Emergent Syntax:

Insights from Imperatives

Richard Stockwell

Pembroke College

Department of Theoretical and Applied Linguistics

University of Cambridge

4th June 2015

This dissertation is submitted for the degree of

Master of Philosophy.

Contents

Preface

Abstract

A generative theory of emergent syntax aims to reduce UG to a more plausible size, with increased roles for the PLD and a domain-general Minimax acquisition bias. I examine the syntax of imperatives in this light, advancing an analysis that captures the centrality of the addressee. Imperatives provide acquirers with crucial information about their language, from which they actively generalise.

Word Count

This thesis contains 20,481 words, an allowance having been granted for glosses and translations.

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text.

List of abbreviations

ABS	absolutive
ACC	accusative
ADDR	addressee situation participant
ADS	(adult-to-)adult-directed speech
ANTIPASS	antipassive
Asp	Aspect head
AUX	null modal auxiliary
BCC	Borer-Chomsky Conjecture
BE	Belfast English
BoulP	Boulemaeic force head
C	Complementizer head
c	language-specific category
CDS	child-directed speech
[coin]	coincidence feature
COMP	Complementizer
COND	conditional
DAT	dative
DEF	definite
e-adverbs	extent-adverbs
ERG	ergative
Eval-sit	evaluation situation
Ev-sit	event situation
EXH	exhortative
F	a functional head
[F]	a formal feature
FE	Feature Economy
FFFH	Flexible Formal Features Hypothesis
Fin	Finiteness head
FUT	future
f-valuation	valuation by a higher functional head
GEN	genitive
I	Inflection head
[IF]	imperative/illocutionary force
[iF]	an interpretable formal feature
IG	Input(/Intake) Generalisation
IMP	imperative
[imp]	imperative feature
IMP(neutral)	neutral imperative
INF	infinitive
INFL	Inflection
J	Jussive head

LF	Logical Form interface
LPH	'light' performative hypothesis
MASC	masculine gender
MP	Minimalist Program
MUG	Minimal UG
m-valuation	morphological-valuation
NBH	No Base Hypothesis
NEG	negator
[NEG]	formal negative feature
Neg	negative head
NOM	nominative
NOMZ	nominalizer
NONFIN	non-finite
NSP	Null Subject Parameter
OBL	oblique
Op¬	covert negative operator
p	propositional content
PAST	past tense
PH	Performative Hypothesis
PL	plural
PLD	Primary Linguistic Data
POL	politeness marker
Pov	Point-of-view head
pred-valuation	predicate-valuation
PRF	perfect
PRES	present tense
PRM	promissive
Pro-sit	pronominal situation argument
PRS	prospective suffix
PSH	Parametric Substantiation Hypothesis
P&P	Principles and Parameters
QT	quotative particle
ReCoS	Rethinking Comparative Syntax
REFL	reflexive
RFH	Rich Functional Hierarchy
S_0	initial state of the language faculty
S_S	steady state adult grammar
SBJV	subjunctive
SG	singular
SOFT	soft imperative
SP	Subset Principle
T	Tense head
t	movement trace
TNI	True Negative Imperative
TRANS	transitive

UBH	Universal Base Hypothesis
[*u*coin]	unvalued coincidence feature
[uF]	an uninterpretable formal feature
UG	Universal Grammar
UoL	Unit of Language
USH	Universal Spine Hypothesis
V	lexical Verb
v	little v head
v*	external argument-introducing little v
WALS	World Atlas of Language Structures
XP	phrasal projection of X
YOU	imperative subject
*	ungrammatical
¬	not
∅	empty head
>	takes scope over
2	second person
3	third person
κ	universal category
Π	sound
Σ	meaning
ΣP	polarity phrase

List of figures

List of tables

Emergent syntax: insights from imperatives

1   Introduction

This thesis considers imperatives from the perspective of emergent syntax. Neo-emergentism seeks to radically reduce the language-specific innate content in Universal Grammar (UG). A reduced role for UG is urged by considerations of parsimony, evolutionary plausibility, and a tractable acquisition mechanism. Compensating for this reduction requires an increased role for domain-general acquisition mechanisms, and thorough consideration of what Primary Linguistic Data (PLD) provides. I consider imperatives in this light. Structurally, I argue for the centrality of the addressee in imperatives in substantiating a point-of-view position between I and v. This and other aspects of imperatives provide crucial acquisition evidence, with the hallmarks of emergentism in the generalisations acquirers make from what they find.

In outline, §2 introduces and refines a generative theory of emergent syntax, as advanced by some of the research on the Cambridge Rethinking Comparative Syntax (ReCoS) project (European Research Council Advanced Grant No. 269752). §3 justifies the choice of imperatives to develop this theory, based on their relative frequency in contrast with other clause-types, useful featural content, and evidence of their early analysis by acquirers. §4 evaluates which aspects of previous formal analyses of imperatives are retained and which discarded in the analysis of imperatives in §5. This analysis comprises standard affirmative imperatives, negative imperatives, hortatives, and the novel category of gerund imperatives. §6 explores the consequences of this analysis from a neo-emergentist perspective. Without appealing to a rich innate component, we will see what imperatives signal to acquirers about their language, and how acquirers manipulate and generalise this information. §7 concludes and considers directions for future research.

2   Emergent syntax

This section outlines the theory of emergent syntax which I aim to develop here. Traditional theories of UG seem wedded to nativism, but have failed to offer a robust theory for how pre-given material is successfully deployed in acquisition. This failure might be taken as a victory for

traditional emergentist theories (e.g. O'Grady 2005; Goldberg 1999). However, a better alternative for acquisition that is both emergent and generative is emerging from ReCoS research on parameter hierarchies.[1] The conception of this theory was made possible by theoretical advances regarding features and lexicocentric parameters from Principle and Parameters (P&P) theory (Chomsky 1981) to Minimalism (Chomsky 1995), and by embracing the role of domain-general factors (Chomsky 2005) in acquisition. I finish by offering some refinements to the format of ReCoS-ian parameter hierarchies.

2.1 Universal Grammar

Traditional UG theories claim that acquisition amounts to the setting of parameters, which are innately given. However, no satisfactory solution has been found to the implementation problem of how an acquirer links up abstract parameters with concrete input data.

Chomsky (1965) developed UG in response to behaviourist theories of language, notably Skinner (1957), marking the advent of the cognitive revolution in linguistics. UG offers innate guidance in language acquisition in the face of the poverty of the stimulus. The stimulus underdetermines finite mental grammars, which have infinite generative capacity. However, the theory struggled early on to attain Chomsky's (1965) standard of explanatory adequacy – explaining acquisition – while equipped with only a vague evaluation metric for selecting among grammars.

The goal of explanatory adequacy appeared within grasp with the advent of the P&P theory of UG (Chomsky 1981), which reduced acquisition to selecting among UG-given (binary) parametric choices (Guasti 2002:19). For example, the Head Parameter (Koopman 1984) parameterises the principle of X-bar theory (Chomsky 1970; Jackendoff 1977) in the phrase structure module of the grammar. Equipped with this parameter, the child need only select from the head-initial or head-final options provided by UG on the basis of simple input data; e.g., for English, *John kissed Mary* indicates head-initial order. The acquisition task regarding phrase structure would then be complete.

The potential for parametric theory to attain explanatory adequacy was concretised with specific reference to acquisition by Hyams (1983, 1986) for the Null Subject Parameter (NSP). However, the initial promise of parameters began to unravel as more empirical data came to light, and conceptual learnability issues were scrutinised. Empirically, the predictive power of parameters regarding clusters of properties was shown to be more limited than first hoped; for example, Gilligan (1987) on

---

[1] Though this line on emergent features is not shared by all members of the ReCoS group. For example, Sheehan (to appear:39) continues to assume that formal features and functional categories are given by UG.

free inversion and the NSP (Rizzi 1982). Instead, microparametric research came strongly to the fore (Kayne 1989 et seq.).

Conceptually, questions were raised regarding the implementation of parameter setting. First there is the Linking Problem (Baker 1979, Pinker 1984, Biberauer et al. 2014:105), also known as the Epistemological Problem (Dresher 1999), of how abstract parameters could be related to concrete input data. Any solution to the Linking Problem seems to require adding further innate information to parameters. Lightfoot (1991) associates each parameter with a cue based on which it could be set. The notion of parameter expression (Clark 1992; Clark and Roberts 1993; Roberts and Roussou 2003) reconceptualises cues as triggers, substrings which can only be analysed when a parameter is set to a particular value (cf. Fodor 1998 on unambiguous triggers). But this just pushes the Linking Problem up a level: how does the acquirer know that an input string is indeed a trigger?

The second conceptual implementation problem for parameter setting is Dresher's (1999) Credit Problem: how does the acquirer know which parameter is responsible for which effect in the input? One potential response is that parameters could be reset at random when faced with an unanalysable input, as implemented in Gibson and Wexler's (1994) Triggered Learning Algorithm. However, certain orderings of parameter setting were found to lead the acquirer into local maxima, incorrect settings of parameter values from where it became impossible to converge on the target. Gibson and Wexler found this effect when considering only three parameters, and Kohl (1999) demonstrates that local maxima become more likely as more parameters are included. These findings suggested that the ordering of parameters must be fixed, writing yet another piece of innate information into UG.

Overall, acquisition theories in the P&P framework found themselves writing parameters into UG in a rigidly fixed order, along with associated triggers or cues, and all still without resolving the Linking Problem. Faced with this conundrum, perhaps it would be better to discard language-specific innate content entirely?

2.2   Emergentism

Emergentism captures a broad range of theories, effectively defined by their opposition to Chomskyan UG. While it is universally accepted that acquisition is innately guided (O'Grady 2008:620), emergentists claim that there is nothing faculty-specific about language acquisition. Instead, emergentists invoke exclusively domain-general "simple learning mechanisms (essentially, inductive generalizations) that extract statistical regularities from experience" (O'Grady 2010:275). With reference to syntax, two examples of emergentist theories include O'Grady's (2005) theory of

computational routines, whereby all syntactic patterns derive from the linear operation of an efficiency-driven processor (cf. Hawkins 1994, 2004); and Construction Grammar (Goldberg 1999, Tomasello 2003), which reduces syntax to stored pairings of form and function.

However, language goes far beyond pattern-matching. As we will see, much of syntax consists in the ways form does not map straightforwardly onto function. An attractive way of capturing the systematicity of language while circumventing the Linking Problem would be to strip UG of much of its content, leaving behind the basic tools for an acquirer to construct a formal syntactic system for themselves. This approach is the one explored in some recent ReCoS work, and its spirit is adopted here. Conceiving of UG and acquisition in this way was a step made possible by the theoretical advances of the Minimalist Program (MP) (Chomsky 1995) and beyond (Chomsky 2005).

2.3   Minimalism and feature-based syntax

The minimalist emphasis on features offered an alternative to the view of parameters as binary choice points tied to principles. On the lexicocentric view, parameters[2] and the cross-linguistic variation they regulate are associated with the features of lexical items. This position is known as the Borer-Chomsky Conjecture (BCC), which was introduced by Borer (1984) and Fukui (1986), adopted by Chomsky (1995), and formulated by Baker (2008:353):

(1) The BCC:

All parameters of variation are attributable to differences in the features of particular items (e.g., the functional heads) in the lexicon.

The BCC collapses together formal features and parameters. However, the BCC does not in itself take us further regarding the innateness of parameters, because the standard assumption remains that features are innately given by UG. Theories usually follow Chomsky's (2001:10) "conventional assumption" that the language faculty specifies a set of features, from which there is a "one-time selection" for a particular language. Some theorists strengthen this position to the extent of claiming that every language manifests the same set of grammatical features, e.g. Miyagawa's (2010) Strong Uniformity Thesis.

However, this standard view of innate features raises learnability, methodological, and evolutionary problems. First, invoking innateness still does not account for how abstract, UG-given features are connected up with concrete data; enter the Linking Problem once again. Second, MP in its methodological guise urges us to posit as little as possible, in the name of doing good science

---

[2] Where they are considered to retain a role in the research program; see e.g. Boeckx (2014) in opposition.

(Chomsky 1995). Third, MP in its substantive guise urges us to go beyond explanatory adequacy (Chomsky 2004) in asking why the human language faculty is the way it is. One aspect of this metaphysical question concerns evolutionary plausibility. It is highly implausible that very much content could have entered the human genome during an evolutionary window in the order of a hundred thousand years (e.g. Hornstein 2009).

All this urges the pursuit of the idea that syntactic features are not innate. This idea is not all that radical in the context of theoretical linguistics in view of progress in phonology (e.g. Mielke 2008; Dresher 2009). The approach taken here argues that features, and thereby, by BCC, parametric variation, are emergent, incorporating roles for the data and general cognitive biases, while retaining a crucial role for a minimal UG. Before introducing this approach in detail, I consider three theories that took steps in this direction.

2.4   Babysteps towards emergent features

The idea that features are emergently constructed by the acquirer has three important theoretical antecedents: Hegarty (2005), Gianollo, Guardiano and Longobardi (2008), and Zeijstra (2008).

2.4.1   Hegarty (2005)

Hegarty argues that functional categories, such as C or T, are not primitives, but vessels for bundles or matrices of features. Language acquisition consists in determining the correct bundlings of features into matrices, which are then merged into the syntax as projecting functional heads. Thus acquisition involves constructing functional heads from primitive features. Here we will see that the features can be constructed too.

