Reductiō ad discrīmen: Where features come from

Elizabeth Cowper and Daniel Currie Hall
University of Toronto; Saint Mary's University

Abstract

This paper addresses two fundamental questions about the nature of formal features in phonology and morphosyntax: what is their expressive power, and where do they come from? To answer these questions, we begin with the most restrictive possible hypothesis (all features are privative, and are wholly dictated by Universal Grammar, with no room for cross-linguistic variation), and examine the extent to which empirical evidence from a variety of languages compels a retreat from this position. We argue that there is little to be gained by positing a universal set of specific features, and propose instead that the crucial contribution of UG is the language learner's ability to construct features by identifying correlations between contrasts at different levels of linguistic structure. This view resonates with current research on how the interaction between UG and external 'third factors' shapes the structure of language, while at the same time harking back to the Saussurean notion that contrast is the central function of linguistic representations.

1. Introduction

The use of formal features in generative grammar derives from the work of Trubetzkoy (1939) and others in the Prague School.* In early generative phonology, Jakobson et al. ([1951] 1969) used distinctive features as a cross-linguistically uniform way of identifying and representing minimal contrasts in sound systems. For them, features encode polar oppositions, are relatively few in number, and can be universal, provided that they are sufficiently abstract. This view of features was continued and expanded on in Halle's (1959) *The Sound Pattern of Russian* (SPR) and Chomsky and Halle's (1968) *The Sound Pattern of English* (SPE). SPR extended the role of features from a purely classificatory one by using them as the basis of a system for formulating phonological rules; SPE added to the list of features and gave them articulatory rather than acoustic names. In the SPR/SPE system, features are binary, and have intrinsic phonetic content. They also form a universal and exhaustive mechanism for representing phonological contrasts; Halle (1959: 19) posits that "differences between segments can be expressed only as differences in their feature composition and that consequently segments (even in different languages) can differ from each other only in a restricted number of ways."

Although many changes to the SPE feature system have been proposed in the succeeding decades, it remains highly influential, and some recent work, particularly in Optimality Theory, which has tended to de-emphasize the importance of representations, has returned to essentially this view of features. One of the primary departures from the SPE system has been the proposal that some or all features are monovalent rather than binary.

While Halle (1959) takes the position that phonological features are universally binary, this is not the only logical possibility. Monovalent features gained currency with the rise of autosegmental phonology: as features came to be seen as entities in their own right, rather than merely as properties of segments, it became natural to speak of contrasts between the presence and the absence of a feature, and autosegmental spreading promised to reduce or eliminate the need for alpha notation (Goldsmith 1981; 1985). Some theories of phonology, such as Element Theory (Ewen and van der Hulst 1985, Harris and Lindsey 1995, Backley 2011) and the Parallel Structures Model (Morén 2003; 2006, Iosad 2012), adopt the strong view that all features are privative; others (e.g., Goldsmith 1985, Sagey 1990, Clements and Hume 1995, Halle

*We are grateful to Sarah Clarke, Elan Dresher, Pavel Iosad, Betsy Ritter, and two anonymous reviewers for helpful comments, and to the participants at the 2013 Dog Days of Summer Workshop in Toronto, and the CASTL Conference on Features in Phonology, Morphology, Syntax, and Semantics for excellent questions and stimulating discussion. All errors are ours.

et al. 2000) employ a combination of privative and binary features, or make a distinction between privative organizing nodes and binary terminal features.

Chomsky (1965) uses binary feature notation in the syntax, primarily to express the syntagmatic restrictions of subcategorization and selection; only in the final chapter does he discuss paradigmatic features such as gender, number, and case. In the spirit of Halle (1959) and Chomsky and Halle (1968), he takes features to be binary, with the possibility of underspecification. Underspecified values are defined as non-distinct from specified values, giving a limited degree of ternarity in the feature system.

More recently, building on prior work in phonology, privative features in dependency relations have been used in morphosyntax as well, as in the work of Harley (1994), Harley and Ritter (2002a;b), Béjar (2003; 2008), Cowper (2005), and McGinnis (2005; 2008), though Harbour (2013) argues for equipollent features in morphosyntax.

There are thus several degrees of freedom in the construction of a theory of grammatical features,[1] giving a relatively large range of possibilities. In what follows, we explore just how restrictive the theory can be while still accounting for the range of linguistic data that features have been brought to bear on.

2. The retreat

It is useful to start from the most restrictive possible hypothesis, and then see to what extent this hypothesis is contradicted by the data. At each step, when revising the hypothesis we will retreat only just enough to permit an account of the data under consideration.

2.1. A starting point

In the case of features, the maximally restrictive hypothesis would be one that minimizes both their formal expressive power and the potential for cross-linguistic variation. Since privative features, *ceteris paribus,* have more limited expressive power than do binary features, we begin with them. We will call the maximally restrictive position Hypothesis 1:

> **Hypothesis 1:** *There is a universal set of privative features, with consistent substantive content, all of which are active in all languages.*

If this hypothesis is correct, then there should be no cross-linguistic variation in grammatical features or in their interpretation. All languages make the same distinctions, at all levels of the grammar, and mark those distinctions in the same way. In the phonology, all languages should use the same phonological features, and they should have cross-linguistically consistent phonetic implementation. In the syntax, all languages should use the same formal features, and they should have cross-linguistically consistent semantic import.

It is, of course, impossible in principle to demonstrate that a feature is wholly absent from a given language, rather than present but grammatically inert. The word *active* is thus essential to the falsifiability of the strong universalist claim in Hypothesis 1. We can say that a feature is active if the grammar crucially refers to it in any way. In the case of potential morphosyntactic features, then, it is not enough to show that the semantic content of a feature is present in the encyclopedic meaning of some lexical item; rather, to be considered active it must be involved in inflectional paradigms, or trigger syntactic movement or agreement, or play some other demonstrably formal role. On the phonological side, an active feature is one that spreads, or blocks spreading, or is otherwise referred to by phonological rules or constraints, or even merely serves to distinguish one phoneme from another. It is not sufficient for the phonetic substance of the feature to be

[1] Throughout, we use the term *grammatical* in its broad sense, to include all components of the grammar; i.e., at least phonology, morphology, and syntax, as well as the phonology–phonetics interface and the syntax–semantics interface.

present on some segments; to be deemed active, the feature must be phonologically consequential.

As expected, this hypothesis is demonstrably too strong, both for morphosyntactic features and for phonological ones. In morphosyntactic number systems, for example, languages exhibit anywhere from no grammatical number contrasts, as in Kawi (Becker and Oka 1974), to four or five grammatically distinct numbers, as in Lihir (Corbett 2000). The Kawi pronoun paradigm, as shown in (1), distinguishes the person of the referent and the degree of intimacy between the speaker and the hearer ('close' or 'distant'), but does not mark number at all:

(1) Kawi (Javanese) pronoun paradigm (Becker and Oka 1974: 232)

	CLOSE	DISTANT
SPEAKER	*aku*	*kami*
HEARER	*ka(N)u*	*kita*
OTHER	*ia*	*sira*

Becker and Oka (1974: 232) note that Kawi also does not mark grammatical number on nouns, although semantic plurality can be expressed through quantifiers and coordination. While the concept of multiplicity is within the expressive power of the language, there is no evidence that categories such as singular or plural are formally encoded in the grammar.

Old Church Slavonic, in contrast, has a robust three-way grammatical number contrast, as illustrated in the inflectional paradigm in (2):

(2) Old Church Slavonic (South Slavic) neuter noun and adjective suffixes (Béjar and Hall 2000: 2)

	SINGULAR	PLURAL	DUAL
NOM./ACC.	-o	-a	-æ
GENITIVE	-a	-ɵ	-u
LOCATIVE	-æ	-æxɵ	-u
DATIVE	-u	-omɵ	-oma
INSTRUMENTAL	-emɔ	-ɨ	-oma

In the OCS system, number is clearly grammatically active; it is obligatorily marked on nouns, pronouns, and adjectives, and reflected in patterns of agreement and concord. Where Kawi offers no evidence of number features at all, the grammar of OCS must be capable of formally representing the categories of singular, dual, and plural. Other languages, such as Lihir (Oceanic; Corbett 2000: 25), include additional number categories such as trial and paucal in their pronoun systems, but offer less conclusive evidence for their grammatical encoding, as their involvement in agreement is less systematic (see Cable (2010) for discussion). In any case, examples like OCS show that a language may make grammatically relevant distinctions among up to at least three different number categories, and examples like Kawi show that a language need not featurally encode number at all.

