Introduction

In recent years, a growing number of linguists have arrived at the conviction
that inflectional paradigms play an essential role in the definition of a lan-
guage's grammar and lexicon. Research on the properties and significance of
inflectional paradigms is now being conducted in many subdisciplines of lin-
guistics, including grammatical theory (e.g. Stump 2001, Blevins 2006, Acker-
man, Blevins and Malouf 2009, Round 2013, O'Neill 2014, and the contribu-
tions to Plank 1990), language typology (Carstairs 1987, Baerman et al. 2005,
2007, 2010, Chumakina and Corbett 2013, Stump and Finkel 2013), histori-
cal linguistics (Fuß 2005, Maiden et al. 2011, Cruschina et al. 2013, Gardani
2013, Fertig 2013), psycholinguistics (Baayen and Schreuder 2003, Bittner
et al. 2003, Milin et al. 2009) and computational linguistics (Beesley and
Karttunen 2003, Brown and Hippisley 2012).

My purpose here is to examine the theoretical indispensability of inflec-
tional paradigms, and in particular, their role as a grammatical interface. Thus,
at the most general level, I am concerned with asserting two hypotheses about
inflectional paradigms: (i) the *irreducibility hypothesis*, according to which
some morphological regularities are, irreducibly, regularities in paradigm
structure; and (ii) the *interface hypothesis*, according to which a language's
paradigms are the interface of its inflectional morphology with its syntax and
semantics. At a more concrete level, I will propose and justify a formal theory
of inflectional morphology that is compatible with the irreducibility and inter-
face hypotheses. This theory, the *paradigm-linkage theory*, was first proposed
by Stump 2002, and was subsequently elaborated on by Ackerman and Stump
2004, Ackerman, Stump and Webelhuth 2011, Spencer and Stump 2013, Stew-
art and Stump 2007, Stump 2006, 2010, 2012, 2014a, 2015, Stump: to appear
A, Stump: to appear B. O'Neill 2011, 2013 and Round 2013 have also advanced
proposals that are closely akin to the paradigm-linkage theory, and the analysis

1

2   *Introduction*

of Nepali verb inflection proposed by Bonami and Boyé (2008, 2010) has important similarities as well.[1]

The gist of the paradigm-linkage theory is that the definition of a language's inflectional morphology is based on three interlocking kinds of paradigms. A lexeme L's **content paradigm** identifies the range of morphosyntactic property sets with which L is associated in syntax and which determine L's semantic interpretation in accordance with its syntactic context; a stem X's **form paradigm** identifies the range of property sets for which the various word forms arising from X are inflected; and a lexeme L's **realized paradigm** associates each of L's fully inflected word forms with the content that it expresses. Canonically, a lexeme's content paradigm is isomorphic to the form paradigm of its stem and to its realized paradigm; but this isomorphic relationship is often disrupted by a variety of disparate morphological phenomena, including defectiveness, overabundance, syncretism, suppletion, heteroclisis, homomorphy, deponency, polyfunctionality and morphomically conditioned inflection.

This new conception of inflectional paradigms affords new explanations for synchronic mismatches between a word's content and its morphological form. It is a novel but natural extension of the inferential-realizational theories of inflection proposed by Matthews 1972, Anderson 1992 and Stump 2001. Like them, it entails that morphology is an autonomous domain of linguistic structure; but it is set apart by its premise that the definition of a language's inflectional morphology must account not only for patterns of inflectional exponence, but for the sometimes complex linkage between content and form that these patterns entail.

The book can be seen as comprising two parts. The first part (Chapters 1–6) lays the book's conceptual groundwork; the second (Chapters 7–14) proposes and motivates the paradigm-linkage theory.

Chapter 1 ("What are inflectional paradigms?") introduces the notion of an inflectional paradigm and its relevance to morphology, syntax and semantics. Current theories of grammar are in stark disagreement over the significance of inflectional paradigms; some hold that paradigms are merely an epiphenomenon of principles of morpheme combination and therefore have no role in the definition of a language's inflectional morphology; others maintain that the definition of a language's inflectional morphology makes essential reference to the structure of paradigms. I discuss the numerous shortcomings of a morpheme-based conception of inflectional morphology and present two hypotheses that assume a paradigm-based approach.

