I am not just interested in computer natural language processing. I am interested in understanding how in the world communication is possible at all. To me, it is astonishing that the huge space of ambiguity and uncertain information that language presents somehow gives rise to accurate and efficient communication. What we do cognitively, in language processing, is nothing short of miraculous. We bring together knowledge of the language, knowledge of the world, knowledge of the subject matter, knowledge of the situation, in such a way as to never even see the landscape of what the linguistic possibilities are....
>
> Now is not the time to stop exploring discrete methodologies, but rather to look for hybrid methodologies that exploit the complementary strengths of discrete and statistical methodologies. That is the only way that language technology will ever develop to play the role it must in the technology of tomorrow.

Finally, no appreciation of Sag's career and its impact on the fields of computational linguistics and especially linguistics would be complete without remarking on the energy and verve he brought to the social aspects of doing science: He had a keen understanding of the importance of community and communication in the doing of science, and worked tirelessly to promote both. He organized the first International Conference on Head-Driven Phrase Structure Grammar in Columbus, Ohio, in 1993 in order to provide a forum for HPSG researchers scattered across the globe to exchange ideas; the 20th iteration of that conference was held in Berlin in 2013. But his notion of community extended beyond staid academic discourse: He also started the band Dead Tongues with Geoffrey Nunberg in the early 1980s, and played keyboards in every iteration of that band (with ever-changing, but always linguistically motivated, membership) through the roof-raising performance at Ivan Fest at CSLI in April 2013. He also organized the rental of sorority houses and the hiring of gourmet chefs for shared housing at every Linguistic Institute he attended. These opportunities for social engagement helped build the community, not just of HPSG researchers, but of linguists more generally, which in turn supports the lively exchange of ideas on which Sag thrived and to which he contributed so much.

References

Boas, Hans Christian and Ivan A Sag. 2012. *Sign-Based Construction Grammar*. CSLI Publications, Stanford, CA.

Bresnan, Joan and Ronald M. Kaplan. 1982. Lexical-Functional Grammar: A formal system for grammatical representation. In Joan Bresnan, editor, *The Mental Representation of Grammatical Relations*, pp. 29–130, MIT Press, Cambridge, MA.

Bresnan, Joan, Ronald M. Kaplan, Stanley Peters, and Annie Zaenen. 1982. Cross-serial dependencies in Dutch. *Linguistic Inquiry*, 13(4):613–635.

Callmeier, Ulrich. 2002. Preprocessing and encoding techniques in PET. In Stephan Oepen, Daniel Flickinger, J. Tsujii, and Hans Uszkoreit, editors, *Collaborative Language Engineering. A Case Study in Efficient Grammar-based Processing*, pp. 127–140, CSLI Publications, Stanford, CA.

Carpenter, Bob and Gerald Penn. 1994. ALE: The attribute logic engine user's guide, version 2.0.1. Carnegie Mellon University, Department of Philosophy, Paper 526.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Foris Publications, Dordrecht, Holland.

Copestake, Ann. 2002. *Implementing Typed Feature Structure Grammars*. CSLI Publications, Stanford, CA.

Copestake, Ann, Dan Flickinger, Rob Malouf, Susanne Riehemann, and Ivan Sag. 1995. Translation using minimal recursion semantics. In *Proceedings of the Sixth International Conference on Theoretical and Methodological Issues in Machine Translation*, pp. 15–32, Leuven.

Copestake, Ann, Dan Flickinger, Carl Pollard, and Ivan A. Sag. 2005. Minimal recursion semantics: An introduction. *Research on Language & Computation*, 3(4):281–332.

Crysmann, Berthold and Woodley Packard. 2012. Towards efficient HPSG generation for German, a non-configurational language. In *Proceedings of COLING 2012*, pages 695–710, Mumbai.

Culy, Christopher. 1985. The complexity of the vocabulary of bambara. *Linguistics and Philosophy*, 8(3):345–351.

Flickinger, Dan. 2000. On building a more efficient grammar by exploiting types. *Natural Language Engineering*, 6 (1) (Special Issue on Efficient Processing with HPSG):15 – 28.

Flickinger, Dan. 2011. Accuracy v. robustness in grammar engineering. In Emily M. Bender and Jennifer E. Arnold, editors, *Language from a Cognitive Perspective: Grammar, Usage and Processing*. CSLI Publications, Stanford, CA, pages 31–50.

Gawron, Jean Mark, Jonathan King, John Lamping, Egon Loebner, E. Anne Paulson, Geoffrey K. Pullum, Ivan A. Sag, and Thomas Wasow. 1982. Processing English with a Generalized Phrase Structure Grammar. In *Proceedings of the 20th Annual Meeting of the Association for Computational Linguistics*, pages 74–81, Toronto.

Gazdar, Gerald, Ewan Klein, Geoffrey Pullum, and Ivan Sag. 1985. *Generalized Phrase Structure Grammar*. Harvard University Press, Cambridge, MA.