The notion that morphosyntactic features have universal substantive content is seriously challenged by work on Infl by Ritter and Wiltschko (2009), who argue that the features of Infl can be temporal (as in English), locational (as in Halkomelem), or personal (as in Blackfoot). In all three systems, the features serve to anchor clauses deictically and play essentially the same role in the syntax, but each of the three languages uses a different substantive dimension of deixis to ground the system.

In Halkomelem, as illustrated in (3), an auxiliary obligatorily identifies each clause as proximal or distal, anchoring it to the deictic centre spatially, but not temporally:

(3) Halkomelem (Ritter and Wiltschko 2009: 155)

a. í qw'eyílex tú-tl'ò
AUX.PROXIMAL *dance* *he*
'He is/was dancing (here).'

b. lí qw'eyílex tú-tl'ò
AUX.DISTAL *dance he*
'He is/was dancing (there).'

In Blackfoot, Ritter and Wiltschko (2009) argue, the anchoring is personal rather than spatial or temporal. In support of the proposition that Blackfoot person prefixes on verbs represent Infl itself, and not merely agreement, they note, citing Frantz (1971), that the prefixes *nit-* and *kit-* in sentences such as those in (4) indicate only the presence of discourse participants in the event, and not their specific grammatical or thematic roles. Thus the second-person *kit-* appears in (4b) and (4c) alike (in contrast with (4a), which has no second-person argument at all); whether the addressee is the subject or the object is indicated by the suffixes glossed as 'direct' (*-a*) or 'inverse' (*-ok*).

(4) Blackfoot (Ritter and Wiltschko 2009: 182)

a. nit-iik-wákomimm-a-wa n-itan-wa
1-very-love-DIR-*3*SG *1-daughter-3*SG
'I love my daughter.'

b. kit-iik-wákomimm-a-wa n-itan-wa
2-very-love-DIR-*3*SG *1-daughter-3*SG
'You love my daughter.'

c. kit-iik-wákomimm-ok-wa n-itan-wa
2-very-love-INV-*3*SG *1-daughter-3*SG
'My daughter loves you.'

In the syntax of nominals, systems of determiners and classifiers differ significantly among languages, as can be seen in various papers in Ghomeshi et al. (2009). Some languages have rich systems of classifiers, while others do not. In some languages, demonstratives are part of a grammatical paradigm that also includes determiners, and thus seem to spell out grammatical features, while in others, they function more like modifiers and may bear no grammatically relevant features at all.

Morphosyntactic features in a single language can also change over time. As argued by Cowper and Hall (2013), the feature Modality was absent from the English Infl system until about 1500, when the syntactic properties of the class of modal verbs were established. Before the change, there were verbs whose lexical semantics included root or epistemic modal meanings (much as do adjectives like *possible* or nouns like *obligation*), but no obvious role for any featural morphosyntactic encoding of this property; after the change, the modal verbs are distinguishable not only by their meanings, but also by their inflectional properties and syntactic position.

Turning to phonological features, an obvious problem for Hypothesis 1 is the existence of both signed and spoken languages, as noted by Mielke (2008: §1.5.1), among others. If features have consistent substantive content, then signed and spoken languages require very different phonological features (see, e.g., Sandler (1989) and Brentari (2011) for signed language feature geometries). Alternatively, if signed and spoken languages make use of the same phonological features, then the content of those features must either be able to differ significantly from language to language, or be so abstract as to make a claim of consistent substantive content essentially meaningless. While there have been many productive efforts to unify the abstract formal structures of signed and spoken phonology (see, e.g., Brentari (1998) on syllable structure and sonority, or Morén (2003) on the geometric organization of place and manner features), it seems clear that there are many substantive phonetic properties that are phonologically relevant in signed languages but not in spoken ones, and vice versa. To take an example from Morén (2003), signed and spoken languages may both have C-Place and V-Place organizing nodes, but only signed languages are likely to have PINKIE as a dependent of those nodes, and only spoken ones are likely to use DORSAL. While manner features such as OPEN or CLOSED can readily apply to either modality, features that identify articulators must differ.

It is not only different modalities that present problems for Hypothesis 1 at the phonological level. Even within the same modality, phonemic inventories differ significantly from language to language. If all languages actively employ the same set of phonological features, then it is mysterious that the phonological behaviour of segments seems to depend not solely on their phonetic properties, but also on what they contrast with. (See Calabrese 1995, Dresher 2009, Hall 2007, Mackenzie 2009, Nevins 2010, among others, for various recent treatments of the role of contrast in determining phonological feature specifications and patterning.)

Consider for example the consonant inventory of Hawaiian, given in (5):

(5) Phonemic consonant inventory of Hawaiian (Polynesian)

p		k	ʔ
			h
m	n		
w	l		

Assuming that /k/ is characterized as DORSAL in languages with robust classes of dorsal and coronal obstruents, Hypothesis 1 would dictate that Hawaiian /k/ also be specified as DORSAL. However, Hawaiian /k/ patterns quite differently from /k/ in, say, English. Phonetically, it can be realized as [t]:

(6) Interchangeability of [t] and [k] in Hawaiian (Schütz 1995: 77–78)

a. [kanaka] ∼ [tanata] ‘people’
b. [ko] ∼ [to] ‘sugar cane’
c. [tabetee] ∼ [kabekee] ‘cabbage’

In borrowings from English and other languages, Hawaiian /k/ is substituted for all non-labial, non-glottal obstruents, as in the examples in (7):

(7) Borrowings into Hawaiian from English (Herd 2005: 80)

a. /mekala/ < *medal*
b. /keaka/ < *theatre*
c. /kaakini/ < *dozen*
d. /kanuwika/ < *sandwich*

As Herd (2005) points out, the position of /k/ in the Hawaiian inventory is equivalent to that of /t/ in the inventory of Tahitian, shown in (8).

(8) Phonemic inventory of Tahitian (Polynesian)

p	t	ʔ
f		h
v		
m	n	
	r	

And /t/ plays the same role in borrowings into Tahitian that /k/ plays in borrowings into Hawaiian:

(9) Borrowings into Tahitian from English (Herd 2005: 80)

a. /roti/ < *rose*
b. /tapati/ < *sabbath*
c. /tamanu/ < *salmon*
d. /tiunu/ < *June*

In characterizing the phonological systems of these two languages, there is nothing to be gained by adopting a view of features that compels Hawaiian /k/ and Tahitian /t/ to be represented differently, and there is at least arguably something to be lost. In the absence of a phonological contrast between coronal and dorsal consonants, requiring the specification of the features CORONAL and DORSAL (or their equivalents) does nothing but obscure the fact that Tahitian /t/ and Hawaiian /k/ each serve as the sole obstruent member of a natural class that might more perspicuously be identified as lingual consonants.

A similar problem is found with vowel inventories. For example, Standard Arabic distinguishes three phonemic vowel qualities normally transcribed as /i a u/. Phonologically, there is a two-way height contrast in this vowel system, which might be characterized as high /i u/ versus non-high /a/, or as low /a/ versus non-low /i u/. Phonetically, these vowels exhibit considerable variation, as illustrated in Figure 1, and each of the two height categories encroaches on phonetic space that would be identified as 'mid' in languages that distinguish three heights. If we assume that there are universally specified privative features HIGH and LOW, with cross-linguistically invariant phonetic substance, then we cannot say that Arabic /i/ and /u/ consistently meet the definition of HIGH, or that Arabic /a/ consistently meets the definition of LOW. As far as the phonology and phonetics of Arabic are concerned, all that is relevant about the heights of these vowels is that /i u/ are higher than /a/; attempting to impose on this system a set of features designed to accommodate languages with more fine-grained height distinctions only muddies the picture.

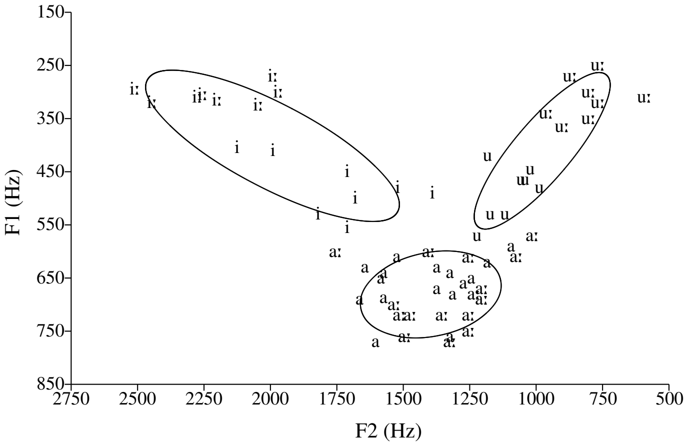


Figure 1: Variation in Standard Arabic vowels (Hall 2011b: 22; data from Abou Haidar 1994)

2.2. *Retreat from universal activity*

Since Hypothesis 1 is too strong, we must weaken it. While more than one of the claims of Hypothesis 1 has been challenged, we will begin by changing just one of them, namely the claim that all features are used in all languages.