1 I discuss these similarities in Chapter 8.

Drawing on the principles of canonical typology (Corbett 2005, 2009, Brown et al. 2013), **Chapter 2** ("Canonical inflectional paradigms") develops the notion of a canonical inflectional paradigm: a typological idealization relative to which the inflectional paradigms of natural languages may be compared. A morpheme-based approach to inflection would suffice if inflectional paradigms were always canonical; but as I demonstrate at length in this book, actual inflectional paradigms deviate from the canonical ideal in a variety of ways, and each such deviation engenders a different kind of problem for morpheme-based inflection. Even so, the notion of a canonical inflectional paradigm provides a crucial point of reference for the discussion of such phenomena.

The inflectional paradigm of a lexeme may be seen as a set of cells, each cell being the pairing of a word's lexical and morphosyntactic content with its morphological form. In order to develop and refine this preliminary conception of inflectional paradigms, it is essential to understand its fundamental components. Chapters 3–6 accordingly present a detailed explanation of four basic notions: morphosyntactic properties, lexemes, stems and inflection classes.

Chapter 3 ("Morphosyntactic properties") is a detailed examination of the nature of morphosyntactic properties (or "features," Corbett 2012). Morphosyntactic properties are part of the shared vocabulary of morphology, syntax and semantics. They serve in syntax to determine a word form's distribution with respect to other constituents and to regulate its relations with other parts of a sentence; in morphology, they determine the inflectional exponents involved in a word form's phonological expression; and at least some morphosyntactic properties are associated with specific semantic content. In order to understand the structure of inflectional paradigms, it is essential to be precise about the characteristics of morphosyntactic properties – their association with syntactic categories, the ways in which they may be associated with word forms, the ways in which they combine, the logical relations that may exist between different sets of morphosyntactic properties, property constraints and their satisfaction, and the nature of morphosyntactic properties' exponence.

Chapter 4 ("Lexemes") presents a detailed discussion of the notion "lexeme," a lexical abstraction allowing distinct word forms to be classified according to their shared lexical content and contrasting morphosyntactic content. Lexemes are central to understanding the organization of a language's lexicon, but cannot simply be identified with lexical entries. A language's lexemes are intuitively regarded as differing from one another in both content and form, but this, too, is an oversimplification.

Chapter 5 ("Stems") is a detailed account of the role of stems in the definition of inflectional paradigms. In the simplest cases, the same stem serves as

4   *Introduction*

the basis for every word form in a lexeme's paradigm. But a lexeme's inflection often depends on more than one stem. Stem alternations within a lexeme's inflectional paradigm are of various kinds. Sandhi alternations are purely an effect of automatic phonology. Among nonautomatic alternations, some are phonologically conditioned and others grammatically conditioned; crosscutting this distinction is a distinction between **class-determined** alternations (alternations that follow from membership in a particular inflection class) and **class-independent** alternations (alternations that are not simply an effect of membership in a particular inflection class); and grammatically conditioned stem alternations may themselves be morphosyntactically conditioned (in which each alternant is invariably associated with a particular morphosyntactic property set) or **morphomic** (Aronoff 1994), following a distributional pattern whose significance is purely morphological, with no invariant phonological, syntactic or semantic correlate.

Languages with rich inflectional systems frequently exhibit contrasting inflection classes, each of which is associated with its own particular inventory of inflectional markings; examples are the Latin declension classes and conjugation classes. In **Chapter 6** ("Inflection classes"), I discuss the properties of such classes, distinguishing between global inflection classes (which determine full paradigms) and segregated inflection classes (which determine specific subparadigms). I address the important question of what inflection classes are classes of; as I show, the standard assumption that they are classes of lexemes is difficult to reconcile with the phenomenon of heteroclisis (Stump 2006), which instead favors the assumption that they are classes of stems. While inflection classes are often distinguished by different inventories of affixes, they are very frequently distinguished by their patterns of stem formation and stem alternation.

If all inflectional paradigms conformed to the canonical ideal described in Chapter 2, there would be no reason to attribute any theoretical significance to them, since each of a lexeme's word forms could be seen as arising through a simple "spelling out" of its associated morphosyntactic properties. But inflectional paradigms rarely conform to the canonical ideal; on the contrary, there are numerous ways in which content and form may be misaligned in a lexeme's inflectional realization; such misalignments invariably involve patterns defined not over individual word forms but over inflectional paradigms. The second part of the book comprises detailed examinations of the different kinds of misalignment observed in the world's languages and the development of a theory of inflectional morphology that is compatible with the full range of observed misalignments.