2.4.2   Gianollo, Guardiano and Longobardi (2008)

Gianollo, Guardiano and Longobardi propose a parameter schema for the acquisition of formal features. The schema aims to account for cross-linguistic variation in the presence and behaviour of features in a language. It consists of five questions. The first asks whether or not a feature [F] is grammaticalised in the language. If it is, then four further questions specify the spreading and checking behaviour of the feature.

However, features are not truly emergent on this view. The domain of [F]s considered by the schema is the set of innate, UG-given features. Thus Gianollo, Guardiano and Longobardi provide a mechanism for how Chomsky's (2001:10) "one-time selection" might work, but, like Hegarty, they do not question the primitive status of features themselves.

2.4.3 Zeijlstra (2008)

Zeijlstra (2008) goes further than Hegarty and Gianollo, Guardiano and Longobardi in proposing that the set of syntactic formal features in a language is derived during acquisition. His proposal adopts the standard minimalist (Chomsky 1995, 2000, 2001) conception of lexical items as feature bundles, comprised of phonological, formal syntactic, and semantic features. As diagrammed in Figure 1, formal features overlap with semantic features as interpretable features [iF], which are legible at the Logical Form (LF) interface. By contrast, uninterpretable formal features [uF] have no semantic content, so are illegible at LF. By the Principle of Full Interpretation (Chomsky 1995), uninterpretable features must therefore be deleted before reaching the interface.
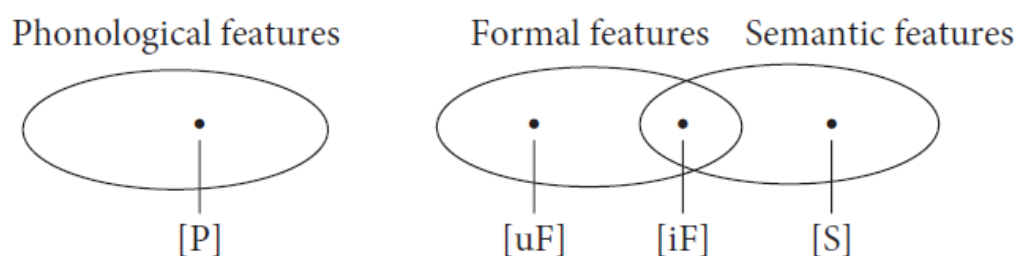
Figure 1: Zeijlstra's (2008:146) minimalist conception of the features of lexical items

According to Zeijlstra's (2008) Flexible Formal Features Hypothesis (FFFH), formal features are posited by the child on the basis of doubling phenomena in the PLD; without such evidence, a feature remains a purely semantic feature.

Features are still not entirely emergent on Zeijlstra's approach. FFFH raises some parameters beyond the BCC to a level higher than the features of lexical items: to whether a given semantic operator is grammaticalised or not. However, the search space for potential formal features is the set of semantic operators, which are (implicitly) innate. Thus Zeijlstra (2008:169) characterises the acquisition procedure: "Each grammar makes a particular selection of semantic operators that can be realised as [Functional Projections] based on the language input available during [first language] acquisition." Unlike the standard (Chomsky 2001) position of Hegarty and Gianollo, Guardiano and Longobardi, Zeijlstra derives formal features from a syntax-external source; but the source is still an innate set of features, which may even be internal to the language faculty. By contrast, the approach here will be fully emergent in granting the grammaticalisation process access to domain general cognition (cf. Ramchand and Svenonius 2014 and §6.6).

2.4.4 Summary

These three proposals demonstrate the potential within a lexicocentric minimalist framework to give the acquirer a greater role in determining which features are involved, and how, in their language. To advance the position that features themselves are emergent requires embracing a final theoretical advance regarding the role played by domain-general factors in acquisition.

2.5 The third factor

Prior to Chomsky (2005), acquisition was conceived of as guided by the interaction between UG and the PLD. The initial state of the language faculty (S_0), which amounted to UG, developed spontaneously when exposed to the PLD through a series of intermediate grammars (S_1, S_2, …) before reaching a steady state adult grammar (S_S). To UG and the PLD, Chomsky (2005:6)[3] adds a third factor, consisting of "Principles not specific to the faculty of language"; in particular, "(a) principles of data analysis that might be used in language acquisition and other domains; [and] (b) principles of structural architecture and developmental constraints … including principles of efficient computation". With the introduction of this third factor, on top of a lexicocentric minimalist framework, the way is paved for a truly neo-emergentist theory of syntactic acquisition.

2.6 ReCoS and the three factors

The ReCoS-ian view developed here brings together lexicocentric minimalism and Chomsky's (2005) three factors into a theory of emergent features. It adopts the lexicocentric BCC by situating the acquisition of features within lexical learning. This major advantage of the BCC follows from Borer's (1984:29) original observation: "Associating parameter values with lexical entries reduces them to the one part of a language which clearly must be learned anyway: the lexicon."

To learn the lexicon of their language, children must discover the arbitrary pairings between words and meanings. Words are signs in the Saussurean (1916) sense that word forms arbitrarily signify meanings. In standard emergentist theories, such as Construction Grammar (Goldberg 1999, Tomasello 2003), acquisition does not go far beyond learning form-meaning mappings; abstract construction schemas are admitted, but these generalizations are unsystematic and unconstrained. By contrast, the ReCoS-ian view recognises that form-meaning mappings do not explain syntax. Instead, formal features are involved in formal syntactic operations. Syntactic competence emerges from the interaction of the three factors of Chomsky (2005).

---

[3] Though see already Chomsky (2004:105-6).

The first factor is UG. The contribution of UG is minimal, but, in contrast to the eliminativism of standard emergentist theories, remains crucial. UG provides the formal syntactic operations Merge and Agree, and possibly Copy and Label,[4] along with an [attribute:value] template for formal features. However, UG does not provide the features themselves, which instead emerge from the interaction of factors two and three.

The second factor is the PLD. The PLD should properly be characterised as intake rather than input (Evers and van Kampen 2008; Lidz and Gagliardi 2015). Intake is the data that the child actually processes, as determined by factors such as attention, computational capacity, and the current state of the grammar. The PLD contains more than direct form-meaning mappings. Rather, it contains systematic departures from direct, item-for-item Saussurean arbitrariness, which reveal a more general, higher arbitrariness at the structural level (Biberauer 2014a, 2015a). Expanding from Zeijlstra's (2008) emphasis on doubling, these departures from maximally straightforward form-meaning mappings are taken to include agreement, (optional) silence, movement, and multifunctionality (Biberauer and Roberts 2014:7):

Table 1: Departures from item-for-item Saussurean arbitrariness signalled in the PLD, as summarised by Biberauer and Roberts (2014:7)

PLD characteristic	Form	Meaning
agreement	two	one
Case	yes	no
multifunctionality	one	several
empty categories	no	yes
movement	one	several positions

The central claim is that children are sensitive to these systematic departures from Saussurean arbitrariness. In particular, they respond by postulating formal features. The purpose of formal features is thus to regulate these systematic contrasts in the PLD. For example, negative concord (Biberauer and Zeijlstra 2012) signals the presence of a formal feature NEG, since it is an instance of agreement: two negative forms together convey a single negative meaning.

Such emergentism might appear to predict unconstrained variation. As Biberauer (2014b) notes, it is true that this approach predicts categorial particularism as opposed to categorial universalism (Haspelmath 2010). Since categories and formal features are constructed on the basis of language-specific PLD, with varying morphosyntactic and morphosemantic contrasts, they will have language-specific formal identities. Evidence in support of this prediction abounds (Wiltschko 2014; cf. §4.3).

---

[4] It is in this sense that syntax is universally the same; cf. Boeckx (2014).

However, variation is constrained by the third factor. As we saw above, Chomsky (2005:6) deconstructs the non-language-specific third factor into two subparts; in short, subpart (a) refers to acquisition strategies, while subpart (b) amounts to computational economy. The ReCoS approach builds on both subparts in proposing an economy-driven acquisition strategy that concretises Chomsky's (2005) third factor proposal. The overall strategy consists of two principles, Feature Economy (Roberts and Roussou 2003:201) and Input Generalization (Roberts 2007:275):

(2) Feature Economy (FE):

Given two structural representations R and R' for a substring of input text S, R is less marked than R' iff R contains fewer formal features than R'.

(3) Input Generalisation (IG):[5]

If a functional head F sets parameter P_j to value v_i then there is a preference for similar functional heads to set P_j to value v_i.

On their original conception, FE and IG were viewed separately as direct third factor biases (Roberts 2012). However, they can be fruitfully viewed in combination, yielding a minimax search/optimization algorithm (Biberauer 2013a; Biberauer and Roberts 2014) that instructs the acquirer to make maximal use of minimal means. To draw out this minimax character, the two principles could be informally stated as:

(4) FE:

Minimize the postulation of features.

(5) IG:

Maximise postulated features.

This suggests that FE and IG are reflexes of a superordinate third factor principle, Minimax (cf. Biberauer 2014a; Biberauer and Roberts 2015, to appear a):

(6) F3 Minimax:

Make maximal use of minimal means.

We will see evidence for the operation of this minimax algorithm in the acquisition of syntax throughout this thesis. Further circumstantial evidence for a minimax principle can be found elsewhere in linguistics, and beyond language in general cognition.

---

[5] Perhaps more properly In*take* Generalization – see above discussion.

2.7   Minimax beyond syntax

The workings of an overarching psychological principle Minimax can be found beyond the acquisition of syntax in linguistics, at the linguistics-cognition interface, and in domain-general cognition.

Within linguistics, Minimax influences the acquisition of phonology and semantics. In phonology, Dresher's (1998, 2003, 2008, 2009) Successive Division Algorithm proceeds according to a Minimax scheme (Biberauer 2014c; Biberauer and Roberts 2014). From a starting point of assuming no feature contrasts (cf. FE), the sound space is successively divided up into phonemes, where the data demands so (cf. IG). Thus sound systems make maximally efficient use of featural contrasts (Clements 2003, 2009).[6]

In semantics, Minimax is arguably at play in concept construction. The traditional view of semantic concepts parallels the traditional view of syntactic features, arguing that they are innate.[7] Innateness applies to a core set of primitives, such as the primes of Wierzbicka's (1996) Natural Semantic Metalanguage, or, in the extreme, to all concepts (Fodor 1975). By contrast, on the constructivist view (e.g. Mandler 1992) children successively divide the world into concepts along Minimax lines. In this regard, Biberauer and Roberts (2014) point to the Minimax nature of Jaspers' (2005, 2012, 2013) Concept Formation Constraint, as applied to the constructive acquisition of logical connectives, quantifiers, and colour terms.

At the linguistics-cognition interface, Minimax is standardly assumed to be at play in pragmatics. According to Relevance Theory (Sperber and Wilson 1986, 1995), for example,[8] pragmatic reasoning is characterised by a single, double-bounded Minimax principle of relevance, enjoining minimal processing effort for maximal cognitive effect. Hence any departure from minimal form or literal meaning carries additional meaning as a conversational implicature, increasing the cognitive effect in line with the greater processing effort.

Minimax computation is also detectable beyond linguistics in domain-general learning. For example, Mobbs (2015:202) reports findings of non-veridical learning and overregularization (cf. IG) from the general psychology literature (e.g. Gardner 1957, Weir 1964, Bever 1982).

Overall, there is strong evidence for a Minimax computational principle in operation throughout linguistics and general cognition. It is therefore sensible to pursue the hypothesis that it plays a

---

[6] See Mobbs (2015:196ff.) for discussion.
[7] Compare the discussion of Zeijlstra (2008) in §2.4.3.
[8] The intuition goes back to Grice's (1989) Quantity Maxim: be as informative as required, but no more.

significant role in the acquisition of syntax. In syntax, the predicted patterns of acquisition have been modelled as parameter hierarchies.

2.8   Parameter hierarchies

The minimax algorithm predicts NO > ALL > SOME acquisition sequences as the child postulates the minimum number of formal features consistent with the input. Without evidence to the contrary, no feature is assumed to exist by FE. When evidence for the existence of a feature is encountered, it is postulated in violation of FE, and generalised as far as possible by IG. The domain of application of the feature is subsequently restricted to some sub-parts of the original domain, according to the positive evidence present in the PLD, violating IG as minimally as possible. The crucial role of concrete positive evidence means that the Linking Problem does not arise here. By FE, features are created by the acquirer only when concrete evidence demands so.

The NO > ALL > SOME acquisition sequences have been modelled as emergent, lexicocentric parameter hierarchies. One proposed hierarchy relates to headedness (Roberts 2012:321):

Is head-final present?

No: **head-initial**   Yes: present on all heads?

Yes: **head-final**   No: present on [+V] heads?