> **Hypothesis 2:** *There is a universal set of privative features, with consistent substantive content, of which each language uses a subset.*

This position is similar to the one taken by Chomsky (2000), though he does not, in that paper, address the question of whether features are privative or binary. Under this view, languages can differ in how many grammatical distinctions they make, and in which distinctions they choose. In Arabic, for example, there will be a smaller number of features characterizing vowels than there are in Swedish. In English, there will be fewer features characterizing number than than there are in Old Church Slavonic or Lihir. However, crosslinguistic diversity is constrained by the claim that features have consistent substantive content. Thus, if two languages use a given feature, that feature must have the same phonetic or semantic import. Since features are privative, if languages make a given distinction, the same member of the opposition should be marked in both languages.

There is some evidence, both for morphosyntactic features and for phonological features, that Hypothesis 2 is still too strong. Cowper (2005) and Kyriakaki (2006) have argued that in English and Greek, events are marked relative to states. They thus propose that eventive clauses in those two languages are characterized by the presence of a privative feature EVENT. However, Clarke (2013) has shown that in Japanese, stative clauses must be treated as marked relative to eventive clauses, and proposes that they are characterized by the privative feature STATE. In phonology, Casali (2003) has argued that in ATR harmony systems, either advanced (ATR) or non-advanced (RTR, or, in Casali's binary feature system, [−ATR]) may be the active feature. In Maasai, ATR is active, while in Komo, RTR is active, as illustrated in (10) and (11).

In the Maasai system, an ATR vowel in a suffix such as /-íé/ triggers harmony, causing preceding vowels to become ATR, as in (10a). The data in (10b) show that the harmony is based on ATR dominance, rather than directionality: a root with an ATR vowel, such as /duŋ/ or /bol/, will cause both prefix and suffix vowels to harmonize.

(10) Maasai (Eastern Nilotic): ATR dominant (Casali 2008: 514–515)
- a. ATR suffix makes root (and prefix) vowels ATR
  - (i) [ɛ́-sʊ́j] 'he will follow'
  - (ii) [e-suj-íé] 'he will follow using it'
  - (iii) [ɛ́-bɛ́l] 'he will break (it)'
  - (iv) [e-bel-íé] 'he will break (it) using it'
- b. ATR root makes suffix (and prefix) vowels ATR
  - (i) [ɛ-sʊj-ʊ́] 'he will follow toward'
  - (ii) [ɛ-bɛl-ʊ́] 'it will break off toward'
  - (iii) [e-duŋ-ú] 'he will cut it toward'
  - (iv) [e-bol-ú] 'it will be opened toward'

In Komo, RTR (or [−ATR]) is the active feature in harmony. In (11a), the prefix [mo-]∼[mɔ-] harmonizes with mid vowels in the root. In (11b), we can see that the default form of the prefix is /mo-/; this is the form that occurs when the following vowel is high or low. (Only mid vowels participate in the harmony pattern.) If the prefix vowel is underlyingly advanced, then RTR or [−ATR] must be capable of spreading.

(11) Komo (Bantu): RTR (or [−ATR]) dominant (Casali 2003: 314)
- a. Prefix [mo-]∼[mɔ-] harmonizes for ATR with mid root vowels
  - (i) [mo-séphé] 'a fish'
  - (ii) [mo-cócɑ́] 'rapids'
  - (iii) [mɔ-mɛ́ɲɔ] 'hatred'
  - (iv) [mɔ-gɔgɔ] 'sheep'
- b. Default ATR form [mo-] before high or low root vowels
  - (i) [mo-ɗíŋgɑ] 'tap root'
  - (ii) [mo-kumɑ́] 'healing ritual'
  - (iii) [mo-gɑ́ndɑ] 'a youth'

It is possible to accommodate these analyses under Hypothesis 2, by including both EVENT and STATE, and both ATR and RTR, in the set of universally available features. However, this approach has two problems. First, it misses the point that in each case, the same featural dimension is involved. Second, it predicts that there should be languages that make use of both EVENT and STATE, or of both ATR and RTR as marked features. As pointed out by Harbour and Elsholtz (2012), such an approach would effectively disguise a binary feature as a pair of privative features, reducing the empirical interest of feature privativity.

Setting aside the move just described, then, and taking seriously the idea that EVENT/STATE and ATR/RTR each encode the two poles of a single featural dimension, it seems that what is universal may not be the specific features themselves, but rather the featural dimensions whose poles they represent. The choice of which pole of a given dimension is active, or marked, is language-particular. If true, this suggests that in general, features may be more properly thought of as binary, at least at some level.

2.3. Retreat from universal markedness relations

Moving from universal features to universal dimensions brings us to Hypothesis 3.

> **Hypothesis 3:** *There is a universal set of binary oppositions, with consistent substantive content. Each language uses a subset of these, and selects one pole of each opposition to encode in a privative feature; the opposing value is unmarked.*

Under this view, what is universal—and binary—about a given feature is thus the dimension of contrast. But in any given language, features function in a privative manner, in that only one end of the dimension can be referred to. Maasai has privative [ATR], while Komo has privative [RTR]; English and Greek have privative [Event] and Japanese has privative [State]. Crucially, each language marks only one pole of any given opposition.

The highly restrictive definition set forth in Hypothesis 1 has now been reduced to two components: a universal set of featural dimensions with substantive content, from which languages choose, and the privative behaviour of features at the level of the individual grammar. We now consider whether the second of these can be maintained, postponing discussion of the first to section 5.

Both in syntax and in phonology, it has been argued that features must, at least in some cases, be binary rather than privative. Harbour (2011b) shows that an elegant account of the very complex system of number and noun class in Kiowa is possible, with just three binary features: [±singular], [±augmented], and [±group]. Each of these may crucially be positive, negative, or unspecified on a given syntactic head, and the grammar may refer to either specified value. As Harbour (2011b; 2013) argues, it is not possible to choose a single pole of each opposition as the consistently marked value.

On the phonological side, Krämer (2000) argues that a binary feature [±voice] is the best way to account for the three-way contrast in Breton among (1) voiced obstruents that trigger regressive assimilatory voicing (specified as [+voice]), (2) voiced obstruents that undergo progressive assimilatory devoicing (unspecified for [voice]), and (3) voiceless obstruents (specified as [−voice]). Wetzels and Mascaró (2001) likewise argue for binary [±voice] in accounting for patterns in languages where both voiced and voiceless consonants appear to trigger assimilation. Mackenzie (2009; 2013) makes crucial use of binary features in formulating constraints that express co-occurrence restrictions and drive consonant harmony in Bumo Izon, Kalabari Ijo, Dholuo, Anywa, Päri, Chaha, and Aymara. She argues that "[t]he distinction between segments which are contrastively specified for the negative value of a feature and segments for which the feature in question is noncontrastive and hence simply absent must be referred to by the grammar" (Mackenzie 2009: 78).

However, arguments of this sort are rarely completely decisive. The classes defined by a single binary feature [±F] could, in general, be captured by a pair of privative features F and G, with F representing [+F] and G representing [−F]. Harley and Ritter (2002a) use the marked privative features [Participant], [Speaker], and [Addressee], along with a default interpretation of [Participant] as [Speaker], to account for pronoun systems with an inclusive-exclusive distinction. (12) shows the person feature specifications for such a language on the left. Number features dependent on [Individuation] are omitted; the root node RE stands for 'referring expression.'

(12) a. First person:

RE — [Participant] [Individuation] — [+Speaker, −Addressee]

b. Second person:

RE — [Participant] [Individuation]; [Participant] — [Addressee] [−Speaker, +Addressee]

c. Inclusive:

RE — [Participant] [Individuation]; [Participant] — [Speaker] [Addressee] [+Speaker, +Addressee]

d. Third person:

RE — [Individuation] [−Speaker, −Addressee]

The empirical advantages of this account over an otherwise identical account using the two binary features [±Speaker, ±Addressee], as shown on the right, are almost impossible to discern, and indeed Ritter (p.c.) has more recently adopted the latter view.