Hale, John. 2001. A probabilistic Earley parser as a psycholinguistic model. In *Proceedings of the Second Meeting of the North American Chapter of the Association for Computational Linguistics on Language Technologies*, pages 1–8, Pittsburgh, PA.

Hofmeister, Philip, T Florian Jaeger, Inbal Arnon, Ivan A. Sag, and Neal Snider. 2013. The source ambiguity problem: Distinguishing the effects of grammar and processing on acceptability judgments. *Language and Cognitive Processes*, 28(1-2):48–87.

Hofmeister, Philip and Ivan A. Sag. 2010. Cognitive constraints and island effects. *Language*, 86(2):366–415.

Hudson, Richard. 1984. *Word Grammar*. Blackwell, Oxford.

Jurafsky, Daniel. 1996. A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, 20:137–194.

Levy, Roger. 2008. Expectation-based syntactic comprehension. *Cognition*, 106(3):1126–1177.

Makino, Takaki, Minoru Yoshida, Kentaro Torisawa, and J. Tsujii. 1998. LiLFeS — towards a practical HPSG parser. In *Proceedings of the 17th International Conference on Computational Linguistics and the 36th Annual Meeting of the Association for Computational Linguistics*, pages 807–811, Montreal.

Oepen, Stephan, Erik Velldal, Jan Tore Lønning, Paul Meurer, Victoria Rosén, and Dan Flickinger. 2007. Towards hybrid quality-oriented machine translation. On linguistics and probabilities in MT. In *11th International Conference on Theoretical and Methodological Issues in Machine Translation (TMI-07)*, pages 144–153, Skövde.

Penn, Gerald. 2004. Balancing clarity and efficiency in typed feature logic through delaying. In *Proceedings of the 42nd Meeting of the Association for Computational Linguistics (ACL'04), Main Volume*, pages 239–246, Barcelona.

Pollard, Carl and Ivan A. Sag. 1987. *Information-Based Syntax and Semantics. Volume 1: Fundamentals*. CSLI Lecture Notes # 13. Center for the Study of Language and Information, Chicago, IL and Stanford, CA. Distributed by the University of Chicago Press.

Pollard, Carl and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Studies in Contemporary Linguistics. The University of Chicago Press and CSLI Publications, Chicago, IL, and Stanford, CA.

Pullum, Geoffrey K. and Gerald Gazdar. 1982. Natural languages and context-free languages. *Linguistics and Philosophy*, 4:471–504.

Sag, Ivan A. 1992. Taking performance seriously. *VII Congreso de Languajes Naturales y Lenguajes Formales*, pages 61–74.

Sag, Ivan A. 1993. Interview with Anne-Marie Mineur and Gerrit Rentier. *Ta!*, 2(2).

Sag, Ivan A. and Janet D. Fodor. 1994.
Extraction without traces. In *West Coast
Conference on Formal Linguistics*,
volume 13, pages 365–384,
Los Angeles, CA.

Sag, Ivan A. and Thomas Wasow. 2011.
Performance-compatible competence
grammar. *Non-transformational Syntax:
Formal and Explicit Models of Grammar*,
pages 359–377.

Sag, Ivan A., Thomas Wasow, and Emily M.
Bender. 2003. *Syntactic Theory: A Formal
Introduction*. CSLI, Stanford, CA,
second edition.

Shieber, Stuart M. 1985. Evidence against the
context-freeness of natural language.
Linguistics and Philosophy, 8(3):333–343.

Slayden, Glenn C. 2012. Array TFS storage
for unification grammars. Master's thesis,
University of Washington.

Staum, Laura and Ivan A. Sag. 2008.
The advantage of the ungrammatical.
In *Proceedings of the 30th Annual Meeting of
the Cognitive Science Society*, pages 601–606,
Washington, DC.

Staum Casasanto, Laura, Philip Hofmeister,
and Ivan A. Sag. 2010. Understanding
acceptability judgments: Distinguishing

the effects of grammar and processing on
acceptability judgments. In *Proceedings
of the 32nd Annual Conference of the
Cognitive Science Society*, pages 224–229,
Portland, OR.

Suppes, P., D. Flickinger, B. Macken, J. Cook,
and T. Liang. 2012. Description of the
EPGY Stanford University online courses
for mathematics and language arts.
In *International Society for Technology in
Education (ISTE) Annual 2012 Conference*,
pages 1–9, San Diego, CA.

Uszkoreit, Hans, Rolf Backofen, Stephan
Busemann, Abdel Kader Diagne,
Elizabeth A. Hinkelman, Walter Kasper,
Bernd Kiefer, Hans-Ulrich Krieger, Klaus
Netter, Günter Neumann, Stephan Oepen,
and Stephen P. Spackman. 1994. DISCO—
an HPSG-based NLP system and its
application for appointment scheduling.
In *Proceedings of the 15th International
Conference on Computational Linguistics*,
pages 436–440, Kyoto.

Wahlster, Wolfgang, editor. 2000. *Verbmobil.
Foundations of Speech-to-Speech Translation*.
Springer, Berlin, Germany.

Wood, Mary McGee. 1993. *Categorial
Grammars*. Routledge.