As a preliminary to this presentation, **Chapter 7** ("A conception of the relation of content to form in inflectional paradigms") distinguishes three ways of conceiving of inflectional paradigms (Stump 2002, 2006, Stewart and Stump 2007).

- A lexeme L's content paradigm enumerates the morphosyntactic property sets with which L may be associated in syntax and semantics. The cells in this paradigm ("content cells") are therefore pairings of the lexeme L with each relevant morphosyntactic property set σ: $\langle L, \sigma \rangle$.
- A stem X's form paradigm specifies the range of property sets that may be realized through the inflection of X. The cells in this paradigm ("form cells") are therefore pairings of the stem X with each relevant property set τ: $\langle X, \tau \rangle$.
- The realized paradigm of a stem X is the smallest set R such that for each pairing $\langle X, \tau \rangle$ in X's form paradigm, $\langle w, \tau \rangle$ belongs to R if and only if w realizes X and τ.

In general, each content cell is realized by being linked to a form cell whose realization it shares; this form cell is the content cell's **form correspondent**. In the canonical case, a lexeme L has a single stem X such that each content cell $\langle L, \sigma \rangle$ in L's content paradigm has $\langle X, \sigma \rangle$ as its form correspondent, so that $\langle L, \sigma \rangle$ and $\langle X, \sigma \rangle$ share a realization $\langle w, \sigma \rangle$. In noncanonical cases, however, the correspondence between content cells and form cells is more complex; each such case involves one or another kind of mismatch between content and form.

Mismatches can be observed at different levels of granularity. Some can be observed within a single paradigm (Chapters 8–11). Others can only be seen by comparing distinct paradigms belonging to the same category (Chapter 12) or to different categories (Chapter 13). In **Chapter 8** ("Morphomic properties"), I discuss several noncanonical systems of inflection in which the grammatical distinctions relevant for a lexeme's syntax are neither identical nor isomorphic to those relevant for its inflectional realization; these include the systems of subject–verb agreement in Hua and Nepali, the system of tenses in Twi, and the system of verb inflection in Noon. I show that in these systems, a lexeme's inflected forms are not directly determined by the morphosyntactic property sets relevant to their syntax, but instead involve the realization of **morphomic** properties (Aronoff 1994) – properties whose sole motivation is morphological. Thus, these systems involve a mismatch between the property sets distinguishing the cells of a lexeme's content paradigm and the property sets distinguishing the cells of its stem's form paradigm.

6   *Introduction*

Chapter 9 ("Too many cells, too few cells") examines cases in which a con-
tent cell fails to correspond to any form cell (and hence lacks a realization) as
well as cases in which a lexeme has more realizations than expected. Instances
of this latter sort include cases of overabundance, in which a lexeme has more
than one realization for the same morphosyntactic property set (as with English
dreamed and *dreamt*); cases of overdifferentiation, in which a lexeme's inflec-
tion expresses more morphosyntactic distinctions than are normal for mem-
bers of its syntactic category (as with English *am* and *are*); and cases of shape
alternation, in which synonymous word forms are restricted to complementary
phonological or syntactic contexts (as with English *my* and *mine*).

Chapter 10 ("Syncretism") examines the very widespread phenomenon of
syncretism: the realization of distinct cells in a paradigm by the same word
form. In general, syncretism involves two or more content cells sharing a single
form correspondent, but such instances arise in more than one way. Natural-
class syncretisms arise only because no rule of inflectional realization is sen-
sitive to the morphosyntactic distinction between the syncretized cells; other
syncretisms are directly stipulated by rules of morphology. These stipulated
syncretisms include directional syncretisms (which arise when the realization
of one property set systematically patterns after that of some distinct property
set) and morphomic syncretisms (which arise when two property sets that do
not form a natural class are nevertheless alike in their realization and neither
set is associated with that realization independently of the other set).

Chapter 11 ("Suppletion and heteroclisis") focuses on the related phenom-
ena of suppletion (the replacement of one stem by a morphophonologically un-
related stem in a lexeme's inflectional paradigm) and heteroclisis (suppletion
of stems belonging to distinct inflection classes). Suppletion and heteroclisis
reflect two dimensions of variation among class-independent stems: they may
differ in form in a way that is not predicted by their inflection-class member-
ship (and may, in that case, still be members of the same inflection class); in-
stead or in addition, they may differ in their inflection-class membership (and
may, in that case, still be alike in form). Though suppletion is seen as a kind of
irregularity, it does exhibit certain cross-linguistic regularities.