A privative account of Kiowa number would also certainly also be possible, given enough features and complex enough realization rules; however, it is safe to say that it would likely be considerably less parsimonious than the one proposed by Harbour (2011a). In a sense, the positive and negative values of Harbour's features already exhibit the same kind of autonomy that privative features do: they are defined as functions, and opposite values can co-occur in the same representation, with the output of one being taken as the input to the other. For example, Harbour (2013) represents trial as [+minimal]([−minimal]([−atomic](n))): the minimal non-minimal non-atomic number, or smallest number greater than dual. One could give distinct names without plus and minus signs to the functions [+minimal] and [−minimal], but that would do nothing more than paper over the fact that their denotations are opposites of each other.

In phonology, Hall (2009) and Iosad (2012) have proposed accounts of the Breton voicing facts that employ monovalent features, based on the assumption that voicing features are dominated by a LARYNGEAL organizing node (Clements and Hume 1995: §3.3.1), and that this node itself can be contrastively absent underlyingly or deleted by the phonological computation. Similarly, Iverson and Salmons (2003) argue for privative VOICE in Dutch, contra Wetzels and Mascaró (2001).

Thus, the question of whether features can or should be treated as universally privative is not, in the end, a purely empirical one. The answer depends entirely on how theoretically attractive privative features are relative to binary ones; in other words, on how theoretically costly binary features are compared to other formal devices. In principle, one could resort to any number of formal tricks to avoid using binary values.

Let us therefore retain Harbour's (2011a) assumption that privative features are more restrictive than binary features and should thus be preferred, all else being equal. However, binary features are also possible; let us assume that Harbour is correct in using them for Kiowa, and that Mackenzie (2009) is right about consonant harmony. This brings us to hypothesis 4:

Hypothesis 4: *There is a universal set of features, all of which have two possible values. Each language chooses a) whether or not to use the feature at all, and b) whether one or both poles of the opposition are active in the grammar.*

For a given featural dimension, then, there are three possible implementations in a given grammar. Using the Event–State dimension as an example, these are shown in (13).

(13) a. Only the Event pole is active: the grammar uses the privative feature [Event].
b. Only the State pole is active: the grammar uses the privative feature [State].
c. Both poles are active: the grammar uses the binary feature [±Event] (or, alternatively, [±State]).

3. Excursus on privativity and notation

Phonologists'—and, more generally, linguists'—understanding of privative and binary features owes a great deal to Trubetzkoy's (1939) discussion of privative, equipollent, and gradual oppositions, which continued in the spirit pioneered by Saussure (1916; 1959). For Saussure (1959: 120), "in language there are only differences [...]. The idea or phonic substance that a sign contains is of less importance than the other signs that surround it." Like Saussure, Trubetzkoy wrote about oppositions between segments, rather than about features *per se,* and because he focused on oppositions between pairs of segments, all the relations he described are, in a basic sense, binary. In a privative opposition, one of the two members is marked and the other unmarked; Trubetzkoy (1939: 67) gives as examples oppositions between voiced and voiceless, nasalized and non-nasalized, and rounded and unrounded. In a gradual opposition, the distinction is one of degree, as in vowel height or pitch, rather than of presence or absence. Finally, in an equipollent opposition, neither member is more marked than the other, nor do the two members occupy different positions on a single continuum; their properties are either equally marked or perhaps incommensurable. Here, Trubetzkoy (1939: 67) cites German /p/–/t/, where there is a minimal opposition between labial and coronal places of articulation, two different values of the same phonetic dimension that do not obviously present themselves either as differing degrees along a single scale or as the presence and absence of any single property. He also refers to an equipollent opposition between German /f/ and /k/, which are presumably incommensurable because they differ on multiple dimensions.

For Trubetzkoy, an opposition that phonetically ("logically" or "potentially," in his terms) appears to be one kind of relation might pattern phonologically ("actually") as another. To determine the "actual" (phonological) status of any opposition, one must take into account how it functions in the inventory and in the phonological system more generally. For example, an opposition between /t/ and /d/ could be a privative contrast in the presence or absence of voicing or a privative contrast in the presence or absence of tension, or, if neither member patterns as unmarked, the opposition could be equipollent. Alternatively, if /t/ and /d/ further contrast with another phoneme that differs from both of them along the same dimension (perhaps an even more voiced /^{n}d/ or an even more tense /t'/), the opposition may be gradual.

Figure 2, adapted from Trubetzkoy (1939: 69), shows how different kinds of phonetic oppositions may map to phonological oppositions in Trubetzkoy's system. Notably, oppositions of any phonetic kind may be phonologically equipollent.

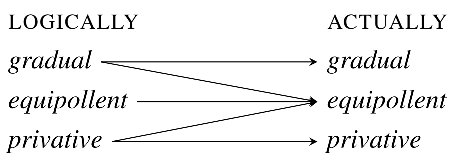


Figure 2: Trubetzkoy's (1939: 69) typology of oppositions

Turning from oppositions to features, we propose to distinguish three kinds of formal representations: equipollent binary features, which represent opposing values of a single (phonetic, morphosyntactic, or semantic) dimension as equally marked; privative monovalent features, in which one value is marked and the other contrastively unmarked; and monovalent modifier features, in which there is a single marked value whose absence is not interpreted as contrastive. We set aside here (but do not necessarily rule out) the possibility of gradual or multivalent features, which have been employed much less frequently than one- and two-valued features. (One domain in which gradient features have been proposed is the representation of vowel height; see Pulleyblank (2011: §5) for an overview of various approaches.)

3.1. *Equipollent binary features*

For equipollent binary features, we follow the familiar notational convention of positive and negative values enclosed in square brackets. Each binary feature represents a single property whose presence and absence function as equally marked in the grammar. Binary features can thus encode some of the kinds of oppositions that Trubetzkoy (1939) described as equipollent, provided that only a single dimension of contrast is involved; others of Trubetzkoy's examples of equipollent oppositions, such as /f/–/k/ or /k/–/l/, would presumably be encoded by multiple features, except perhaps in very sparsely populated phonological inventories.

Binary features are necessary only when both the presence and the absence of a particular property are grammatically active in some sense. What it means to be 'active,' of course, depends on orthogonal assumptions about how the mechanisms of the grammar operate, and, as discussed above, binary features can always or almost always be translated into monovalent ones, though sometimes at the expense of elegance. In general, however, a grammatically active feature is one that is directly referred to and manipulated by rules or constraints. In phonology, for example, a binary feature allows both positive and negative values to spread (or to block spreading) in autosegmental processes. Binary features are thus useful for phenomena such as regressive voicing assimilation in Czech, in which both voiced and voiceless obstruents serve as triggers. An account of regressive assimilation could operate as in (14), with leftward spreading of equipollent binary [±voice] (though cf. Hall (2007) for a monovalent alternative).

(14) Equipollent binary [±voice] as one possible analysis of Czech regressive assimilation

a. /s domem/ → [zdomem] 'with a house'

s domem

[−voice] [+voice]

b. /z pole/ → [spole] 'from a field'

z pole

[+voice] [−voice]

3.2. *Privative monovalent features*

Monovalent features, when contrastive, are privative in Trubetzkoy's (1939) sense. With a privative feature, only one pole of the opposition is linguistically active, but its absence, in a context in which it could have appeared, is interpreted as contrastively signifying the opposite pole.

As a notational convention, we write names of privative monovalent features in SMALL CAPITALS, and, where relevant, we make their contrastive absence explicit by writing (*non*-FEATURE). In the formal computational system, the absence of a privative feature is nothing more than an absence—(*non*-FEATURE) cannot be spread or copied—but at the interface with semantic or phonetic interpretation, the absence of FEATURE is interpreted as the absence of the property that FEATURE signifies. (The contrastive interpretation of absence crucially requires the representation in question to be within the contrastive scope of the privative feature in question, in the sense discussed by Dresher 2009. For example, the absence of privative PAST might be contrastively interpreted as non-past time reference on verbs but not on pronouns, or the absence of privative VOICE might be contrastively interpreted as voicelessness on obstruents but not on sonorants.)

Thus, in a language that distinguishes first, second, and third persons using the two privative features PARTICIPANT and ADDRESSEE, as in (15), a pronominal π with neither of these features is contrastively interpreted as neither first nor second person. A pronominal marked with PARTICIPANT but not with ADDRESSEE is contrastively interpreted as not second person—i.e., as first person.