Yes: **head-final**   No: present on ...
in the clause only
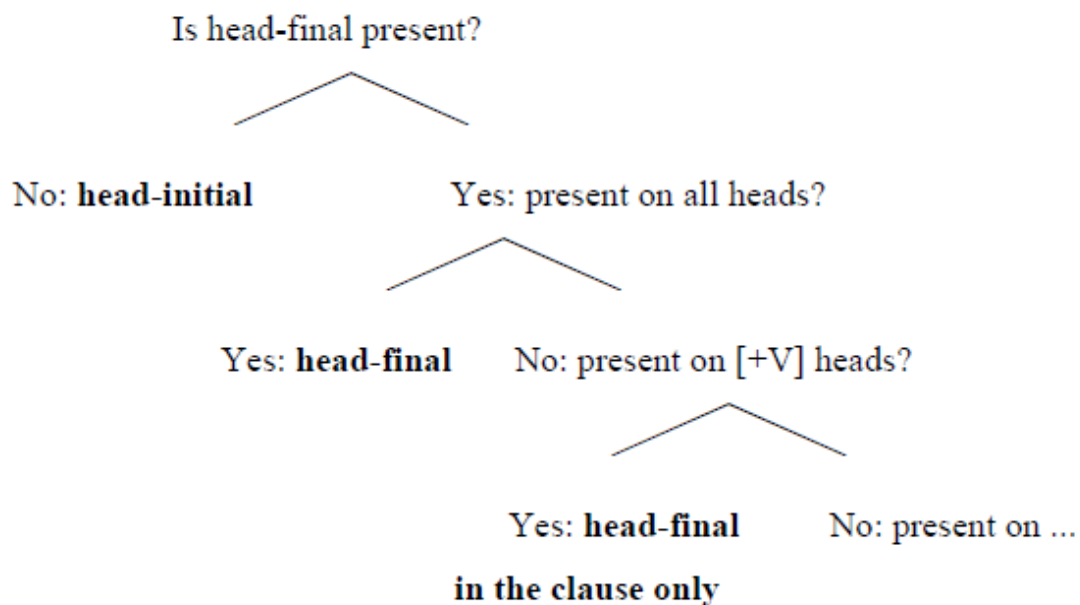
Figure 2: An emergent parameter hierarchy for headedness, from Biberauer et al. (2014:110)

Different depths in the hierarchy can de distinguished descriptively (Biberauer and Roberts 2012:268), according to whether a parametrically variant value of a feature applies to: all relevant functional heads – macroparameters; a naturally definable class of functional heads – mesoparameters; a small subclass of functional heads – microparameters; or a few individual lexical items – nanoparameters.

The parameter hierarchies are claimed to embody predictions for diachronic change, acquisition, and typology (Roberts 2012). However, in trying to pack all three into a single representation, the hierarchies fail to accurately capture the characteristics of any of the three. I will outline these problems, and sketch alternative possible representations, with each of the three areas requiring a different perspective.

2.8.1 Diachrony

The algorithm predicts two paths for diachronic change, depending on the robustness of a feature in the PLD (Biberauer 2015a:38)[9]. A robust feature is predicted to be able to spread to other domains by IG. Conversely, a feature that is not robustly attested is predicted to become restricted to ever smaller domains, and eventually to disappear, by FE. This second pathway of change would lead to upward leaps in parameter hierarchies (Biberauer 2015a:39). Biberauer and Roberts (to appear b) explore this second pathway with reference to conditional inversion in the history of English. English once had a mesoparametric setting for V2, which applied to all inflected verbs, attracting them to C. Conditional inversion survived the loss of general V2, as V2 became a microparametric setting, restricted to auxiliaries. Finally, conditional inversion became nanoparametric, restricted to just one modal (7a), and certain forms of have (7b) and be (7c):

(7) a) Should Cameron win, Ian will be fuming.

b) Had Miliband won, Ian would have been content.

c) ?Were Farage to win, Ian would be stunned.

From the mesoparameter stage to the microparameter stage, English C to I movement became more marked, applying to a more specific group of verbs further down the word order hierarchy. From the microparameter stage to the nanoparameter stage, a radical simplification has taken place, leaping up the hierarchy and leaving behind a few nanoparametric lexical exceptions.[10]

The current presentational format of the parameter hierarchies captures this change from an acquisition perspective. Rather than traversing all the way down the hierarchy, the acquirer remains at the top, and learns nanoparametric lexical items as exceptions:

---

[9] See also Biberauer and Roberts (2009), though not yet in these terms.
[10] Compare Yang (2013) on tipping points to a rule.
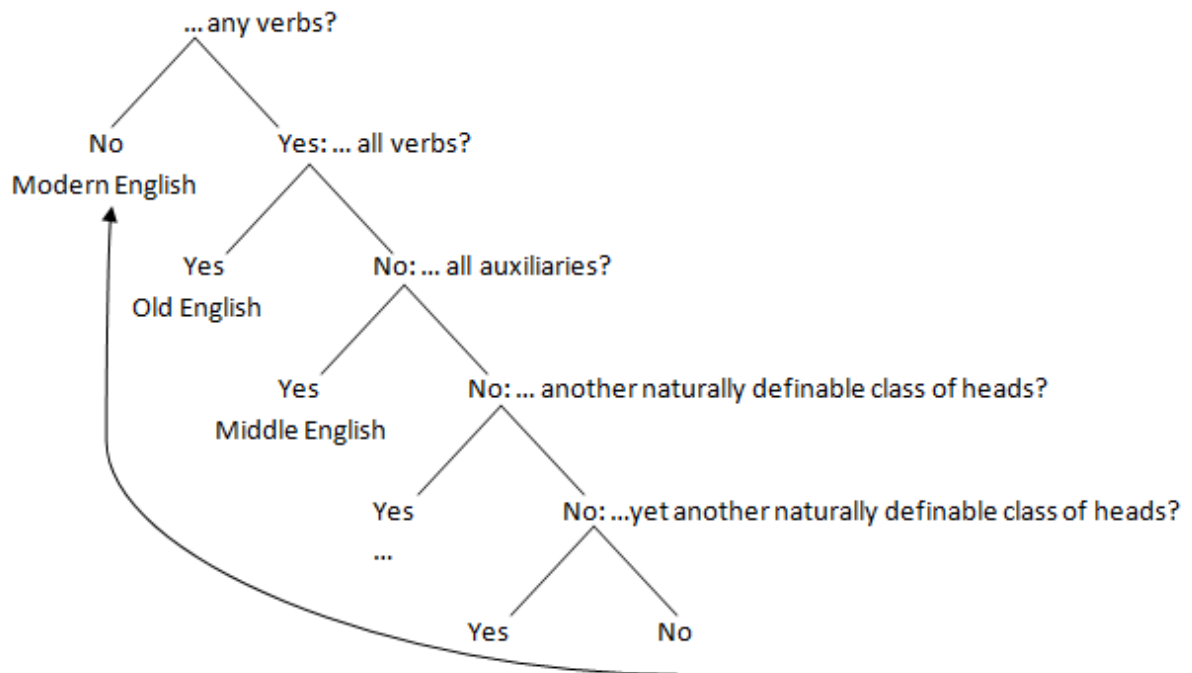
Figure 3: An upward leap in the headedness hierarchy in the history of English from an acquisition perspective; one generation of acquirers traverses all the way down the hierarchy, but the next generation remains at the top

Figure 3 reflects the fact that this is a big change in the system in acquisition terms. However, the distance between the position leapt from and the position leapt to does not reflect their closeness as just a single step in diachronic terms. This closeness is brought out in the alternative representation offered below:

Table 2: A tabular representation of the diachronic change in headedness in the history of English

Conditional C attracts:	Verbs		Stage
all verbs	███████████████████		Old English
all auxiliaries	████████████████		Middle English
another naturally definable class of heads	██████████		…
yet another naturally definable class of heads	███████		…
no verbs[11]			Modern English

2.8.2   Acquisition

Central to our focus here, the presentational format of parameter hierarchies does not adequately capture the minimax algorithm's predictions for acquisition. In particular, the question format is misleading, as are binary branching SOME options.

---

[11] Except nanoparametric lexical items.

First, the question format is not appropriate from an acquisition perspective. Looking first to the top of the hierarchy, the child does not interrogate the data to determine whether or not a feature is present; if the feature is absent, no question arises. Take the headedness hierarchy in Figure 2 as an example: in a consistently head-initial language, the question of whether a head-finality feature is present does not arise. This has been increasingly recognised in ReCoS research by Biberauer (2013b), Biberauer et al. (2014:124), and Biberauer and Roberts (2014:10, 2015:10). Rather, as Biberauer and Roberts (2014:11) note, the fact that the first question does not arise means that "a natural notion of default (or 'unmarked parameter setting') emerges without stipulation". This notion properly derives from FE, whereby no feature is posited without evidence. The question format of the parameter hierarchies does not accurately reflect how the child approaches the data, nor does it bring out the notion of defaults.

Second, the inappropriateness of the parameter hierarchy format persists further down the tree in relation to SOME options. As Biberauer (2015a:37, 2015b:14) emphasises, the SOME options are not successively considered by the acquirer. Rather, the acquirer will only postulate the SOME option that is motivated by the PLD of the language they are acquiring. For example, in Figure 2 the move to questioning whether the head-final feature is present on [+V] categories, and subsequent questions, involve SOME options of this type.

Sheehan (to appear:39ff.) identifies these relations as negative dependencies. The dependencies in a parameter hierarchy are negative when it is the NO side that branches. The properties involved are not cumulative, but mutually exclusive: the child can restrict the property in question to whatever coherent sub-class is dictated by the PLD. Consider for example, Sheehan's (to appear:38f.) parameter hierarchy for ergativity, where the dependencies from P2-P5 are negative:[12]

---

[12] Sheehan argues that only P2-4 are negatively dependent; but P5 can be reordered too, as long as P6, which is dependent upon it, moves also.

P1: Does transitive v assign a theta-related case (ERG) in a language L?

N - accusative   Y – non-accusative
P2: is this *generalised* to all vs in L?

Y – Basque   N – P3: is this *extended* to intransitive [volition] vs in L?

Y – Hindi   N – P4: is this *restricted* to [instigator] vs in L?

Y – Tsez, Lezgian   N – P5: does v^{erg} bear an EPP in L?

N – Walpiri, Niuean(?)   Y – P6: are v^{erg}'s φ-features suppressed in L?

N – West Greenlandic, Tagalog, Cavineña(?) Yidiɲ, Kuikuro, Tongan   Y – Chamorro, Mam, Q'anjobal, Dyirbal, Seediq, Trumai

Figure 4: Sheehan's (to appear:38f.) parameter hierarchy for ergativity

For Sheehan, negative dependencies are a problem.[13] They are not restrictive enough, in that they allow successive levels of the hierarchy not to stand in subset-superset relations. By contrast, it is only when the YES side of the hierarchy branches that there is a truly dependent, hierarchical relationship between the two levels. Sheehan argues that parameter hierarchies should be arranged to maximise branching YES options, as each successive layer of the hierarchy moves from a subset to a superset. Then, if the hierarchy is assumed to model an acquisition path, acquisition becomes the process of selecting grammars of increasing size and complexity, where this is apparently measured in terms of description length. Sheehan claims that organising the parameter hierarchies in this way might provide a solution to the subset problem as identified by Wexler and Manzini (1987) and Manzini and Wexler (1987). However, it is not clear that subset-superset relations are an issue here, due to the way the ReCoS acquisition algorithm interacts with the PLD. To demonstrate this requires an excursus on the Subset Principle.

2.8.2.1 Is there a subset problem?

Sheehan's (to appear) concerns relate to the Subset Principle (SP) (Berwick 1985).[14] The SP can be informally defined as the principle that "the learner must guess the smallest possible language compatible with the input at each stage of the learning procedure" (Clark and Roberts 1993:304-5).

---

[13] Biberauer and Roberts (2015:6) though not in the same terms, likewise argue that parameter hierarchies must not branch from the NO side, but only from the YES side of binary branching parameters, in order to be constrained, deterministic representations of the learning path.
[14] Though cf. Dell (1981) for phonology.

The SP is often invoked because there is no negative evidence routinely available to acquirers (Baker 1979, Baker and McCarthy 1981); and even where it is available, it is generally ignored (Brown and Hanlon 1970). Thus, if an acquirer guesses a grammar that is a superset of the target grammar, they will not receive the corrective negative evidence that urges them to retreat to the subset target grammar. As such, an acquirer must conform to the SP, or risk entering a superset trap. The SP has been discussed at length in the literature (e.g. Biberauer and Roberts to appear a), and leads to an evolutionarily implausible conception of UG, along the same lines as encountered in §2.1. However, SP does not turn out to be an issue for our model.

For Hale and Reiss (2003:242), the SP makes "a logical necessity" of the Innateness of Primitives Principle. They attribute this principle to Pylyshyn (1973), its stout defence to Fodor (e.g. 1975), and its succinct formulation to Jackendoff (1990:40): "In any computational theory, 'learning' can consist only of creating novel combinations of primitives already innately available." Since learning consists in parsing inputs into representations, it cannot begin without representational primitives. Learning is therefore characterised as the relaxing of highly specified early representations, admitting successive supersets of entities. This is the reverse of the ReCoS position, entailing an 'overspecification' view of UG, where the set of UG primitives determines the set of possible grammars. SP compliance therefore comes at the high cost of an evolutionarily implausible view of UG.

Fodor and Sakas (2005) take their lead from Hale and Reiss (2003) in identifying further innately given information required for a learner to observe the SP. This is because the SP is beset by an implementation problem (cf. Biberauer and Roberts 2009:59,fn.2). If they are to avoid superset traps, the learner must know the subset-superset relations between parameter settings. To do so means writing default parameter values into UG, as well as a fixed order for their consideration (cf. §2.1), all in addition the representational primitives argued for by Hale and Reiss (2003). In our terms, this would mean that both parameter hierarchies and features must be innate, making the proposed conception of UG further evolutionarily implausible.

However, the SP and the innateness and learnability issues it raises can be circumvented. Hale and Reiss (2003:226) identify two valid critiques of a claim that Q is necessary for acquisition, and that Q is therefore innate. One way would be to show that Q is not necessary for acquisition. The second is to show that Q derives from more basic innate entities or processes. The second line of criticism is the one generally taken by constructivists, with their appeals to general learning mechanisms. Here, however, we have the ReCoS minimax algorithm as a concrete proposal for a domain-general learning mechanism.

