

# The Subset Principle: Consequences and Conspiracies<sup>1</sup>

**William Gregory Sakas**

Department of Computer Science, Hunter College

Ph.D. Programs in Linguistics and Computer Science, The Graduate Center  
City University of New York

Gold (1967) showed that the Subset Principle (*SP*), though not by that name, is necessary (but not sufficient) for learning from positive data. *SP* is most familiar to linguists through the work of Baker (1979), Pinker (1979), Dell (1981), Berwick (1985), Manzini & Wexler (1987), Clark (1992), and others. The penalty for not obeying *SP* is very well understood. In this talk, we examine problems that emerge as we begin to think about how to apply *SP* in various contexts.

We begin by adopting the psychologically attractive assumption that the learning mechanism (*LM*) is memoryless; during the course of learning *LM* has no ability to recall past hypotheses that were entertained<sup>2</sup> or prior sentences that were encountered. Given this memoryless assumption, we present a safe definition of *SP*:

***SP*:** When *LM*'s current language is incompatible with a new input sentence *s*, *LM* should hypothesize a UG-compatible language which is a smallest superset of  $\{s\}$ .

By “smallest superset”, we mean a language that contains *s* and has no proper subset that also contains *s*. Although *SP* as defined above is safe (i.e., will not lead to chronic overgeneration errors), it is problematic since previous facts that were correctly learned may have to be abandoned if the *next* input does not exhibit them. Intuitively, in order for a learner with no memory for past learning events to abide by *SP*, each newly encountered sentence is essentially the first sentence the learner has heard.

In the worst-case, we prove that even a finite (e.g. parameterized) domain is not learnable unless every potential target language in the domain contains a *subset-free-trigger*: a sentence *s* such that the target language includes no proper subset

---

<sup>1</sup> This is joint work with Janet Dean Fodor, City University of New York much of which is drawn from Fodor & Sakas (2005).

<sup>2</sup> It is important to note that Gold-style learnability studies standardly assume that *LM* has access to an enumeration, or ordering, of hypotheses where subsets appear before supersets. Under our memoryless assumption, we disallow such an enumeration in much of the discussion here. Thus, some of the conclusions we draw may seem to contradict well-established formal results when this is not the case.

languages that also contain *s*. Note that a subset-free-trigger is not necessarily an unambiguous trigger, so a single encounter may not correctly pinpoint the target language. But a learner with a fair text and no bias against the target will eventually converge. We give examples of how and to what extent languages and/or parameter settings conspire to create a scenario in which learning fails. It is surprising that SP and memoryless learning are incompatible (except in a domain with unnatural properties, e.g. where every language contains subset-free-triggers and there are no conspiracies).

Finally, we consider how memory in the form of a partially-ordered enumeration of grammars that places subsets before supersets, together with a psychologically plausible means of exploiting it, holds promise for SP compliancy. Preliminary data drawn from simulation experiments run on the CUNY CoLAG word-order domain will be presented.

Baker, C. L. (1979). Syntactic theory and the projection problem. *Linguistic Inquiry* 10, 233-280.S

Berwick, R.C. (1985) *The Acquisition of Syntactic Knowledge*. Cambridge, MA: MIT Press.

Clark, Robin (1992) The selection of syntactic knowledge. *Language Acquisition* 2.2, 83-149.

Dell, François (1981) On the learnability of optional phonological rules. *Linguistic Inquiry* 12.1, 31-37.

Fodor, J.D., and Sakas, W.G. (2005) The Subset Principle in Syntax: Costs of Compliance. *Journal of Linguistics*, 41, 513-569.

Gold, E. M. (1967) Language identification in the limit. *Information and Control*, 10, 447-474.

Manzini, M. Rita and Wexler, Kenneth (1987) Parameters, binding theory, and learnability. *Linguistic Inquiry* 18.3, 413-444.

Pinker, S. (1979). Formal models of language learning. *Cognition*, 7, 217-283.