(15) Privative ϕ-features (Harley and Ritter 2002a)

a. First person

π
|
PARTICIPANT
|
(*non*-ADDRESSEE)

b. Second person

π
|
PARTICIPANT
|
ADDRESSEE

c. Third person

π
|
(*non*-PARTICIPANT)

The same contrastive behaviour can be seen in Cowper's (2005) treatment of viewpoint aspect in English. Events are distinguished from states by the privative feature EVENT, and imperfective viewpoint aspect is further specified by the privative feature INTERVAL, dependent on EVENT. In a clause containing the feature EVENT, the absence of INTERVAL is contrastive, giving an interpretation of perfective viewpoint aspect. However, since INTERVAL is dependent on EVENT, its absence is contrastive only in eventive clauses. Stative clauses thus do not exhibit a viewpoint aspect distinction.

(16) Privative INTERVAL in English (Cowper 2005)

a. Perfective
(*She wrote a book*)

INFL
|
EVENT
|
(*non*-INTERVAL)

b. Imperfective
(*She was writing a book*)

INFL
|
EVENT
|
INTERVAL

c. Stative
(*She knew the answer*)

INFL
|
(*non*-EVENT)

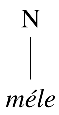
In their interpretation, privative monovalent features are similar to equipollent binary features. The absence of a privative feature F in a context where F could have appeared receives an interpretation equivalent to [−F]. In the morphosyntactic or phonological computation proper, however, as opposed to the interpretive interfaces with semantics or phonetics, privative monovalent features have an inherent asymmetry that equipollent binary features do not: the negative value is structurally not present. The unmarked value (*non*-F) cannot project syntactically, or host other features, or be copied, or spread, or block spreading, or otherwise be grammatically active.

3.3. Monovalent modifier features

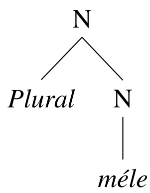
In the literature, however, the absence of a monovalent feature does not always trigger a contrastively complementary interpretation. For example, Wiltschko (2008) shows how monovalent features can occur as non-contrastive modifiers. In Halkomelem, the absence of plural marking on a nominal does not imply singular reference, exemplifying what has traditionally been described as the plural vs. general number contrast.

(17) Optional *Plural* in Halkomelem (Wiltschko 2008: 641–642)

a. méle
child
'child(ren)'



b. má+mele
Plural+child
'children'



There is a structural difference between these two uses of monovalent features. For Wiltschko (2008), when they are non-contrastive they are attached as adjuncts to the head they modify, as shown in (17). When they are contrastive, they are dependents of a functional head, and thus participate systematically in deriving the grammatical paradigms of the language.

The interpretation of an unmarked (i.e., absent) monovalent feature thus depends crucially on a) whether the language makes use of the feature at all, b) if the language does use the feature, whether it is a contrastive feature or an optional modifier feature, and c) if the feature is contrastive, whether the representation in question is within its contrastive scope. Answering these questions requires access to the system of contrasts in the grammar, and is thus not locally obvious in a given representation if absent features are simply not there.

It is therefore useful to make a notational difference between these two uses of monovalent features, while still distinguishing contrastive privative features, with only one active value, from truly binary features, with two active values, as in the examples above. We thus use *italics* to distinguish modifier features, and their absence is simply an absence, with no contrastive default interpretation. Our notational conventions are summarized in Table 1.

Type of feature	Positive	Negative	Negative active?	Negative contrastive?
equipollent binary	[+feature]	[−feature]	yes	yes
privative monovalent	FEATURE	(*non*-FEATURE)	no	yes
monovalent modifier	*Feature*		no	no

Table 1: Summary of notation and interpretation for different types of features

4. The language learner's task

If the reasoning above is correct, the language learner has several things to figure out in the course of arriving at an adult grammar, as follows:[2]

(18) a. Which of the universally available features does the language use?
b. If the language uses a feature F, is it a modifier, or a contrastive feature?
c. If F is contrastive, is it privative or binary?
d. If it is privative, which pole of F is used?

[2]These are questions that must be answered wherever features appear in the grammar, and are thus stated in relatively general terms. There are, in addition, component-specific questions that must be answered as the learner acquires their language. For example, the learner must figure out which syntactic projection hosts a given morphosyntactic feature, and exactly which segments carry a given phonological feature.

This is rather more work than would have been required under Hypothesis 1, and we must therefore consider what innate tools the learner might bring to bear on the task. Along these lines, Blaho (2008: 40) proposes that "Features are not innate, but the ability to make generalisations over data and posit categories is." Following Trubetzkoy (1939), Jakobson (1949), and Dresher (2009), among many others, we assume that learners are innately predisposed to look for contrast in the linguistic input. In other words, learners are innately predisposed to construct grammatical systems of contrast.

But what exactly does it mean to look for, and identify, a linguistic contrast? How is this task different from simple categorical perception, which is found not only in human infants, but in many other animals (Kuhl 1987)? We propose that finding contrast means identifying which of the many differences in the surface forms of utterances have grammatically relevant correlates. Finding linguistic contrasts is thus not just learning to tell things apart. Rather, it involves assigning linguistic significance to the differences by systematically correlating differences at one level with differences at another.

Sometimes the correlation is between a substantive phonetic or semantic property and a grammatical one. For example, the presence or absence of vocal fold vibration signals the difference between one word and another, as in (19).

(19) Phonetic distinction and lexical distinction:
[fil] 'feel' vs. [vil] 'veal'

As shown in (20), nominals with plural denotations trigger a particular form of agreement on the verb, even when the singular and plural nouns have exactly the same form.

(20) Semantic distinction and syntactic agreement:
The sheep is white vs. *The sheep are white.*

However, the correlation does not always involve a transparently substantive phonetic or semantic property. It is the correlation, rather than the substance, that is crucial. Sometimes the correlation is between two abstract properties, such as a phonological pattern and a lexical distinction.

For example, a learner acquiring Nupe must discover that, among the phones realized as [a], there is one that triggers palatalization of a preceding consonant, one that triggers labialization, and one that triggers neither (Hyman 1970). Different phonological behaviours can therefore mark lexical differences just as different phonetic forms can.

(21) Phonological patterning and lexical distinction (Nupe, Hyman 1970: 62):

PALATALIZING			LABIALIZING	
[ēgʲī]	'child'		[ēgʷũ̄]	'mud'
[ēgʲē]	'beer'		[ēgʷó]	'grass'
[ēgʲà]	'blood'	[ēgā] 'stranger'	[ēgʷā]	'hand'

Hyman (1970) argues for an analysis with three contrasting underlying low vowel phonemes /ɛ a ɔ/, all of which surface as [a].[3] The vowel at the end of 'blood' is thus associated with the same feature of /i/ and /e/ that triggers palatalization, and the vowel at the end of 'hand' is associated with the same feature of /u/ and /o/ that triggers labialization.

Similarly, a learner acquiring Northern Alaskan Iñupiaq must discover that there is a 'strong' [i] that palatalizes following consonants, as in (22a), and a 'weak' [i] that (like the [u] in (22c)) does not, as in (22b) (Kaplan 1981, Compton and Dresher 2011).

[3]Alternative analyses would involve positing either underlying sequences that do not follow the usual phonotactics of the language, or whole contrasting series of palatalized and labialized consonants that are neutralized everywhere except before /a/. See also Harms (1973) and Hyman (1973) for further discussion.

(22) Phonological patterning and lexical distinction (Iñupiaq, Compton and Dresher 2011: 206):

a.	[iki]	'wound'	[iki-ʎu]	'and a wound'	[iki-ɲik]	'wounds' (oblique)
b.	[ini]	'place'	[ini-lu]	'and a place'	[ini-nik]	'places' (oblique)
c.	[iɣlu]	'house'	[iɣlu-lu]	'and a house'	[iɣlu-nik]	'houses' (oblique)

The different behaviour of the two kinds of [i] is not predictable from the segmental context. Historically, strong [i] corresponds to Proto-Eskimo *i, and weak [i] to *ə. Synchronically, there is an arbitrary association between particular lexical items with [i] and the presence or absence of palatalization, in the same Saussurean sense in which it is arbitrary that 'house' ends in [u]. The learner must assign some feature to the phonological representations of words such as 'wound' and 'place' that will yield the correct palatalization effects; in Compton and Dresher's analysis, this means positing two distinct underlying phonemes, /i/ and /ə/, distinguished by the presence or absence of the privative feature CORONAL.