The interaction of the ReCoS algorithm with the data allows apparent superset errors to be cured, as positive evidence forces a retreat from full IG compliance (Biberauer and Roberts to appear a; Branigan 2012; contra Fodor and Sakas 2005:517). Against the concerns of the SP, the ReCoS approach predicts that acquisition proceeds top-down from supersets (ALL) to subsets (SOME), rather than bottom-up from subsets to supersets. As Bazalgette (2015:§1.3.3) notes,[15] however, this aligns with the fact that language data stand in intersection relations, rather than inclusion subset-superset relations. In light of the headedness hierarchy in Figure 2, consider the intersecting data generated by a grammar where all heads are head-final (G_m) at a level higher in the hierarchy than a grammar where only [+V] heads are head-final (G_n):

L_m

I literature of students met

I students met

I students of literature met

L_n
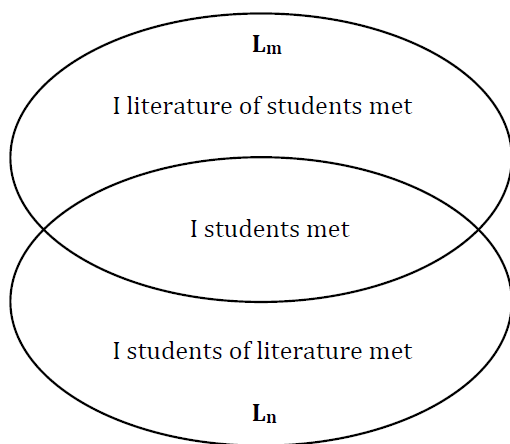
Figure 5: Diagram from Bazalgette (2015:§1.3.3) of the intersecting data of two languages L_m and L_n, generated respectively by a grammar G_m where all heads are head-final, and a grammar G_n where only [+V] heads are head-final

Therefore, adequate positive evidence exists in the form of intersecting data to allow a minimax acquisition algorithm to circumvent the subset problem.[16]

Overall, by postulating minimal innate content, our theory circumvents the conceptual challenge of the subset problem, just as it did the linking problem. Consequently, we do not need to tether ourselves to an evolutionarily implausible view of UG, nor – from an acquisition perspective – do we need to be concerned about negative dependency relations in the parameter hierarchies.

2.8.2.2   The format of acquisition

However, we are still left with the problems of the question format of the parameter hierarchies, and the fact that not all SOME options are successively considered: options only arise if they are

---

[15] Biberauer and Roberts (2009:59) make observations in the same spirit for the macro-settings of the headedness and null subject parameters; as did Hyams (1986:154-6) for the latter.
[16] Though subset-superset relations may have a role to play in diachronic changes that eliminate free variation from the syntax (Biberauer and Roberts 2009).

relevant in the language being acquired. In fact, from the perspective of the acquirer, there are no options: they go wherever the interaction of the data and Minimax leads them, until there are no further contrasts to be captured. Rather than the hierarchy in Figure 2, perhaps Table 3 would therefore be a more accurate depiction of the acquisition of head-finality from the acquirer's perspective:

Table 3: A tabular representation of the acquisition of head-finality

	Consistent head-initial	Consistent head-final	Head final in the clause [+V] only	Head final within vP [+v] only
NONE	●	● ↓	● ↓	● ↓
ALL		●	● ↓	● ↓
SOME			●	● ↓
SOME				●

2.8.3   Typology

The original typological significance of the parameter hierarchies was in capturing typological skewing (Greenberg 1963). Roberts (2012) argues, contrary to Baker (2008), that macroparameters can be viewed as aggregates of microparameters acting in concert due to IG. Given the above discussion of parameter hierarchies from an acquisition perspective, we can make a refinement that allows us to capture the notion of a typological equivalence class (Biberauer 2015a:37, 2015b:14). Just as an acquirer only considers the successive, negatively dependent SOME options that are relevant in their PLD, language groupings cannot be defined by a successive set of terminal nodes. Instead, the successive SOME questions are freely reorderable, and define mutually exclusive options. Sheehan (to appear:40), on the suggestion of an anonymous reviewer, presents an alternative version of her ergativity hierarchy, in which only one negative dependency remains (that between P3' and P5'):

P1: Does transitive v assign a theta-related case (ERG) in a language L?

N - accusative   Y – non-accusative
                        P4': is this *generalised* to all transitive vs in L?

N (some)– Tsez, Lezgian   Y – P3': is this *extended* to intransitive vs in L?

N– P5': does v^{erg} bear an EPP in L?   Y –P2': is this *generalised* to all intransitive vs?

N – Walpiri, Niuean(?) Y –   N (some)– Hindi   Y - Basque
                        P6': are v^{erg}'s φ-features suppressed?

        N – West Greenlandic,   Y – Chamorro, Mam,
        Tagalog, Cavineña(?)   Q'anjobal, Dyirbal,
        Yidiɲ, Kuikuro, Tongan   Seediq, Trumai

Figure 6: Sheehan's (to appear:40) presentation of an anonymous reviewer's alternative to Figure 4, which minimises negative dependencies

Instead, the mutual exclusivity of negatively dependent typological options would be better represented by admitting n-ary branching representations. Then every group of negatively dependent SOME options defines a different language group, but all at the same level of the typology. The representation below thus draws out the notion of a typological equivalence class (Biberauer 2015a:37, 2015b:14):

P1'': transitive v assign a theta-related case (ERG) in a language L

N - accusative   Y– non-accusative

P2'': generalised to   P3'': extended to intransitive   P4'': restricted to   P5'': v^{erg} bears EPP
all vs   [volition] vs   [instigator] vs

Basque   Hindi   Tsez, Lezgian   N – Walpiri,   Y – P6'': v^{erg}'s φ-features
                                                                                 Niuean   suppressed

                                                                                 N – West Greenlandic,   Y – Chamorro,
                                                                                 Tagalog, etc.   Mam, etc.
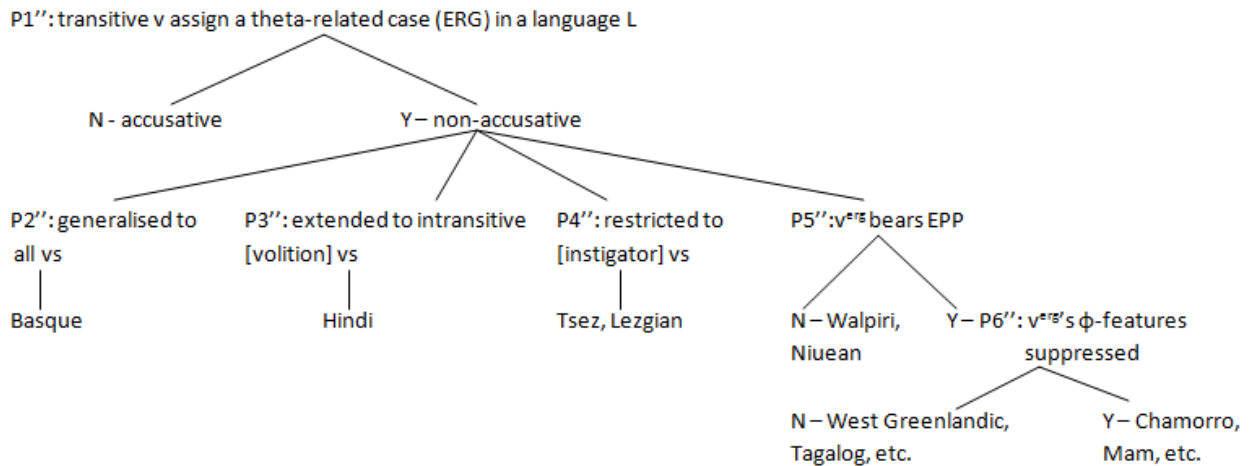
Figure 7: A typological parameter hierarchy of ergativity, alternative to Figures 4 and 6

2.8.4   Summary of parameter hierarchies

Overall, attempting to capture diachrony, acquisition, and typology together in a single parameter hierarchy format results in a misleading characterisation of all of three. This finding suggests it would be unwise to draw conclusions from literal readings of the parameter hierarchies, as do Biberauer et al. (2014) for calculating the relative complexity of languages. Instead, each domain

should be represented in its own way, which more accurately captures the differing perspectives involved.

2.9   Summary of emergent syntax

This section has shown that it is possible to pursue a generative theory of emergent syntax that is couched in a feature-based, minimalist, three factors framework. I will seek to develop this theory, in particular in the direction of emergent syntactic structure, with insights from imperatives in the rest of this thesis.

3   Why imperatives?

Before considering previous formal analyses of imperatives, and ultimately proposing my own, I should justify my choice of imperatives to develop a theory of emergent syntax. Biberauer and Roberts (2015:7) "take the acquirer to be sensitive to particular aspects of PLD such as movement, agreement, etc., readily encountered in simple declaratives, questions and imperatives." Children apply Minimax to these readily available aspects of the PLD to construct the formal features of their language. In this thesis I examine imperatives. I justify this choice from: (i) the relatively high frequency of imperatives in the PLD, providing contrast with other clause-types; (ii) cases where imperatives are disproportionately useful for the acquirer; and (iii) evidence that imperatives are analysed and manipulated as a very early component of intake. I end this subsection by urging methodological caution against over- or under-estimating the significance of imperatives in acquisition.

3.1   Relative frequency and contrast

Imperatives are fairly common in child-directed speech (CDS). Salustri and Hyams (2006:164,166) report that 14.9% of verbs in their sample of child-directed Italian were imperatives, and 36% in German. For child-directed English, Newport, Gleitman and Gleitman (1977:122) report 18% imperatives. The significance of the frequency of imperatives in CDS is highlighted when compared with their frequency in (adult-to-)adult-directed speech (ADS). Salustri and Hyams (2006:164) report 5.6% imperatives in their sample of adult-directed Italian. Newport, Gleitman and Gleitman (1977:122) describe the proportion in adult-directed English as "negligible", though suggestive is their finding that English CDS contains 30% declaratives, as compared with 87% declaratives in English ADS. These percentages counter the intuitively plausible view that the input data might be biased to allow children to start from declaratives. Instead, the input data provide children with a

range of contrasting clause-types. The hypothesis pursued here is that the various clause-types, both individually and in contrast with each other, provide the child with rich acquisition evidence.

3.2 Useful imperatives

In addition to being relatively frequent in CDS compared to ADS, imperatives can provide acquirers with useful evidence for formal features that are under-represented elsewhere in the language. This consideration applies to two of the phi-feature categories: gender and number.

3.2.1 Gender

Imperative data can offer a 'way-in' for acquiring gender. For example, as Aikhenvald (2010:123f.) reports, gender marking is generally not very prominent in Lakhota (Siouan). However, gender is marked on imperative particles, which differ by the sex of the speaker. Similarly, gender is distinguished in special imperative-only lexemes like 'come here!' and 'look out!' in Zargulla (Omotic, Afroasiatic) (Aikhenvald 2010:318). Thus imperatives can be crucial in providing acquisition evidence for gender features, in a language where they are otherwise rare.

3.2.2 Number

Another phi-feature for which imperatives can offer a 'way-in' is number. Some languages have more number distinctions in imperatives than in their other verbal paradigms. Birjulin and Xrakovskij (2001:29) list Nivkh (isolate), Klamath (Plateau Penutian), Mongolian (Mongolic), Japanese, and Lezgi (Northeast Caucasian) as languages of this type. Thus imperatives can be the main source of evidence for acquiring number features in a language.

3.3 Imperative intake

Despite their relatively high frequency, and the useful featural information they often contain, it might be objected that the mere presence of imperatives in the PLD does not guarantee that acquirers process them as early intake. To counter this objection, I present production evidence for the early acquisition and analysis of imperatives, and their resulting diachronic stability.

3.3.1 Early production and analysis

It is resoundingly reported that children acquire imperatives very early. Imperatives are very frequent in early production in a wide range of languages, including Georgian (Kartvelian) (Imedadze and Tuite 1992:59), Modern Hebrew (Semitic) (Berman 1985:268), Kalui (Kutubuan) (Schieffelin

1985:536), Warlipi (Pama-Nyungan) (Bavin 1992:327), and West Greenlandic (Eskimo-Aleut) (Fortescue and Olsen 1992:153).

It could still be objected that such early and frequent production of imperatives is not significant for grammatical acquisition, as it represents only unanalysed imitation of the input. At this point, the distinction between input and intake (Evers and van Kampen 2008; Lidz and Gagliardi 2015) becomes crucial. Taking Kalui (Schieffelin 1985) as a test case, it might at first seem that imperatives are unanalysed imitations of forms that are frequent in the input. The simple imperative discourse markers *mena* 'Come!' and *bɔba* 'Look!' are among the first verbs to be produced (Schieffelin 1985:536). However, evidence that acquirers intake the imperative input comes from their use of the imperative as a base in attempting to derive other verb forms (Schieffelin 1985:569); for example, a child will utter (8) with the intended meaning of adult (9):

(8) ne menɛ
     I come:IMP

(9) ne yɔl
     I come:1:PRES
     'I am coming.'

Another piece of evidence that children analyse their imperative input comes from patterns of overgeneralisation in Japanese. Clancy (1985:383ff.) observes that Japanese CDS contains a high proportion of explicit directive imperatives, which are marked by *–te*. This *–te* morpheme is overgeneralised beyond the imperative clause-type as a finite tense/aspect marker in early child production. This pattern is an example of the ReCoS algorithm in operation: a single, early-acquired feature (FE) is generalised to further contexts (IG) in preference to postulating another feature, before evidence is found to the contrary. That the direction of generalisation is from imperatives to other clauses suggests the significance of imperatives for acquisition in Japanese.