Another well-documented phenomenon involving a mismatch of content
and form is that of deponency. In **Chapter 12** ("Deponency and metaconju-
gation"), I discuss deponent paradigms, in which morphology that ordinarily
serves to realize one class of morphosyntactic property sets is instead used to
realize a contrasting class of property sets. In Latin, for example, deponent
verbs possess the morphology usual for passives but exhibit the syntax and
semantics of active verbs. Deponency involving morphosyntactic properties

other than properties of voice are observable in a number of languages. I distinguish deponency from metaconjugation, the realization of content-level morphosyntactic contrasts as form-level distinctions in inflection-class membership.

In **Chapter 13** ("Polyfunctionality"), I discuss a final phenomenon in which a difference of content between two paradigm cells coincides with a similarity in form. This is the phenomenon of polyfunctionality, the systematic use of the same morphology for different purposes. Instances of polyfunctionality vary widely in their characteristics. In some cases, the same morphology has more than one use in the inflection of the same class of lexemes. In other instances, the same morphology expresses one kind of content in the inflection of one category of lexemes and a distinct kind of content in the inflection of a distinct category of lexemes. Examples of these sorts show that languages often put the same morphology to more than one use in expressing the inventory of grammatical contrasts relevant to syntax; that is, content cells that are different may nevertheless have form correspondents and realized cells that are alike.

Chapter 14 ("A theoretical synopsis and two further issues") presents a summary of the formalization proposed for the paradigm-linkage theory over the course of Chapters 7–13, then addresses two issues pertinent to its formalization and application. The first of these concerns the possibility of paring down the paradigm-linkage theory by adopting a purely abstractive approach to inflectional exponence (Blevins 2006); adopting such an approach would seem to open the possibility of eliminating form paradigms from the paradigm-linkage architecture, but as I show, there are good reasons not to pursue this strategy. The second issue relates to the relevance of paradigm linkage to morphological change; as I show, the proposed theoretical architecture sheds important light on the sometimes conflicting pressures that affect the evolution of inflectional systems.

The evidence presented here shows that inflectional paradigms constitute a theoretically indispensable grammatical interface; this evidence motivates the development of a theory of inflectional morphology with the essential characteristics of the paradigm-linkage theory.

1   *What are inflectional paradigms?*

In this chapter, I examine the defining characteristics of inflectional paradigms (Section 1.1). A central issue in morphological theory is whether inflectional paradigms have theoretical significance. According to paradigm-based theories of inflection (e.g. those of Stump 2001, Blevins 2006, Ackerman, Blevins and Malouf 2009, Brown and Hippisley 2012), the definition of a language's inflectional morphology makes essential reference to the structure of paradigms; but according to morpheme-based theories (e.g. those of Halle and Marantz 1993, Bobaljik 2002, Müller 2002), paradigms are instead merely an epiphenomenal effect of principles of morpheme combination, having no essential role in the definition of a language's inflectional morphology. I compare the morpheme-based perspective (Section 1.2) with the paradigm-based approach (Section 1.3); of these, only the latter is compatible with two central hypotheses for which I argue in this book: the irreducibility hypothesis and the interface hypothesis.

1.1   *What is an inflectional paradigm?*

In a language with inflectional morphology, **morphosyntactic properties** are grammatical properties to which the language's syntax and morphology are both sensitive. In French, for example, the gender properties "feminine" and "masculine" and the number properties "singular" and "plural" are morphosyntactic properties. On one hand, syntactic agreement relations are sensitive to contrasts in gender and number; thus, the noun phrase in (1) is grammatical because its constituents agree in gender and number, and the noun phrase in (2) is ungrammatical because its constituents fail to agree in this way. At the same time, rules of inflectional morphology are likewise sensitive to contrasts in gender and number; thus, the French adjective NATIONAL has distinct

8

feminine and masculine forms in the plural (feminine plural *nationales*
/nasjɔnal/, masculine plural *nationaux* /nasjɔno/), while in the singular, the
feminine and masculine forms of NATIONAL are distinguished orthographically
though not phonologically (feminine singular *nationale*, masculine singular
national, both /nasjɔnal/).