To take a morphosyntactic example, a learner acquiring French must learn that nouns belong to classes that have no clear and consistent semantic properties, but which determine the forms of determiners and adjectives in construction with them. The two noun classes sometimes align with natural gender, as in (23a-i) and (23b-i), but in nouns whose referents have no biological sex, they are not predictable either on the basis of meaning, as illustrated in (23a-ii)–(23a-iv) and (23b-ii)–(23b-iv), or on the basis of form, as in (23a-v) and (23b-v).

(23) Lexical class and determiner choice (French):

a.	MASCULINE			b.	FEMININE		
	(i)	*un homme*	'a man'		(i)	*une femme*	'a woman'
	(ii)	*le lit*	'the bed'		(ii)	*la table*	'the table'
	(iii)	*le violon*	'the violin'		(iii)	*la contrebasse*	'the double bass'
	(iv)	*le vocabulaire*	'the vocabulary'		(iv)	*la syntaxe*	'the syntax'
	(v)	*le tour*	'the turn'		(v)	*la tour*	'the tower'

We thus propose that grammatical features, both phonological and morphosyntactic, are the grammatical manifestation of systematic contrasts in a language. These contrasts are identified by the language learner based on correlated patterns at different linguistic levels. This brings us, finally, to the question of whether the set of possible grammatical features is given by UG, or extracted from the linguistic input by the language learner.

5. UG and features

On the assumption that anything attributed specifically to Universal Grammar amounts to a theoretical axiom, we must ask, for any such axiom, what value it adds to the theory, either in terms of empirical gains or in terms of theoretical elegance. The question before us, then, is whether any value is added by positing a universal set of features. The alternative to such a universal set is to claim that learners somehow extract the features from the linguistic input.

First, we have shown that even if the set of possible features is given by Universal Grammar, the language learner must determine which of those features their language uses. As for morphosyntactic features, Cinque and Rizzi (2008) estimate that there are as many as 400 functional projections (i.e., grammatical features) present in natural language. For them, every feature is present in every language; what the language learner needs to do under this view is discern which features are active and which are inert, a task that amounts to essentially the same thing as deciding which features are used and which are not.

In phonology, as discussed above in §2.1, the differences between spoken and signed languages suggest either that it is possible for languages to make use of very different sets of phonological features, or that the phonetic content of features may vary drastically from one language to another. There is also a substantial body of work (e.g., Archangeli 1984; 1988, Archangeli and Pulleyblank 1989, Steriade 1987, Dresher et al. 1994, Dresher 2009, Hall 2007; 2011b, Mackenzie 2009; 2013; see Archangeli 2011 and Hall 2011a for discussion and additional references) arguing that the specification of phonological features on segments varies from one language to another depending on the system of contrasts present in each. If this is correct, then learners must determine what features are specified on what segments in the context of all the contrasts the language makes and the relative scope of each of those contrasts, not simply on whether a given segment appears in the inventory. And if we further take into account the fine-grained phonetic variation between broadly similar (or phonologically analogous) sounds in different languages, it seems likely that learners must also acquire language-specific information about the phonetic implementation of whatever phonological features their language employs.

It is thus not completely obvious either that a universal set of features significantly reduces the task of language acquisition, or that it significantly simplifies the theory itself. While positing a universal set of features permits a purely descriptive account of the fact that many cognitively salient substantive properties, such as 'edible,' 'dangerous,' or 'poisonous,' are never grammaticalized as formal features,[4] it does not explain that fact. If features are emergent rather than given, then we must confront the question of why some substantive properties are grammaticalized and some are not. In the absence of a ready answer to that question, we can say for the moment that the value added by a universal set of features depends on how difficult it is for the learner to extract the features from the input.

We know that very young infants exhibit categorical perception; that is, they have the ability to assign categorial boundaries to perceptual phenomena (Eimas et al. 1971, Eimas 1975). While Eimas et al. took this as evidence for innate, universal distinctive features, this conclusion has since been argued to be unwarranted. First, other animals also exhibit categorical perception in the same way that human neonates do (Kuhl 1987), which means that whatever innate capacity makes this possible is not specific to human language, or even to human beings. In addition, Goldstone et al. (1996) have argued that categorical perception does not, in fact, require the postulation of an innate set of categories, as was originally suggested by Eimas et al. (1971). Rather, novel categories can be learned, and can form the basis for categorical perception. While Goldstone et al.'s (1996) experiment used visual, rather than auditory, stimuli, the results strongly suggest that a pre-existing set of categories (i.e. distinctive features) is not required in order for categorical perception to take place. In other words, if learners already know how to categorize perceived phenomena, then why do they need a (very large) *a priori* set of possible features to choose from? Why not simply make use of the tools they already have, and identify the features in whatever way makes sense?

In fact, Blaho (2008), Mielke (2008), and Samuels (2011), among others, present arguments for acquired or emergent features, though Hale and Reiss (2003: 225) claim that features must be innate, because, as they put it, "Ya gotta start with something!" Essentially, Hale and Reiss assume that perception of linguistic input can only be in terms of pre-existing categories. See also Hall (2010) and Dresher (this volume) for rebuttals of Hale and Reiss's (2003) position, and the references in Hall and Mielke (2011) for further discussion.

It is important to remember, however, that the acquisition of linguistic contrasts and their implementation as grammatical features is not just a matter of categorical perception. As we have claimed, the categorization of phenomena on one level has to be systematically correlated with a categorization on another level. This is what distinguishes simply telling things apart from establishing systematic grammatical contrasts. Let us hypothesize that what is specific to human language is an innate predisposition to correlate categorical distinctions, rather the categories themselves. Taken together with the independent capacity for categorical perception, this innate predisposition permits an account of the acquisition of grammatical contrasts without the need for a universal set of features to choose from.

This move follows the same line of thought as Hauser et al. (2002: 1573), who hypothesize that the narrow faculty of language "comprises only the core computational mechanisms of recursion [i.e., Merge]

[4]It is perhaps worth noting that the Dyirbal *balan* noun class that inspired the title of Lakoff's (1987) *Women, Fire, and Dangerous Things* also includes fireflies, crickets, platypuses, and water; it seems much more like a typical semantically arbitrary noun class with tantalizing hints of regularity than like a clear case for grammaticalized dangerousness.

as they appear in narrow syntax and the mappings to the interfaces." While their proposal addresses the question of the mechanisms that are responsible for building linguistic structures, it does not address the question of where the elements that participate in the mechanisms come from. Given two linguistic objects, Merge provides a means of constructing a more complex linguistic object, but Merge says nothing about what the basic linguistic objects consist of. We propose that they consist precisely of the contrastive features that have been acquired based on correlated differences in the input. Further, we claim that the ability to posit formal features from correlated categorical distinctions, along with Merge, is all that Universal Grammar provides as the basis for narrow syntax. As far as the basic building blocks of linguistic representations go, then, Saussure was essentially correct in saying that "there are only differences."

For phonological features, the same basic acquisition mechanism applies: the learner identifies phonological features by correlating differences in lexical meaning, differences in phonetic realization, and differences in phonological behaviour. Dresher et al. (1994), inspired by Jakobson and Halle (1956), propose an algorithm for constructing phonological representations that successively divides the phonemic inventory by assigning contrastive feature specifications; the correlation-of-differences mechanism complements this algorithm by inducing the features to be assigned.

Our central hypothesis is thus as stated in (24).

(24) ***Reductiō ad discrīmen:*** *The ability to search for systematic contrast in the linguistic input, by correlating differences at various levels, is the only mechanism required to account for the abstract building blocks that make up those mental structures: the formal features of grammatical systems.*

This minimalist view of the contribution of UG to formal representations gives rise to some new typological questions. If features are not directly supplied by UG itself, why are languages so similar in the sets of formal features they use? Is there any limit to what substantive properties, semantic or phonetic, can play a role in linguistic contrast, and, if so, where do those limits come from? We suspect that the answers to these questions lie in the properties of the conceptual–intentional (C–I) and articulatory–perceptual (A–P) systems with which the grammar must be able to communicate (Chomsky 1995).