Overall, these two case studies of the active manipulation of imperatives by acquirers suggest that imperatives are part of the early intake, not just imitations of the input.

3.3.2 Diachronic stability

Further evidence for the significance of imperatives in acquisition comes from their diachronic stability. Aikhenvald (2010:339-41) discusses the stability of canonical second person singular imperatives, which are often one of the most archaic forms in a language. For example, canonical imperatives in the Arawak languages of South America have not changed from the proto-language.

This stability indicates reliable transmission of imperatives over successive generations, which adds to the evidence that imperatives form part of very early acquisition intake.

3.4   Methodological caution

Having made the case for the significance of imperatives in emergent acquisition, I sound two notes of caution, without which their importance could be over- or under-estimated.

First, imperatives are neither sufficient nor necessary for acquisition. They are not sufficient, because much about the structure of a language cannot be gleaned from imperatives, but can only be found in other clause-types. Moreover, evidence from imperatives cannot be necessary for acquisition, because some languages do not have imperatives; Aikhenvald (2010:398) cites Navajo (Athapascan) and Wardaman (Australian), for example. Therefore, the significance of imperatives for acquisition should not be over-estimated.

Second, it should be recognised that observation has the potential to impact upon language use in fieldwork settings. This is known in science as the 'observer effect', or, pertaining specifically to human subjects modifying their behaviour, the 'Hawthorne effect' (Gillespie 1991). Observation could lead to subjects using more polite, indirect, non-imperative command strategies, reducing the occurrence of grammatical imperatives. In this light, consider Aikhenvald's (2010:304) report that Nuer (Nilotic) has a specialised imperative paradigm (Crazzolàra 1933:140) that is never used (Akalu 1985:63-4); but if the language has an imperative paradigm, it must be used in order for children to acquire it! Imperatives may therefore be under-represented in observational fieldwork, which should not lead to the significance of imperatives for acquisition being under-estimated.

3.5   Summary of why imperatives

Overall, the importance of imperatives for a theory of emergent syntax is apparent from their relative frequency, in contrast with other clause-types, and the useful featural information they can provide. Moreover, children clearly make use of imperative input data, analysing it as part of their intake. Next I will consider previous formal analyses of imperatives, assessing what they can offer to a theory of imperatives from a neo-emergentist perspective.

4   Formal analyses of imperatives

In this section I consider previous formal analyses of imperatives. In each case, I outline the analysis before evaluating which aspects to retain, and which to discard. This review of the literature will

pave the way for a novel analysis of imperatives from a neo-emergentist perspective in the next main section.

4.1   Formalising the imperative

First and foremost, the imperative should be formalised as a clause-type. Aikhenvald (2010) includes non-imperative command strategies in her typological survey, but takes pains to recognise from the outset that a formal characterisation of 'imperative' is necessary. Aikhenvald (2010:2) distinguishes between imperatives as a category of language and commands as a parameter in the real world, on analogy with the distinction between tense and time. I follow Lyons (1977) in eschewing meaning-based distinctions, and focusing on imperatives as a category of language. More specifically, I follow Sadock and Zwicky (1985) in analysing the imperative as a basic clause-type, opposed to declaratives and interrogatives.

Still, the relevance of directive speech act content has not gone unnoticed in formal work. This was originally represented by Katz and Postal (1964:74-9) as an IMP(erative) morpheme, and in subsequent generative syntactic work has been standardly represented as an [imp(erative)] feature in the C-domain, which types the clause (see van der Wurff 2007:21 for references). The presence of [imp] in C is taken to account for the near universal lack of focus markers or focus constructions in imperatives (Aikhenvald 2010:108), which would compete for the same structural position. It is also taken to account for the root clause restriction on true imperatives, which cannot be embedded (van der Wurff 2007:22-7; contra Aikhenvald 2010:109ff.).[17]

4.1.1   Evaluation

I adopt the idea that the directive semantic content of imperatives is syntactically encoded by an [imp] feature under C.

4.2   Alcázar and Saltarelli (2014)

The first recent analysis of imperatives we will consider is Alcázar and Saltarelli (2014), who propose an updated 'light' version of the Performative Hypothesis (PH) of Ross (1970).

Ross' PH was conceptually rooted in the philosophy of language tradition (see van der Wurff 2007:4 for references). PH claims that root clauses are in fact complement clauses of higher covert lexical

---

[17] Imperatives can be embedded in, for example, Korean (Zanuttini, Pak and Portner 2012), under particular structural circumstances. I argue in §§5.5 and 6.3 that these imperatives are not clausal, so there is no [imp] in C to block embedding.

predicates: *I say that* for declaratives, *I ask that* for questions, and *I order that* for imperatives. However, as van der Wurff (2007:7f.) summarises, serious questions remained about the operation deleting the performative hypersentence, and how to capture the peculiar word order, negation, and subject optionality of imperatives. Further work made little to no headway on these issues, and the PH fell into disrepute.

However, it has become increasingly recognised that syntax is sensitive to the context. Alcázar and Saltarelli (2014:75-96,§3.3) cite indexical shift, logophoric pronouns, and conjunct-disjunct marking as syntactic phenomena that can only be accounted for with reference to the context. Similar conclusions have been reached by, among others, Speas and Tenny (2003), Bianchi (2003), Sigurðsson (2004), Giorgi (2010), Hill (2013), Haegeman and Hill (2013), Haegeman (2014), Wiltschko (2014) and Heim et al. (2014), Wiltschko (2015). These works variously encode the context of utterance (Kaplan 1989) in the syntax in terms of speaker, addressee, time, place, and world. This syntacticization of the context and properties of speech acts revives the PH in a different guise. Embedding under a lexical verbal predicate is replaced by embedding the overt sentence structure under functional projections, comprising a further extended projection of the clause.

In this vein, Alcázar and Saltarelli (2014) advance the 'light' performative hypothesis (LPH) as an analysis of imperatives. According to LPH, the syntax-semantics relationship in imperatives is mediated by the two clausal phase heads, v and C. Imperative/illocutionary force [IF] in C licences a directive functional light verb as the lower v phase head. This functional light verb has the meaning of prescription (Birjulin and Xrakovskij 2001). The relationship CP(vP) is a function from context to content, creating a predicate argument relation between syntax and the context of utterance. The thematic role of addressee is assigned to the subject argument of the vP that introduces the lexical verb V. A deontic relation, or δ-relation, holds over the speaker and addressee speech act participants, constraining them to be first and second person respectively. In canonical imperatives, the addressee is also the performer. Overall, this yields the following structure:
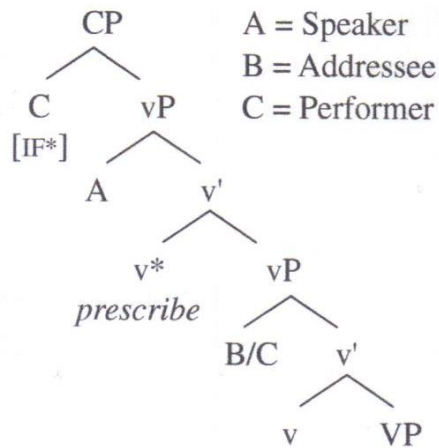
A = Speaker
B = Addressee
C = Performer

As well as capturing canonical imperatives, LPH subsumes hortatives as a type of imperative, in opposition to exceptionalist analyses (for references see Alcázar and Saltarelli 2014:124). First person imperatives are often termed (ex)hortative, while third person imperatives are termed jussives or injunctives (see Aikhenvald 2010:48 for references). I will follow Alcázar and Saltarelli (2014) in using 'hortatives' to refer collectively to first and third person imperatives. Alcázar and Saltarelli (2014) argue that hortatives are of the same imperative clause-type as canonical second person imperatives. In doing so, they follow the inclusion of hortatives in typological surveys of imperatives in Xrakovskij (2001), by Aikhenvald (2010) and for the World Atlas of Language Structures (WALS): van der Auwera and Lejeune (2013a) and van der Auwera, Dobrushina and Goussev (2013). Support for this position comes from morphological, typological and semantic evidence.

First, a language can happily have a morphological imperative paradigm which encompasses all three persons. Sanskrit is one such language (Aikhenvald 2010:48, and pp.49-52 for further examples). Kobon likewise has a homogenous imperative paradigm (Alcázar and Saltarelli 2014:36). Second, the above typological surveys uncover two implicational relationships between hortatives and imperatives (Alcázar and Saltarelli 2014:40ff.): (i) if a language marks a person distinction in the hortative, then it also does so in the imperative; and (ii) second person is unmarked in imperative-hortative paradigms, whereas third person is unmarked in declaratives and interrogatives. Third, semantically the addressee mediates between the speaker and a third party in hortatives (Alcázar and Saltarelli 2014:6). In other words, the addressee is still the subject. Birjulin and Xrakovskij (2001:5f.) likewise split the addressee and performer, and thereby admit hortatives into imperative paradigms. As van der Auwera, Dobrushina and Goussev (2013) explain, both imperatives and hortatives express a wish of the speaker about a future state of affairs, and both appeal to the

addressee to help make it come true. The difference lies in who is in control: the addressee in imperatives, and some other person in hortatives.

Overall, the morphological, typological and semantic evidence for integrating hortatives with imperatives leads to the following structure for hortatives as 'causative imperatives', with an additional causative little v projection:

```
        CP          A = Speaker
       /  \         B = Addressee
      C    vP       C = Performer
     [IF*]  /\
          A    v'
              /  \
            v*    vP
         prescribe /\
                 B    v'
                     /  \
                    v    vP
                  cause  /\
                       C    v'
                           /  \
                          v    VP
```
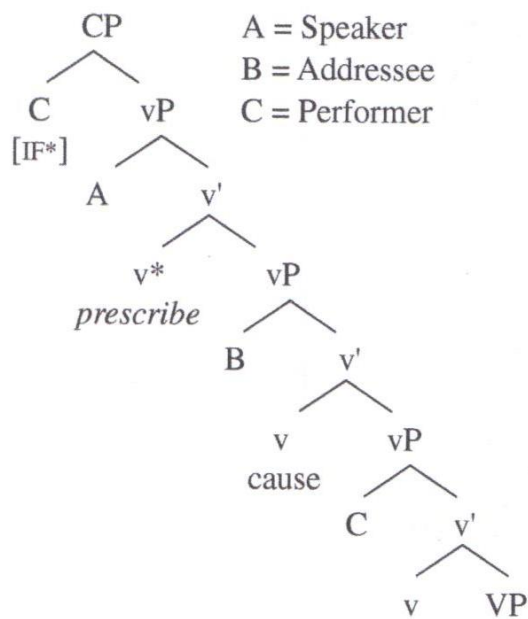
Figure 9: The structure of a hortative as a causative imperative, according to Alcázar and Saltarelli (2014:114)

4.2.1   Evaluation

I adopt the idea that syntax must make reference to some aspects of the context; in particular, I argue for the central role of the addressee's point-of-view in imperatives. I also argue that hortatives can be captured within the standard imperative structure, though my analysis will not require an extra causative position. Moreover, care must be taken to distinguish true *let*-imperatives from optatives.

I abandon the phasal CP(vP) relation, maintaining the traditional position that directive force comes from [imp] in C, which makes the presence of a 'prescription' verb tautologous. I abandon the δ-relation for similar reasons: the speech act role of addressee can be established by a direct relation between syntax and the context, without a deontic binding relation involving the speaker. In fact, unlike the addressee, I will not represent the speaker within the clausal domain. Alcázar and Saltarelli (2014:137f.) ponder a question: "what independent evidence is there for an additional vP to introduce a "logical" subject (Speaker or argument A) in imperative sentences?" This question

turns out to be rhetorical, as Alcázar and Saltarelli do not provide any such evidence, and I have none to advance here regarding imperatives.[18]

Finally, Alcázar and Saltarelli (2014) omit the I position between C and v.[19] The I position will be crucial to my analysis, following the proposal of Ritter and Wiltschko (2014).

4.3   Wiltschko (2014) and Ritter and Wiltschko (2014)

Wiltschko (2014), developing joint work with Ritter (Ritter and Wiltschko 2009, 2014), argues that there is a universal spine of functional categories, which are variably realised by language-specific content. Among these universal categories are I(NFL(ection)) and Point-of-view (Pov). Wiltschko's theory has implications for the mechanics of acquisition, while Ritter and Wiltschko's (2014) analysis of imperatives strongly influences my analysis in §5.

4.3.1   The universal spine

Wiltschko (2014) tackles the problem of how to reconcile syntactic universalism with the plethora of attested linguistic variation. The two major strands of existing theories fail to address this problem. One, the No Base Hypothesis (NBH), argues that there are no universal categories in language; but this predicts unsystematic variation, in the face of strong systematicity. The other, the Universal Base Hypothesis (UBH), sees UG as a repository for a strictly ordered hierarchy of functional categories identified by their content, as in Cinque's (1999) theory of Cartography. UBH is challenged by three kinds of variation (Wiltschko 2014:19ff.). First, not every category provided by UG is attested in every language; for example, Blackfoot (Algonquian) has no category tense. Second, categories are not formally identical across languages; for example, tense marking is obligatory in English, but optional in Halkomelem (Salish). Third, the logic of the UBH dictates that any category attested in a single language must be an innate part of UG; therefore direct/inverse person marking, found in Blackfoot, must be included in the UG hierarchy.