(1)   *les   musées   nationaux*
        the.PL   museum.MASC.PL   national.MASC.PL
        'the national museums'

(2)   **le   musées   nationales*
        the.MASC.SG   museum.MASC.PL   national.FEM.PL

Some morphosyntactic properties have specific semantic correlates; thus, noun
phrases with singular reference tend to be headed by nouns in their singular
form. But morphosyntactic properties may also lack any obvious semantic cor-
relate; thus, while noun phrases with female reference may tend to be headed by
feminine nouns, some feminine nouns (e.g. *recrue* 'recruit,' *sentinelle* 'sentinel')
ordinarily have male reference and others – in fact, the large majority – fail to
refer specifically to either sex (*souris* 'mouse,' *table* 'table,' *invention* 'invention,'
and so on). Even so, a French noun's gender may be seen as part of its content, if
content is assumed to encompass properties whose significance is either seman-
tic or primarily grammatical.

     The **inflectional categories** ("features")[1] of gender and number crosscut each
other in the inflection of French adjectives: "feminine" and "masculine" com-
bine with "singular" and "plural" to define a matrix of up to four (orthographic)
word forms for each adjective, as in (3). Each gender/number combination is a
combination of morphosyntactic properties arising in syntax, and each adjecti-
val word form expresses one or more such combinations.

(3)   Feminine   Masculine
        Singular   *nationale*   *national*
        Plural   *nationales* *nationaux*

In order to understand the structure of such matrices, it pays to be precise about
a number of conceptual distinctions. First, the semantico-grammatical content
shared by the word forms in such a matrix constitutes the **lexeme** realized
by those word forms; thus, the content shared by the word forms in (3) is the

---

1 I favor the term "inflectional category" over "feature"; the latter term tends to be used impre-
cisely, sometimes referring to true inflectional categories (e.g. tense) and sometimes to specific
morphosyntactic properties (e.g. past tense).

10   *What are inflectional paradigms?*

French adjectival lexeme NATIONAL. Because a lexeme is composed of elements of semantic and grammatical content, it is not a linguistic form; rather, it is expressed by linguistic forms. One cannot equate a lexeme with the stem common to all the linguistic forms that express it, since for some lexemes, there is no such stem; in French, for example, the lexeme ALLER 'go' is realized by *vais* /vɛ/ 'I go,' *allons* /alɔ̃/ 'we go,' and *iras* /iʁa/ 'you (sg) will go' (among other word forms), which share no part of their form. I follow the convention of representing lexemes in SMALL CAPITAL letters; but such representations are to be understood as referring to an abstract combination of semantic and grammatical properties.[2]

Each of a lexeme's word forms expresses one or more of its cells, where a lexeme L's **cells** are form–content pairings. The content of each such pairing includes L and a complete and coherent morphosyntactic property set σ compatible with L; the pairing's form is the inflected word form w that realizes both L and σ. Thus, a cell pairing the lexical content of L and the morphosyntactic content of σ with the word form w may be represented as $\langle L, \sigma : w \rangle$. A lexeme's complete set of such cells is its **paradigm**. For example, the paradigm of French NATIONAL is the set of cells in (4).

(4)   The paradigm of French NATIONAL
         { ⟨NATIONAL, {fem sg}: *nationale*⟩,   ⟨NATIONAL, {masc sg} : *national*⟩,
            ⟨NATIONAL, {fem pl} : *nationales*⟩,   ⟨NATIONAL, {masc pl} : *nationaux*⟩ }

In a cell $\langle L, \sigma : w \rangle$, the word form w **realizes** (is the **realization** of)

- the lexeme L,
- the property set σ,
- the combined content of L and σ, and
- the cell itself.

In cases of gross irregularity, the word form w in a cell $\langle L, \sigma : w \rangle$ is stipulated lexically; but more often, w is deducible from L and σ. In particular, it is a language's inflectional morphology that ordinarily determines the form w from the lexeme L and the property set σ in a cell $\langle L, \sigma : w \rangle$.

Both within and across languages, lexemes' paradigms vary widely in shape and size. If a lexeme's realization is sensitive to a large number of crosscutting morphosyntactic properties, its paradigm comprises a correspondingly large number of word forms. This correspondence is complicated by the fact that

2 In Chapter 4, I examine the defining properties of lexemes in systematic detail.