At the C–I interface, the semantic interpretation of sentences imposes certain consistent requirements on syntactic representations. For example, illocutionary force is an essential part of the interpretation of propositions, and so the features of the C system include content that is relevant for clause typing (Cheng 1991). To be evaluated, propositions must be deictically anchored. The features that perform this anchoring may be temporal, personal, or locative (Ritter and Wiltschko 2009), and there is also variation among different kinds of temporal anchoring systems (Cowper 2005), but the shared property of deictic anchoring identifies an apparently universal class of Infl features. In the internal structure of events, we find that languages in general have features that represent inner aspect (Clarke 2013) and that introduce arguments (Kratzer 1996, Pylkkänen 2008); the cross-linguistic recurrence of these kinds of features presumably reflects the way events are structured in the C–I system. If meaning in general is constructed through functional application (Frege 1891), we need features that identify the semantic types of nominals (Ghomeshi et al. 2009) and clauses, and mechanisms such as Case and agreement that track the relations between predicates and their arguments (Nichols 1986).

On the A–P side, properties of the human vocal tract and auditory system determine the range of articulatory and acoustic properties that can ground phonological features of oral languages. Analogously, properties of the human upper-body anatomy and the human visual system determine range of physical movements and positions that can ground phonological features of signed languages. Mielke (2008) uses the fact that these limitations are inherent in the human body to argue that they need not also be stipulated in UG, and that phonological features can be emergent rather than innate. In Mielke's emergent feature theory, it is also possible for features to be based solely on phonological patterning; a feature can simply pick out "the segments that do X" (Mielke 2008: 99). Mielke primarily attributes the fact that most phonological patterns involve phonetically natural classes to the way in which such patterns develop diachronically; under the view of features we have proposed here, the prevalence of natural patterns may also follow from the fact that the learner posits features not through phonological patterning alone, but by correlating differences in phonological patterning with differences in phonetic realization and/or lexical meaning. Cross-linguistic consistencies thus arise, both in phonology and in morphosyntax, from a combination of interface requirements and properties of the learning mechanism.

References

Abou Haidar, Laura. 1994. Norme linguistique et variabilité dialectale: Analyse formantique du système vocalique de la langue arabe. *Revue de Phonétique Appliquée* 110: 1–15.

Archangeli, Diana. 1984. *Underspecification in Yawelmani Phonology and Morphology*. Ph.D. thesis, Massachusetts Institute of Technology.

Archangeli, Diana. 1988. Underspecification in phonology. *Phonology* 5: 183–207.

Archangeli, Diana. 2011. Feature specification and underspecification. In *The Blackwell Companion to Phonology*, edited by Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice, vol. 1, pp. 148–170. Wiley-Blackwell, Oxford.

Archangeli, Diana and Douglas Pulleyblank. 1989. Yoruba vowel harmony. *Linguistic Inquiry* 20 2: 173–217.

Backley, Phillip. 2011. *An Introduction to Element Theory*. Edinburgh University Press, Edinburgh.

Becker, A. L. and I Gusti Ngurah Oka. 1974. Person in Kawi: Exploration of an elementary semantic distinction. *Oceanic Linguistics* 13 1/2: 229–255.

Béjar, Susana. 2003. *Phi-Syntax: A Theory of Agreement*. Ph.D. thesis, University of Toronto.

Béjar, Susana. 2008. Conditions on phi-agree. In *Phi Theory: Phi-Features across Modules and Interfaces*, edited by Daniel Harbour, David Adger, and Susana Béjar, pp. 130–154. Oxford University Press, Oxford.

Béjar, Susana and Daniel Currie Hall. 2000. Marking markedness: The underlying order of diagonal syncretisms. In *Proceedings of the Eastern States Conference on Linguistics (ESCOL) 1999*, edited by Rebecca Daly and Anastasia Riehl, pp. 1–12. Cornell Linguistics Circle Publications, Ithaca, NY.

Blaho, Sylvia. 2008. *The Syntax of Phonology: A Radically Substance-Free Approach*. Ph.D. thesis, Universitetet i Tromsø.

Brentari, Diane. 1998. *A Prosodic Model of Sign Language Phonology*. MIT Press, Cambridge, MA.

Brentari, Diane. 2011. Handshape in sign language phonology. In *The Blackwell Companion to Phonology*, edited by Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice, vol. 1, pp. 195–222. Wiley-Blackwell, Oxford.

Cable, Seth. 2010. Number categories beyond singular and plural. Ms., University of Massachusetts, Amherst.

Calabrese, Andrea. 1995. A constraint-based theory of phonological markedness and simplification procedures. *Linguistic Inquiry* 26 3: 373–463.

Casali, Roderic F. 2003. [ATR] value asymmetries and underlying vowel inventory structure in Niger-Congo and Nilo-Saharan. *Linguistic Typology* 7 3: 307–382.

Casali, Roderic F. 2008. ATR harmony in African languages. *Language and Linguistics Compass* 2 3: 496–549.

Cheng, Lisa Lai-Shen. 1991. *On the Typology of Wh-Questions*. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. MIT Press, Cambridge, MA.

Chomsky, Noam. 1995. *The Minimalist Program*. MIT Press, Cambridge, MA.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by Step*, edited by Roger Martin, David Michaels, and Juan Uriagereka, pp. 89–155. MIT Press, Cambridge, MA.

Chomsky, Noam and Morris Halle. 1968. *The Sound Pattern of English*. Harper and Row, New York.

Cinque, Guglielmo and Luigi Rizzi. 2008. The cartography of syntactic structures. *CISCL Working Papers* 2: 42–58.

Clarke, Sarah. 2013. *Aspectual Scope and Contrast in English and Japanese*. Ph.D. thesis, University of Toronto.

Clements, G. N. and Elizabeth Hume. 1995. The internal organization of speech sounds. In *The Handbook of Phonological Theory*, edited by John Goldsmith, pp. 245–306. Blackwell, Oxford.

Compton, Richard and B. Elan Dresher. 2011. Palatalization and "strong *i*" across Inuit dialects. *Canadian Journal of Linguistics* 56 2: 203–228.

Corbett, Greville G. 2000. *Number*. Cambridge University Press, Cambridge.

Cowper, Elizabeth. 2005. The geometry of interpretable features: Infl in English and Spanish. *Language* 81 1: 10–46.

Cowper, Elizabeth and Daniel Currie Hall. 2013. A neoparametric approach to variation and change in English modals. Presented at GLOW 36, Lund University, April 2013.

Dresher, B. Elan. 2009. *The Contrastive Hierarchy in Phonology*, vol. 121 of *Cambridge Studies in Linguistics*. Cambridge University Press, Cambridge.

Dresher, B. Elan. 2014. The arch not the stones: Universal feature theory without universal features. *Nordlyd* 41 2: 165–181.

Dresher, B. Elan, Glyne Piggott, and Keren D. Rice. 1994. Contrast in phonology. *Toronto Working Papers in Linguistics* 13: iii–xvii.

Eimas, Peter D. 1975. Speech perception in early infancy. In *Infant Perception: From Sensation to Cognition*, edited by Leslie B. Cohen and Philip Salapatek, pp. 193–231. Academic Press, New York.

Eimas, Peter D., Einar R. Siqueland, Peter Jusczyk, and James Vitorito. 1971. Speech perception in infants. *Science* 171 3698: 303–306.

Ewen, Colin J. and Harry van der Hulst. 1985. Single-valued features and the nonlinear analysis of vowel harmony. In *Linguistics in the Netherlands 1985*, edited by Hans Bennis and Frits Beukema, pp. 39–48. Foris, Dordrecht.

Frantz, Donald G. 1971. *Toward a Generative Grammar of Blackfoot*. Summer Institute of Linguistics of the University of Oklahoma, Norman, OK.

Frege, Gottlob. 1891. *Function und Begriff*. Hermann Pohle, Jena.

Ghomeshi, Jila, Ileana Paul, and Martina Wiltschko (eds.) . 2009. *Determiners: Variation and Universals*. John Benjamins, Amsterdam.

Goldsmith, John. 1981. Subsegmentals in Spanish phonology: An autosegmental approach. In *Linguistic Symposium on Romance Languages 9*, edited by William Cressey and Donna Jo Napoli, pp. 1–16. Georgetown University Press, Washington.

Goldsmith, John. 1985. Vowel harmony in Khalkha Mongolian, Yaka, Finnish and Hungarian. *Phonology Yearbook* 2: 253–275.

Goldstone, Robert L., Mark Steyvers, and Kenneth Larimer. 1996. Categorical perception of novel dimensions. In *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society*, pp. 243–248. Lawrence Erlbaum Associates, Hillsdale, N.J.

Hale, Mark and Charles Reiss. 2003. The subset principle in phonology: Why the *tabula* can't be *rasa*. *Journal of Linguistics* 39 2: 219–244.

Hall, Daniel Currie. 2007. *The Role and Representation of Contrast in Phonological Theory*. Ph.D. thesis, University of Toronto.