In response, Wiltschko and her colleagues have advanced a theory of the universal spine which divorces universal categories from their language-specific substantive content. Ritter and Wiltschko (2009:156) introduce this thesis as the Parametric Substantiation Hypothesis (PSH):

---

[18] My scepticism applies only intraclausally. There is persuasive evidence that the speaker, along with the addressee, is represented above C in the syntactic speech act domain: Heim et al. (2014), Wiltschko (2015), Biberauer and Vikner (2015). Compare my analysis of optative hortatives in §5.4.

[19] It is unclear whether for expository or theoretical reasons.

(10) PSH:

The substantive content of a given functional category is subject to parametric variation, constrained only by the universally determined core function of that category.

By PSH, there are universal categoremes, which are substantiated by language-specific allocategories (Ritter and Wiltschko 2009). Each universal category is defined by its unique spinal function (Dechaine and Wiltschko 2010). Wiltschko (2014) combines these ideas as the Universal Spine Hypothesis (USH). The universal spine provided by UG is a hierarchical organisation of universal categories κ (Wiltschko 2014:28):

(11) C_{UG} = κ:discourse-linking > κ:anchoring > κ:point-of-view > κ:classification

κ is transitive, relational, and requires substantiation (Wiltschko 2014:310-3): it relates two pronominal situation arguments (Pro-sit) via its unvalued coincidence [$u$coin] feature,[20] which requires substantiation by language-specific content:
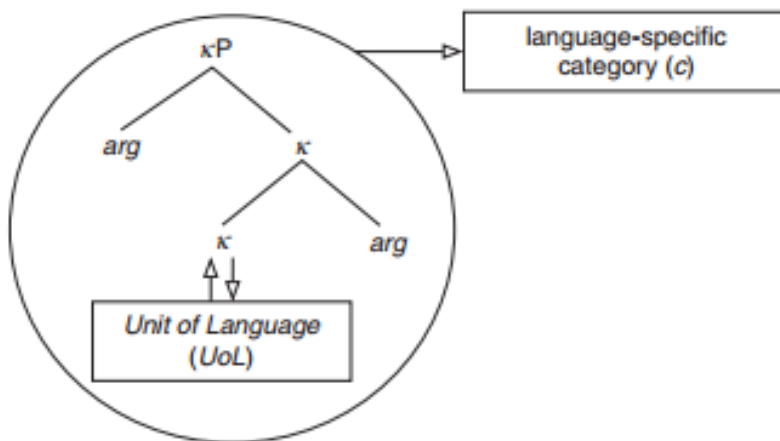
Figure 10: The universal structure of categories and their language-specific instantiations, from Wiltschko (2014:310)

Language-specific categories c are constructed from the set of universal categories κ, provided by the universal spine, and language-specific units of language (UoL) by the operation Associate (Wiltschko 2014:84):

(12) Associate:

c = κ + UoL

Thus the theory dissociates language-specific categories from their universal function and their language-specific form. κ mediates the relation between a form and its interpretation, analogous to how syntax mediates between PF and LF (Wiltschko 2014:303).

---

[20] On the notion of coincidence, see Hale (1986:238).

Take, for example, κ:anchoring.[21] This universal category performs the function of anchoring the reported event to the utterance. Its [*u*coin] feature is m(orphologically)-valued by different substantive content in different languages. In English, the substantive content is tense: present tense asserts the coincidence [+coin] of the event time and the utterance time, whereas past asserts their non-coincidence [-coin]. However, [*u*coin] can equally be valued by other aspects of the context of the situation: place in Halkomelem and person in Blackfoot (Ritter and Wiltschko 2005, 2009, 2014; Wiltschko 2006), and realis – whether or not the state of affairs holds in the real world – in Upper Austrian German (Wiltschko 2014:§4.4.3,130-9).

4.3.2 Point-of-view

Wiltschko (2014:ch.7,249-298) also introduces κ:point-of-view. This category is positioned between κ:anchoring and κ:classification, and relates the event to a point-of-view. In familiar Indo-European languages, κ:point-of-view is substantiated by temporality as outer aspect.[22] The combination of temporality and point-of-view yields Reichenbach's (1947) notion of reference time. Imperfective aspect realises [+coin] when the point-of-view is inside the event, and perfective aspect [-coin] when the point-of-view is outside ([-coin]). However, perfective versus imperfective aspect is not grammatically marked in 121 of the 222 languages in Dahl and Velupillai's (2013) WALS survey. This is because κ:point-of-view can be substantiated by UoLs that do not relate to temporality: control marking in Squamish (Salish), and direct/inverse marking in Blackfoot.

4.3.3 Implications for acquisition

Typologically, the universal spine provides von Humboldt's (1829) *tertium comparationis*, a third element for comparison, for a formal typology based on meaning not structure (Wiltschko 2014:305). Wiltschko (2014) identifies two key diagnostics for κ: multifunctionality and contrast. Multifunctionality arises when a single UoL is associated with two different κs, yielding two different interpretations. Polysemy should therefore be the default methodological assumption, rather than accidental homophony (Wiltschko 2014:96f.). Contrast arises because κ must be obligatorily present, which means that the UoL that substantiates κ may sometimes be only partially interpreted. Given that UoLs are pairings of sound and meaning, Π-Σ, expletive interpretation arises when Π is interpreted, but not Σ. Conversely, zero-marking follows from Σ being interpreted, but not Π. We

---

[21] Also known as INFL (Ritter and Wiltschko 2009, 2014).
[22] Compare inner aspect, which has to do with the internal temporal make-up of an event; see Travis (2010) for detailed definitions.

therefore expect polysemy, syncretism, 'fake forms',[23] expletives, and zero-marking to be rife in language.

While her primary concern is with contrast and multifunctionality as diagnostics for formal typological work, Wiltschko (2014:323) notes that the same diagnostics are predicted to hold as discovery procedures in acquisition. Children will approach the PLD with a default assumption of polysemy, and will respond to it by constructing the language-specific categories c for their language. This prediction is strongly parallel to the proposals of Biberauer (2014a, 2015a) that acquisition is driven by positing formal features in response to departures from Saussurean arbitrariness. The typological diagnostics of USH can therefore also be profitably applied to a neo-emergentist theory of language acquisition.

4.3.4   Implications for the syntax of imperatives

In addition to the implications of Wiltschko's (2014) theory for acquisition, the analysis of Ritter and Wiltschko (2014), within the same framework, has implications for the syntax of imperatives.

So far, we have encountered only m-valuation, whereby [$u$coin] is valued directly by substantive content. However, [$u$coin] can also be valued by external semantic content (Ritter and Wiltschko 2014; Wiltschko 2014:ch.5,145-187). In embedded clauses, [$u$coin] in INFL[24] can be valued externally by a higher lexical predicate, giving pred-valuation. In matrix clauses, [$u$coin] in INFL can be valued externally by a higher functional head, giving f-valuation. The interpretation of Pro-sit in spec,INFL is determined differently in each type of valuation. In m-valuation, it is interpreted deictically with reference to the extra-linguistic context as the utterance situation. In external valuation, Pro-sit is anaphorically bound: by the matrix event argument in pred-valuation, and by an argument in spec,COMP in f-valuation.

The two types of external valuation combined with the two values of [coin] correctly predicts a four-way typology (Ritter and Wiltschko 2014):

---

[23] For example, past tense morphology in English counterfactuals, as in (i), is an example of fake marking (Ritter and Wiltschko 2014:1370f.) It carries no temporal force (as shown by its compatibility with *now*), but instead reflects the fact that INFL is valued [-coin] (cf. past as non-coincidence with the present) by the counterfactual content of C (see the next subsection).

    (i) If John had a bike right now, he would cycle.

[24] In this subsection I follow the terminology of Ritter and Wiltschko (2014), who retain COMP and INFL; compare Wiltschko's (2014) κ:discourse-linking and κ:anchoring respectively.

		coincidence	
		[+coin]	[-coin]
external valuation	pred-valuation	simultaneous infinitives with aspectual verbs, e.g. *start*	future irrealis infinitives with desiderative verbs, e.g. *want*
	f-valuation	imperatives	counterfactuals

Focusing on f-valuation, in imperatives (Ritter and Wiltschko 2014:1367-75) the directive semantic content of the functional head COMP f-values INFL [+coin]. This asserts coincidence between the event situation (Ev-sit) in VP and Pro-sit in spec,INFL. Pro-sit is anaphorically bound by the Plan-set (Han 2001)[25] in spec,COMP. Thus, via Pro-sit, imperatives assert the coincidence of the Ev-sit with the Plan-set:
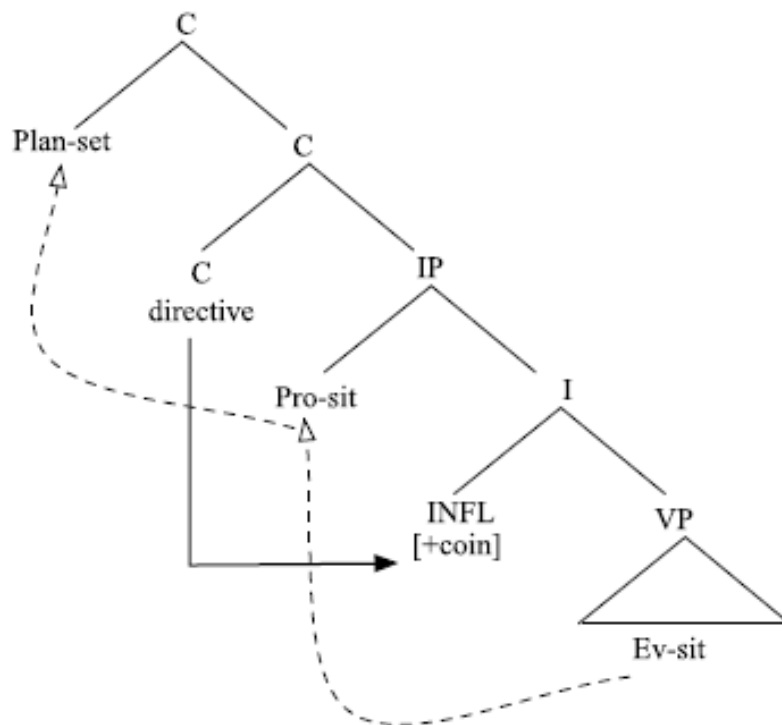
Figure 11: The structure of an imperative, from Ritter and Wiltschko (2014:1370)

In counterfactuals, the counterfactual semantic content of COMP f-values INFL [-coin]. This asserts non-coincidence between Ev-sit and Pro-sit. Pro-sit is anaphorically bound by the evaluation situation (Eval-sit) in spec,COMP. Thus, via Pro-sit, counterfactuals assert the non-coincidence of Ev-sit and Eval-sit:

---

[25] Cf. Portner's (2004) 'to-do list'.
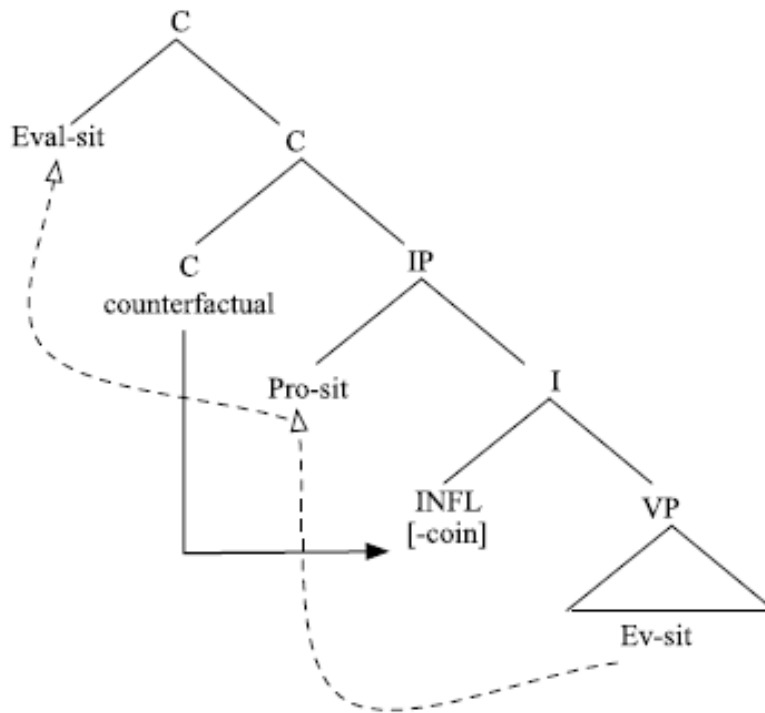
Figure 12: The structure of a counterfactual, from Ritter and Wiltschko (2014:1371)

[*u*coin] can only be valued once, so f-valuation is mutually exclusive with m-valuation. Focusing on imperatives,[26] this mutual exclusivity correctly predicts that there will be no tense in English imperatives. This holds cross-linguistically of tense languages, to the extent that many generative analyses of imperatives involve a defective TP, or a lack of TP altogether (see van der Wurff 2007:21 for references).[27] The lack of tense in English is mirrored in Halkomelem by the obligatory absence from imperatives of locative auxiliaries, which are the UoLs that m-value INFL in matrix clauses. Ritter and Wiltschko (2014:1373) emphasise the significance of the Halkomelem evidence for showing that the obligatory absence of m-marking is a syntactic, not semantic, effect. In tense languages, the obligatory absence of tense could arguably derive from a logical incompatibility, because an addressee cannot be ordered to do something in the past, or something that they are already doing. However, this same argument does not apply to location, because there is no logical restriction on ordering an addressee to carry out an action here, there, or anywhere. Overall, INFL cannot be associated with substantive content in imperatives for the syntactic reason that [*u*coin] is valued externally by f-valuation.

