

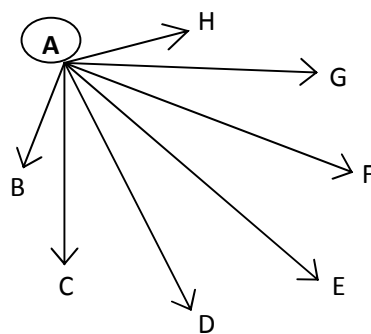
Probabilistic paradigmatics: Principal parts, predictability and (other) possible pieces of the puzzle

Andrea D. Sims

The Ohio State University

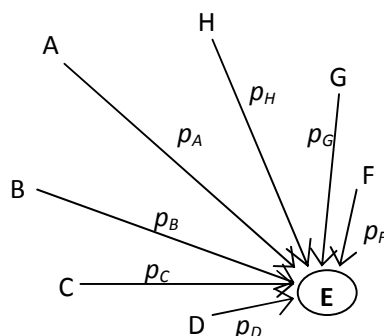
Recent scholarship has refocused attention on how speakers infer inflection class membership from a single inflected form, or a subset of inflected forms of a lexeme. Work has progressed from two directions that are notable here. Finkel and Stump (2007, 2009) approach the issue as a question of predictiveness: what inflected form (or combination of forms) is totally predictive of all other inflected forms of the same lexeme? This is the principal part. In (1) the relationship between a single principal part and the rest of the inflectional paradigm is schematically represented as a series of nodes A-H (representing cells in a morphophonological paradigm) and directed edges (representing predictive relationships).

(1) A schematized network representation of a paradigm representing cell predictiveness



From the other direction, the issue has been approached as a question of predictability: how predictable is the form corresponding to some paradigm cell, given some other inflected form (or combination of forms) belonging to the same lexeme (e.g. Albright and Hayes 2002, Ackerman et al. 2009)?

(2) A schematized network representation of a paradigm representing cell predictability (which occurs with some probability p)



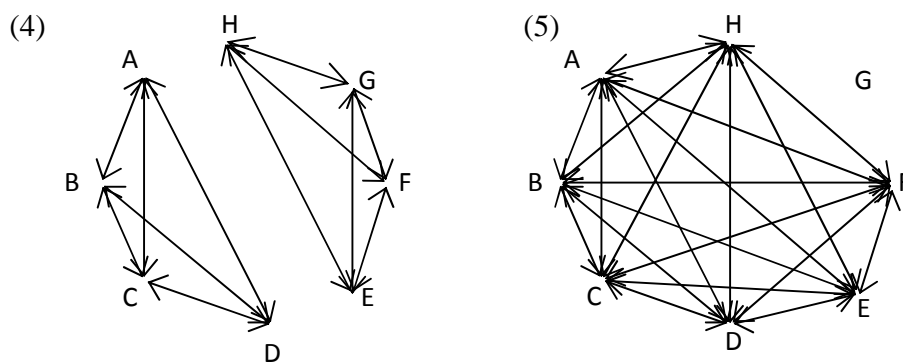
Both perspectives highlight that implicational relations holding among forms are crucial to solving the question of how speakers generalize inflection class membership. However, perhaps because study has focused either on predictability or predictiveness, not both, there seems to often be a tacit assumption that a form must either be predictable from other members of the paradigm, or predictive of other members of the paradigm, or both. In this paper I demonstrate that this is not always true in inflectional systems, and explore consequences for inflectional class structure.

The overarching goal of this paper is to suggest the need for a broader notion – paradigmatic cohesion. I argue that cell predictability and cell predictiveness represent pieces of the inflection class puzzle, but a central question has been overlooked: Is the morphophonological paradigm (necessarily) a cohesive structure? A ‘paradigm’ is typically defined as the set of cells that are licensed by the combinatory possibilities of morphosyntactic properties (e.g. Spencer 2004). The paradigm is thus, virtually by definition, complete. However, a distinction between *morphosyntactic* paradigms and *morphophonological* paradigms (Stump 2002) makes it clear that this definition applies to the morphosyntactic paradigm. Moreover, heteroclisys (a phenomenon in which inflectional exponents for a given lexeme come from two distinct inflection classes) already hints that a morphophonological paradigm need not be complete and cohesive in this same sense.

Looking jointly at cell predictiveness and cell predictability, I present a range of examples in which inflectional forms cannot be represented as a single network of implicational relations. This is equivalent to saying that the relevant morphophonological paradigms cannot be considered cohesive structures. The most notable data here comes from Modern Greek; in some classes, singular formatives, plural formatives, and stress formatives cross-cut each other to an unusual degree, as shown in (3).

| (3) | ‘force’ | ‘mother’ | ‘greengrocer’ | ‘tourist’ | ‘father’ | |
|---------|------------------|-----------------|-------------------|------------------|-----------------|--|
| NOM.SG. | <i>ḍínami</i> | <i>mamá</i> | <i>manávi-s</i> | <i>turísta-s</i> | <i>patéra-s</i> | |
| ACC.SG. | <i>ḍínami</i> | <i>mamá</i> | <i>manávi</i> | <i>turísta</i> | <i>patéra</i> | |
| GEN.SG. | <i>ḍínami-s</i> | <i>mamá-s</i> | <i>manávi</i> | <i>turísta</i> | <i>patéra</i> | |
| NOM.PL. | <i>ḍínami-s</i> | <i>mamá-ḍes</i> | <i>manávi-ḍes</i> | <i>turíst-es</i> | <i>patér-es</i> | |
| ACC.PL. | <i>ḍínami-s</i> | <i>mamá-ḍes</i> | <i>manávi-ḍes</i> | <i>turíst-es</i> | <i>patér-es</i> | (morphological stress difference underlined) |
| GEN.PL. | <i>ḍínam-eon</i> | <i>mamá-ḍon</i> | <i>manávi-ḍon</i> | <i>turíst-ón</i> | <i>patér-on</i> | |

These inflection classes in Modern Greek represent scenarios akin schematically to (4), in which a single morphophonological paradigm constitutes two distinct networks of implicational relations, or (5), an otherwise cohesive network in which one cell is isolated – neither predictive nor predictable. I quantify this lack of paradigm cohesion probabilistically, using Shannon conditional entropy.



Parallel examples are presented from other languages, showing that Greek is not an isolated case.

These patterns are significant in that they suggest that multiple partial morphophonological paradigms are not restricted to ‘marginal’ phenomena like heteroclisys. Ultimately, this paper supports the (now widely accepted) claim that paradigmatic structure is a central organizing principle in inflectional structure, but rejects the frequent implicit assumption of paradigm-based theories that the paradigm necessarily acts as a unitary whole at the morphophonological level.