Hall, Daniel Currie. 2009. Laryngeal neutralization in Breton: Loss of voice and loss of contrast. In *Proceedings of the 2009 CLA*. Canadian Linguistic Association.

Hall, Daniel Currie. 2010. The check-list: Halle's Condition 2 and the form and meaning of phonological features. Invited talk presented at the Sixth North American Phonology Conference (NAPhC 6), Concordia University, Montréal, May 2010.

Hall, Daniel Currie. 2011a. Contrast. In *The Blackwell Companion to Phonology*, edited by Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice, vol. 1, pp. 27–53. Wiley-Blackwell, Oxford.

Hall, Daniel Currie. 2011b. Phonological contrast and its phonetic enhancement: Dispersedness without dispersion. *Phonology* 28 1: 1–54.

Hall, Daniel Currie and Jeff Mielke. 2011. Distinctive features. Annotated bibliography published in *Oxford Bibliographies Online* at http://dx.doi.org/10.1093/OBO/9780199772810-0012.

Halle, Morris. 1959. *The Sound Pattern of Russian*. Mouton, The Hague.

Halle, Morris, Bert Vaux, and Andrew Wolfe. 2000. On feature spreading and the representation of place of articulation. *Linguistic Inquiry* 31 3: 387–444.

Harbour, Daniel. 2011a. Paucity, abundance, and the theory of number. Ms., Queen Mary University of London.

Harbour, Daniel. 2011b. Valence and atomic number. *Linguistic Inquiry* 42 4: 561–594.

Harbour, Daniel. 2013. "Not Plus" isn't "Not There": Bivalence in person, number, and gender. In *Distributed Morphology Today: Morphemes for Morris Halle*, edited by Ora Matushansky and Alec Marantz, chap. 8, pp. 135–150. MIT Press, Cambridge, MA.

Harbour, Daniel and Christian Elsholtz. 2012. Feature geometry: Self-destructed. Ms., Queen Mary University of London and Technische Universität Graz.

Harley, Heidi. 1994. Hug a tree: Deriving the morphosyntactic feature hierarchy. *MIT Working Papers in Linguistics* 21: 289–320.

Harley, Heidi and Elizabeth Ritter. 2002a. Person and number in pronouns: A feature-geometric analysis. *Language* 78 3: 482–526.

Harley, Heidi and Elizabeth Ritter. 2002b. Structuring the bundle: A universal morphosyntactic feature geometry. In *Pronouns: Grammar and Representation*, edited by Horst J. Simon and Heike Wiese, pp. 23–39. John Benjamins, Amsterdam.

Harms, Robert T. 1973. How abstract is Nupe? *Language* 49 2: 439–446.

Harris, John and Geoff Lindsey. 1995. The elements of phonological representation. In *Frontiers of Phonology*, edited by Jacques Durand and Francis Katamba, pp. 34–79. Longman, Harlow.

Hauser, Marc D., Noam Chomsky, and W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298: 1569–1579.

Herd, Jonathon. 2005. Loanword adaptation and the evaluation of similarity. *Toronto Working Papers in Linguistics* 24: 65–116.

Hyman, Larry M. 1970. How concrete is phonology? *Language* 46 1: 58–76.

Hyman, Larry M. 1973. Nupe three years later. *Language* 49 2: 447–452.

Iosad, Pavel. 2012. *Representation and Variation in Substance-Free Phonology: A Case Study in Celtic*. Ph.D. thesis, University of Tromsø.

Iverson, Gregory K. and Joseph C. Salmons. 2003. Legacy specification in the laryngeal phonology of Dutch. *Journal of Germanic Linguistics* 15 1: 1–26.

Jakobson, Roman. 1949. On the identification of phonemic entities. *Travaux du Cercle linguistique de Copenhague* 5: 205–213.

Jakobson, Roman, C. Gunnar M. Fant, and Morris Halle. [1951] 1969. *Preliminaries to Speech Analysis: The Distinctive Features and Their Correlates*. MIT Press, Cambridge, MA. 9th printing.

Jakobson, Roman and Morris Halle. 1956. *Fundamentals of Language*. Mouton, The Hague.

Kaplan, Lawrence D. 1981. *Phonological Issues in North Alaskan Inupiaq*. Native Language Center, Fairbanks.

Krämer, Martin. 2000. Voicing alternations and underlying representations: The case of Breton. *Lingua* 110 9: 639–663.

Kratzer, Angelika. 1996. Severing the internal argument from its verb. In *Phrase Structure and the Lexicon*, edited by Johan Rooryck and Laurie Zaring, pp. 109–137. Kluwer, Dordrecht.

Kuhl, Patricia K. 1987. The special-mechanisms debate in speech perception: Nonhuman species and nonspeech signals. In *Categorical Perception: The Groundwork of Cognition*, edited by Stevan Harnad, pp. 355–386. Cambridge University Press, Cambridge.

Kyriakaki, Maria. 2006. *The Geometry of Tense, Mood and Aspect in Greek*. M.A. thesis, University of Toronto.

Lakoff, George. 1987. *Women, Fire, and Dangerous Things*. University of Chicago Press, Chicago.

Mackenzie, Sara. 2009. *Contrast and Similarity in Consonant Harmony Processes*. Ph.D. thesis, University of Toronto.

Mackenzie, Sara. 2013. Laryngeal co-occurrence restrictions in Aymara: Contrastive representations and constraint interaction. *Phonology* 30 2: 297–345.

McGinnis, Martha. 2005. On markedness asymmetries in person and number. *Language* 81 3: 699–718.

McGinnis, Martha. 2008. Phi-feature competition in morphology and syntax. In *Phi Theory: Phi-Features across Modules and Interfaces*, edited by Daniel Harbour, David Adger, and Susana Béjar, pp. 155–184. Oxford University Press, Oxford.

Mielke, Jeff. 2008. *The Emergence of Distinctive Features*. Oxford University Press, Oxford.

Morén, Bruce. 2003. The Parallel Structures Model of feature geometry. *Working Papers of the Cornell Phonetics Laboratory* 15:

194–270.

Morén, Bruce. 2006. The division of labor between segment-internal structure and violable constraints. In *Freedom of Analysis?*, edited by Sylvia Blaho, Patrik Bye, and Martin Krämer, pp. 313–344. Mouton, The Hague.

Nevins, Andrew Ira. 2010. *Locality in Vowel Harmony*. MIT Press, Cambridge, Mass.

Nichols, Johanna. 1986. Head-marking and dependent-marking grammar. *Language* 62 1: 56–119.

Pulleyblank, Douglas. 2011. Vowel height. In *The Blackwell Companion to Phonology*, edited by Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice, vol. 1, pp. 491–518. Wiley-Blackwell, Oxford.

Pylkkänen, Liina. 2008. *Introducing Arguments*. MIT Press, Cambridge, MA.

Ritter, Elizabeth and Martina Wiltschko. 2009. Varieties of INFL: Tense, location, and person. In *Alternatives to Cartography*, edited by Jeroen van Craenenbroeck, pp. 153–202. Walter de Gruyter, Berlin.

Sagey, Elizabeth C. 1990. *The Representation of Features and Relations in Non-Linear Phonology*. Garland Publishing, New York.

Samuels, Bridget D. 2011. *Phonological Architecture: A Biolinguistic Perspective*, vol. 2 of *Oxford Studies in Biolinguistics*. Oxford University Press, Oxford.

Sandler, Wendy. 1989. *Phonological Representation of the Sign: Linearity and Nonlinearity in American Sign Language*. Foris, Dordrecht.

Saussure, Ferdinand de. 1916. *Cours de linguistique générale*. Payot, Paris. Edited by Charles Bally and Albert Sechehaye.

Saussure, Ferdinand de. 1959. *Course in General Linguistics*. Philosophical Library, New York. Translated from Saussure (1916) by Wade Baskin.

Schütz, Albert J. 1995. *The Voices of Eden: A History of Hawaiian Language Studies*. University of Hawai'i Press, Honolulu.

Steriade, Donca. 1987. Redundant values. *Chicago Linguistic Society* 23 2: 339–362.

Trubetzkoy, N. S. 1939. Grundzüge der Phonologie. *Travaux du Cercle Linguistique de Prague* 8.

Wetzels, Leo and Joan Mascaró. 2001. The typology of voicing and devoicing. *Language* 77 2: 207–244.

Wiltschko, Martina. 2008. The syntax of non-inflectional plural marking. *Natural Language & Linguistic Theory* 26 3: 639–694.