More broadly, Ritter and Wiltschko's (2014) analysis provides a formal explanation for the cross-linguistic tendency for imperatives to have unmarked or minimally marked form (van der Wurff

---

[26] See footnote 23 on counterfactuals.
[27] Compare §§4.4, 4.5 and 4.7.

2007:2). Likewise, Zhang's (1990) cross-linguistic survey of 46 languages from 13 families finds none with a tense distinction in the imperative paradigm.[28] Functionally oriented scholars attempt to derive the tendency for minimal marking from the vague notion of iconicity, whereby the short and simple form of imperatives embodies their brusque directive function (e.g. Aikhenvald 2010:46). On the formal approach here, by contrast, minimal marking in imperatives derives from the mutual exclusivity of m-valuation and f-valuation.

4.3.5   Evaluation

From Ritter and Wiltschko (2014) I adopt the analysis that I(NFL) has a [*u*coin] feature, which is f-valued [+coin] by the directive force content of C(OMP), for me represented as [imp]. From an acquisition perspective, imperatives give children the insight that I can be valued externally, not just by m-marking. This valuing relation derives the core meaning of imperatives, in asserting the coincidence of the Ev-sit with the Plan-set. In adopting the valuation relation, I also adopt Wiltschko's (2014) proposal that κ relates two situation arguments, where the situation comprises aspects of the context such as time, place, participants and world. In addition, I integrate Wiltschko's (2014) κ:point-of-view position into Ritter and Wiltschko's (2014) analysis of imperatives. I argue that κ:point-of-view universally comes to be substantiated by the addressee participant in imperatives.[29]

When it comes to negative imperatives, which are not considered by Ritter and Wiltschko (2014), I argue that negation values κ:point-of-view [-coin]. In this I take my lead from Wiltschko (2014:170-2), who argues that a negative marker in Halkomelem f-values [-coin]. I will expand this assumption to comprise projecting negative heads, which f-value [-coin], as well as adjoined negators, which pred-value [-coin].

The main point I will not adopt is that the universal spine is a pre-given part of UG (Wiltschko 2014). I will develop this line of thinking in §6.6.

---

[28] However, Zhang (1990) finds several languages that use a future suffix or particle on the imperative. The analysis here would argue that this is 'fake' morphology, but there is the question of how non-present morphology can reflect [+coin]. Perhaps the future morphology reflects coincidence with the Plan-set, which pertains to future actions.

[29] If correct, my position contradicts Wiltschko's (2014:287) prediction that the same part of the situation must be used to substantiate both κ:anchoring and κ:point-of-view in a given language. For example, English generally uses tense to m-value κ:anchoring, but, along with all languages, uses the addressee to substantiate κ:point-of-view in imperatives.

4.4   Zanuttini's Jussive head

Zanuttini (2008) and Zanuttini, Pak and Portner (2012) argue that jussive sentences, which subsume imperatives, are characterised by a particular functional projection, the Jussive Phrase (JP). J is positioned between C and T, which is defective in person features. The T head-moves up to J to form a bundle, with J providing the person features:
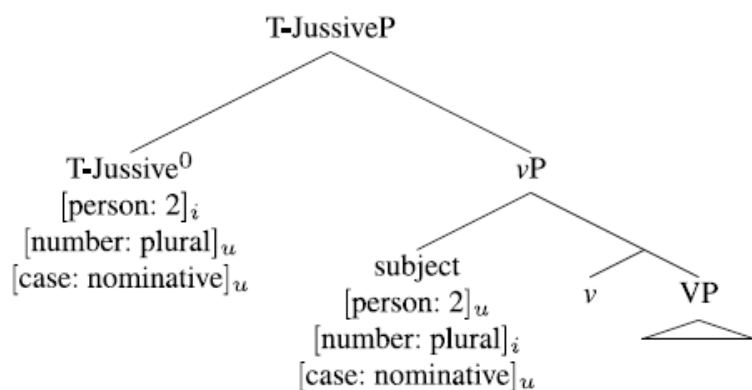
Figure 13: The structure of a canonical imperative, from Zanuttini, Pak and Portner (2012:1246); T moves to J, whose interpretable person features Agree with the person features of the subject

In canonical imperatives, J has second person features, giving the second person subject restriction:

(13) You/*He/*I[30] go!

The person features of J allow it to license null subjects, along the lines of rich agreement analyses for pro-drop languages:

(14) *pro* go!

Proper name and quantificational subjects can come to have second person features, since both can bind a second person reflexive in imperatives. J binds the null D of proper names, and the null partitive phrase within quantifier phrases:

(15) Mary look after yourself.

(16) Everybody look after yourselves.

Furthermore, Zanuttini, Pak and Portner (2012) extend the JP analysis to other jussive clauses, which comprise, in addition to imperatives, promissives and exhortatives. In these sentences, J has first person singular and first person plural inclusive features respectively, binding the speaker in

---

[30] Ungrammatical on the intended imperative reading.

promissives, and the speaker and addressee together in exhortatives. Evidence for these three varieties of J head comes from Korean, where it is argued that each is realised by a different sentence-final particle (Zanuttini, Pak and Portner 2012:1234):

(17) Cemsim-ul   sa -la/   -ma/ -ca

    lunch   -ACC buy-IMP/-PRM/-EXH

    Imperative: 'Buy lunch!'

    Promissive: 'I will buy lunch.'

    Exhortative: 'Let's buy lunch.'

Meanwhile, some languages have third person forms in their imperative paradigm; e.g. Bhojpuri (northern India, Uttar Pradesh state) (Zanuttini, Pak and Portner 2012:1252):

(18) k^ha:

    eat-IMP.2 (informal)

    'Eat!'

(19) k^ha:y

    eat-IMP.3

    'Eat (polite)!'/'Let him eat'

(19) can have both directive and optative-like uses. In its directive use as an imperative, T has person features, which intervene between J and the subject, blocking the second person restriction. Zanuttini, Pak and Portner (2012:1252) leave the optative meaning for future research.

4.4.1   Evaluation

I assume a head which places a second person restriction on imperative subjects, namely Point-of-view (Pov). However, whereas Zanuttini (et al.) argue that JP is present only in jussives, I argue that Pov is present in other clause-types, along the lines of Wiltschko's κ:point-of-view. Furthermore, whereas J is between C and T, my Pov has a lower position between I and v, so as not to interfere with C f-valuing I [+coin]. I retain the idea that proper name and quantificational subjects acquire second person features from J, recast as Pov. However, the position of Pov below T prevents me from adopting this analysis for morphologically third person imperatives, because T does not intervene between Pov and the subject. Instead, I claim that syntax can target either inherent or syntactically valued features. The ambiguity between directive and optative readings for third

person imperatives in Bhojpuri further suggests that the collapsing of hortatives into imperative structures is not as simple as Alcázar and Saltarelli (2014) would have it.

4.5 Johannessen's (2015) prescriptive infinitives

Johannessen (2015) observes that the PLD in Nordic languages (Danish, Faroese, Icelandic, Norwegian, Swedish) exhibit two forms of the imperative. Alongside 'finite imperatives', we find 'child-directed prescriptive infinitives'. Prescriptive infinitives are pragmatically restricted to intimate yet hierarchical settings, from a parent to their child. There are striking syntactic differences between the two:

Table 5: A summary of the syntactic differences identified by Johannessen (2015) between finite imperatives and prescriptive infinitives

	Order of verb and negation	Order of verb and subject	Person of the subject	Person of the object	Pronouns as arguments?
Finite imperative	V – neg	V – subject	2	1	yes
Prescriptive infinitive	neg – V	subject – V	3	3	no
Supporting data[31]	(20)	(21)	(22)	(23)	(24)

(20) Order of verb and negation: (Swedish)

    a.   Kom   inte hit   med dig!   (Teleman et al. 1999:2777)

        come.IMP not   here with you

        'Don't come here, you!'

    b.   *Inte kom hit med dig!

    c.   Inte hälle   mjölken   (Childes, ant23_08.cha)

        not   pour.INF milk.DEF

        'Don't pour the milk!'

---

[31] Data directly from Johannessen (2015), unless otherwise indicated.

(21) Order of verb and subject: (Norwegian)

    a.   Spør   du meg ikkje   (Nordic Dialect Corpus)
        ask.IMP you me not
        'Don't you ask me!'

    b.   Nora sitte   rolig   der   Nora   (Childes, nora2.cha)
        Nora sit.INF quietly there Nora
        'Nora, sit quitely there, Nora'

(22) Person of the subject: (Norwegian)

    a.   Ikke tegn   deg   selv!
        not draw.IMP you.2 self
        'Don't draw yourself!'

    b.   Ikke tegne   seg   selv!
        not draw.INF her/him.3 self
        'Don't draw yourself (*lit*. oneself)!'

(23) Person of the object: (Norwegian)

    a.   Susse   mamma!
        kiss.INF mummy.3
        'Kiss mummy!'

    b.   *Susse   meg
        kiss.INF me.1.SG.ACC
        'Kiss me!'

(24) Pronouns as arguments?   (Norwegian)

    a.   *Hun   sitte   rolig   der!
        she.NOM sit.INF quietly there!

Johannessen (2015) proposes that prescriptive infinitives lack TP. This derives the preverbal negation and subject orders, since there is no position above negation for the verb to move into. The lack of T is also claimed to derive the ban on pronouns:[32] without T, there is no Case-licensing, and pronouns are the only DPs that are overtly case-marked in Norwegian.

4.5.1 Evaluation

Johannessen's (2015) data highlight how the PLD can provide acquirers with significantly more varied information than traditional grammars might suggest.[33] The contrast between finite imperatives and prescriptive infinitives has pragmatic overtones, but also signals different formal properties. Regarding Johannessen's analysis, her link between T and pronouns is suspect in making the very non-standard assumption that only DPs with overt morphological case have abstract Case. Moreover, though prescriptive infinitives lack TP, they remain for Johannessen fundamentally clausal, under a CP projection. I extend the lack of TP to argue more radically that prescriptive infinitives are in fact gerunds, lacking all clausal functional projections above vP. I support my analysis with related data from beyond the Nordic languages, analysing prescriptive infinitives along with generic imperatives as gerund imperatives in §§5.5 and 6.3.

4.6 Potsdam (2007) versus Rupp (2007)

Potsdam (2007) and Rupp (2007) debate the position of *do(n't)* in English imperatives. The pair of articles brings to a head a debate extending back to Potsdam (1998) and Rupp (1999, 2003). Potsdam argues that *do(n't)* is in C, Rupp in I. Rupp's argument is more persuasive in identifying a position in imperatives between T and v, FP, which I take to instantiate Pov.

Potsdam argues for the CP hypothesis that the English imperative can be analysed using the conventional CP-IP-VP clause structure, essentially following the analysis of Beukema and Coopmans (1989), among others (see Potsdam 2007:251 for extensive references). For example, Potsdam would analyse (25) with the subject in spec,IP, and movement of *do(n't)* from I to C:

(25) Do(n't) you have another drink!

(26) $[_{CP} [_C \text{Do(n't)}_i] [_{IP} \text{you}_j [_I t_i] [_{VP} t_j \text{ have another drink}]]]$

This involves the same head movement operation as in polar questions:

---

[32] Which in turn rules out non-third person DPs.
[33] Cf. §3.4 on not under-estimating imperatives in the PLD.

(27) $[_{CP} [_C Do_i] [_{IP} you_j [_I t_i] [_{VP} t_j \text{ like football}]]]$

However, the parallelism between the English imperatives and polar questions breaks down under more detailed scrutiny. Potsdam (2007:252f.) admits Davies' (1986) observation that subject-auxiliary inversion is optional in imperatives; cf. (25):

(28) You do(n't) have another drink!

In this case, *do(n't)* remains in I:

(29) $[_{CP} [_C \emptyset] [_{IP} you_j [_I do(n't)] [_{VP} t_j \text{ have another drink}]]]$

However, as Rupp (2007:313ff.) points out, this optionality greatly weakens the parallelism between imperatives and polar questions, where I-to-C movement is obligatory. There is a further difference regarding the inversion behaviour of separable *do* and *not* (cf. Rupp 2007:315f.):

(30) Do you not like football?

(31) *Do you not have another drink!

In light of these difficulties for the CP hypothesis, Rupp advances the FP hypothesis: *do(n't)* is merged directly in I, where it remains (cf. Zwicky and Pullum 1983). Meanwhile, the subject ordinarily raises from its first merge position in spec,vP only as far as spec,FP, where F intervenes between I and v:

(32) $[_{CP} [_{IP} [_I Do(n't)] [_{FP} you_i F [_{VP} t_i \text{ have another drink}]]]]$

The subject may optionally raise from spec,FP to spec,IP.[34] Thus the optionality regarding the position of the subject arises from subject movement under the FP hypothesis, as opposed to head movement under the CP hypothesis:

(33) $[_{CP} [_{IP} you_i [_I do(n't)] [_{FP} t_i F [_{VP} t_i \text{ have another drink}]]]]$

Neither hypothesis claims that the subject remains in its base position, following conclusive evidence in Potsdam (1998:128-137) from VP ellipsis, passivisation, quantifier float, adverb placement, and aspectual auxiliary placement.

Potsdam and Rupp agree that the significant data for deciding between the CP and FP hypotheses come from adverb placement and negative scope. Both claim the facts to be in their favour, often

---

[34] Cf. Kiss (1996), who argues for an additional subject position between C and I.

due to divergent grammaticality judgements. Overall, however, Rupp's argument is more persuasive.

First, e(xtent)-adverbs (Jackendoff 1972) such as *simply*, *hardly*, and *already*, are only acceptable in clause-internal positions, and not clause-peripheral ones (cf. Potsdam 2007:263-4):

(34) (*Simply) he (simply) must (simply) be (simply) very ill (*simply).

Thus Potsdam (1998) concludes that e-adverbs cannot adjoin higher than I'. He claims support for the CP hypothesis by showing that the pattern of e-adverb placement in imperatives is the same as in polar questions (Potsdam 2007:266,(40c),(42(c)):

(35) *Hardly* should they have worried about that?

(36) *Simply* don't you stand there!

However, whereas (35) is definitely out, I judge (36) to be perfectly acceptable, along with Rupp's (2007:305f.,(20b),(21)) counterexamples:

(37) Simply do not give them your address!

(38) (Just) don't anyone (just) believe what he says!

The contrast between (35) and (36-38) suggests a structural difference between polar questions and imperatives that is not captured by the CP hypothesis. By contrast, the FP hypothesis maintains the analysis that e-adverbs can only adjoin lower than IP:

(39) [IP [I' Simply [I do] not give them your address]]

Overall, therefore, the adverb placement facts favour Rupp's FP hypothesis.

The second set of deciding data concern negative scope. The relevant facts were observed by Schmerling (1982), and are reported by Rupp (2007:308f.):

(40) Everyone didn't get a raise.
     = nobody got a raise   every > not
     = not everybody got a raise   not > every

(41) Don't everyone expect a raise!
     ≠ nobody expect a raise!   *every > not
     = not everyone should expect a raise   not > every

41

In sum, imperative subjects take narrow scope with respect to the preceding negation (Potsdam 2007:257).

Rupp and Potsdam each adopt a different theory of quantifier interpretation to support their position. Rupp (2007) adopts Hornstein's (1995) A-movement theory, whereby Neg can be interpreted in any of its A-positions. According to the FP hypothesis, *don't* in imperatives has only one position: it is first merged in I, and remains there. This correctly derives its single negative scope interpretation. According to the CP hypothesis, by contrast, *don't* has two A-positions: it is first merged in I, before moving to C. On the A-movement theory of quantifier interpretation, this wrongly predicts that both negative scope readings should be available. Therefore, if Hornstein's (1995) theory is adopted, Rupp can claim the scope data for the FP hypothesis.

In response, Potsdam (2007) turns the argument on its head, claiming that (40-41) in fact show that Hornstein's (1995) theory is incorrect. Instead, he sticks to May's (1977, 1985) theory of quantifier raising for (40), and claims that inverted negation in C always takes widest scope to account for (51). In sum, therefore, deciding between the CP and FP hypotheses on the basis of the scope data reduces to deciding between the competing theories of quantifier interpretation.

Overall, while the scope data are equivocal for theory-internal reasons, Rupp's judgements and examples regarding e-adverb placement point tentatively in favour of the FP analysis.

4.6.1   Evaluation

Concluding in favour of the FP analysis is also conceptually appealing, because it means that English imperatives offer acquirers information about clause structure that contrasts with other clause-types. Potsdam (2007:268) argues that imperatives have an unexceptional clause structure, and implies that they would otherwise pose an acquisition problem, since a special syntax would "require exceptional derivational strategies or structures unsupported elsewhere in English." It would therefore be advantageous for different clause-types to have the same structure, so that intake could be generalised between them (cf. van der Wurff 2007:85). Interpreted differently, however, distinctive syntax would helpful to the acquirer. On the approach taken here, we expect different clause-types to provide different information to the acquirer; identity would be a wasted opportunity for acquisition. In particular, the special syntax of the English imperative under Rupp's FP hypothesis provides crucial acquisition evidence for a position between I and v, namely Pov (cf. Wiltschko 2014).[35] Contra Potsdam (2007:268), there is further evidence for this position in

---

[35] Indeed, Rupp (2007:312-3) speculates that the identity of FP could be AspP (Tenny 1987, Ouhalla 1991), which Wiltschko (2014) reinterprets as κ:point-of-view.

imperatives and beyond, in English and cross-linguistically, as we will see in §6.1. Pov plays a crucial role in my analysis of imperatives. Before setting out my analysis, the final body of previous work to consider are theories concerning negative imperatives.

4.7 Analyses of negative imperatives

A large body of research, mostly restricted to Romance languages (see van der Wurff 2007:59ff. for references), investigates the ban on true negative imperatives (TNIs) (Zanuttini 1994). Some analyses ascribe the ban to a syntactic problem in terms of failed head-to-head relationships (Zanuttini 1991 et seq.), others to a semantic problem of negation outscoping imperative force (Han 2001).

4.7.1 Syntactic problems

Syntactic analyses argue that TNIs are banned because: (i) head-to-head relations are blocked by negation; (ii) negation and the imperative are structurally mutually exclusive; or (iii) certain heads go unlicensed in negative imperatives. First, the blocking analysis (e.g. Zanuttini 1991, 1994; Rivero and Terzi 1995) claims that negation blocks the necessary relationship between [imp] in C and the verb, due to some version of Rizzi's (1990) Relativised Minimality. Second, Laka (1990:245-52) argues that negation and the imperative verb are mutually exclusive, because they compete for the same structural slot: the polarity phrase, ΣP.

The third analysis locates the problem in the imperative verb's inability to license tense, finiteness or mood, which are in a selection relationship with Neg. In one version (Zanuttini 1996, Platzack and Rosengren 1998), NegP bears a [finite] feature, which is licensed by TP/FinP, meaning Neg cannot appear without it. However, due to the imperative verb's morphological defectiveness in tense and mood features, it is unable to license TP/FinP. Since NegP requires TP/FinP, but the imperative verb cannot license TP/FinP, TNIs are banned. On another version of this analysis (Zanuttini 1997), the relationship between NegP and other functional projections projections is reversed: Neg selects MoodP, which the imperative verb cannot license.

However, the ban on TNIs is not absolute. In their WALS survey, van der Auwera and Lejeune (2013b) find that 113 out of 495 languages allow TNIs. Accounting for this observation requires an analysis that can allow and ban TNIs according to different structural circumstances in different languages. Zeijlstra (2004:181ff.) stipulates that the difference reduces to physical blocking: where negators are heads they act as minimality barriers for head-to-head relations and block TNIs; but where the negator is phrasal in the specifier of NegP, the null Neg head position allows head to head

relations to be established through it. Meanwhile, Postma and van der Wurff (2007) argue that a minimality problem only arises when Laka's ΣP is collapsed together with the volitional/boulemaeic force projection, BoulP. This collapsing together is morphologically reflected by languages having the same form for their anaphoric negator as for their basic clausal negator (e.g. Portuguese *não* versus English *no/not*), a correlation that holds for all 23 languages surveyed. Where the two negators are identical, the articulated structure with distinct BoulP and ΣP projections would not be acquirable. This analysis is intuitively appealing: in languages with a reduced clause structure, there is not the room to express both negative and imperative at the same time. However, the mechanics of the analysis rely on odd assumptions about movement; namely, that verb movement to C is A'-movement, which is permitted though BoulP, an A'-projection, but not through ΣP, an A-projection. The collapsed BoulP/ΣP then counts as an A-projection, blocking verb movement to C.

4.7.2 Semantic incompatibility

In addition to the syntactic analyses, there are semantic analyses that ascribe the ban on TNIs to the incompatibility of negation and imperative force. Reichenbach (1947:338,342)[36] argues that pragmatic moods, as speaker's instruments, cannot be negated: i.e., *¬f(p). Only the propositional content can be negated: f(¬p). Zeijlstra (2013:871) attributes this insight further back to Frege (1892), and explains that negation cannot outscope the illocutionary force of any speech act: a negative command is still a command, just as a negative assertion is still an assertion, and a negative question is still a question. Therefore, negation logically cannot outscope illocutionary force. Since illocutionary force is standardly assumed to take scope from matrix C, it should be impossible for semantic negation to c-command it from spec,CP. Overall, TNIs should be banned where the structure of the language would position Neg above C in imperatives.

This is the reasoning in Han (2001). Han argues that where negation is a head that procliticises to the verb, and where the negation-verb complex then moves to C, the negator will outscope the imperative operator, leading to semantic incoherence. This analysis would derive the ban on TNIs in languages where negation is a head, and the verb moves to C in imperatives. By contrast, where negation is phrasal, the semantic negation remains below the imperative operator, allowing TNIs. However, Han's analysis is beset by problems. For one, it relies on a non-standard definition of c-command, whereby negation can scope out of its position adjoined to the verb (Postma and van der Wurff 2007:220):

(42) [$_{CP}$ [$_{C}$ [Neg – V] C$_{[imp]}$] IP]

---

[36] As referenced by Alcázar and Saltarelli (2014:10,fn.6) and (2014:113,fn.19).

In addition, as Zeijlstra (2013:874f.) notes, Han's analysis does not account for why the Slavic languages, which have negative heads, allow TNIs. Han argues that the verb's features move covertly across negation from I to C, retaining the scope of the imperative operator over negation. Zeijlstra counters that if covert feature movement is available in Slavic, then it should be equally available to allow TNIs in Romance.

Zeijlstra (2013) proposes a route out of this problem by assuming his (2004) proposal that morphosyntactically negative heads need not be semantically active; instead they can carry an uninterpretable [uNEG] feature. In such instances, the semantically interpretable negation [iNEG] is a covert negative operator Op¬ in spec,NegP. Thus in Slavic, the [uNEG] negative head moves to spec,CP, and the verb moves to $C_{[imp]}$; but since the negative operator remains in spec,NegP, the imperative operator still scopes over negation, avoiding semantic incompatibility.

4.7.3   Evaluation

I adopt the argument that negation cannot outscope imperative force, while arguing further that negation cannot intervene between C and I. Otherwise, negation would block the valuation of I [+coin] by [imp] in C (Ritter and Wiltschko 2014), which would amount to negating the imperative force.

Regarding the syntactic analyses, the typological evidence makes clear that TNIs are only banned in some languages, which is not reflected in the absolutist analyses of Zanuttini (1991) et seq. Meanwhile, the analyses of Zeijstra (2004) and Postma and van der Wurff (2007) assume machinery regarding head versus phrasal status and A- versus A'-movement that do not follow from current minimalist theorising. I look to revise this difference in terms of Wiltschko's (2014) notions of substantiation versus modification.

4.8   Summary of formal analyses of imperatives

Table 6 summarises the each of the evaluation subsections in this section, setting up the analysis in the next section.

Analysis	Adopted and *Extended*	Dropped
1 Formalising the imperative	• [imp] in C syntactically encodes directive semantic content; *and values I [+coin]*	
2 Alcázar and Saltarelli (2014)	• syntax must refer to some aspects of the context; *the addressee is central in imperatives* • hortatives can be incorporated to the standard imperative, *though carefully distinguishing true let-imperatives from optatives*	• phasal CP(vP) relation • δ-relation • presence of the speaker within the clausal domain
3 Wiltschko (2014) and Ritter and Wiltschko (2014)	• [imp] in C f-values I [+coin], asserting the coincidence of the Ev-sit with the Plan-set • κ:point-of-view (*Pov*) between I and v, *universally substantiated by the addressee participant in imperatives* • Negative heads f-value Pov [-coin], *while adjoined negators pred-value Pov [-coin]*	• the universal spine as a pre-given part of UG
4 Zanuttini's Jussive head	• a head with second person features, *Pov, which is also present in clause-types beyond imperatives*	• JP, positioned between C and T
5 Johannessen's (2015) prescriptive infinitives	• prescriptive infinitives lack TP *and higher clausal structure, as gerunds*	
6 Potsdam (2007) versus Rupp (2007)	• Rupp's FP hypothesis, *where F = Pov*	• Potsdam's CP hypothesis
7 Analyses of negative imperatives	• negation cannot outscope imperative force, *nor can it intervene between I and C*	• stipulations regarding head/phrasal status, and A/A'-movement

5   An analysis of the structure of imperatives

In this section I offer an analysis of the structure of imperatives. I introduce the basic structure of affirmative imperatives, followed by the structure of the negative imperative. Hortatives can be collapsed into these structures, but with greater care than has previously been taken. Finally, gerund imperatives, comprising prescriptive infinitives and generic imperatives, have a radically different structure from standard imperatives. But first, I will introduce the central idea behind my analysis.

5.1 Central idea

My analysis captures the fact that imperatives universally hinge on the addressee. I claim that languages universally use the addressee participant part of the situation (ADDR) as the substantive content of Pov. That is, imperatives take the point-of-view of the addressee participant as their reference point.[37] This perspective-taking aligns well with the universal core function of Wiltschko's (2014) κ:point-of-view: to relate the event situation to the point-of-view situation. Imperatives assert the coincidence of the addressee's point-of-view with the Plan-set, which is the intended effect of uttering an imperative. The Pov head carries second person features in imperatives, and can host the imperative subject (YOU), first merged in spec,v, in its specifier.

5.2 Basic structure

Figure 14 depicts the basic structure of an imperative. Following Ritter and Wiltschko (2014), the relation between the Plan-set in spec,CP and the Pro-sit in spec,IP is established by anaphoric binding, which is represented by a dotted arrow. All other relations between situation arguments are asserted by the value of [coin] on the head between them. Coincidence between situation arguments is represented by a lined arrow. I assume that f-valuation can be effected as a kind of composed Agree by functional heads that have themselves been f-valued. Specifically, I, which is f-valued [+coin] by [imp] in C, can in turn value Pov [+coin]. F-valuation of [$u$coin] is represented by a block arrow.

---

[37] For an argument against representing the speaker participant in this position, see §6.5 on particles.
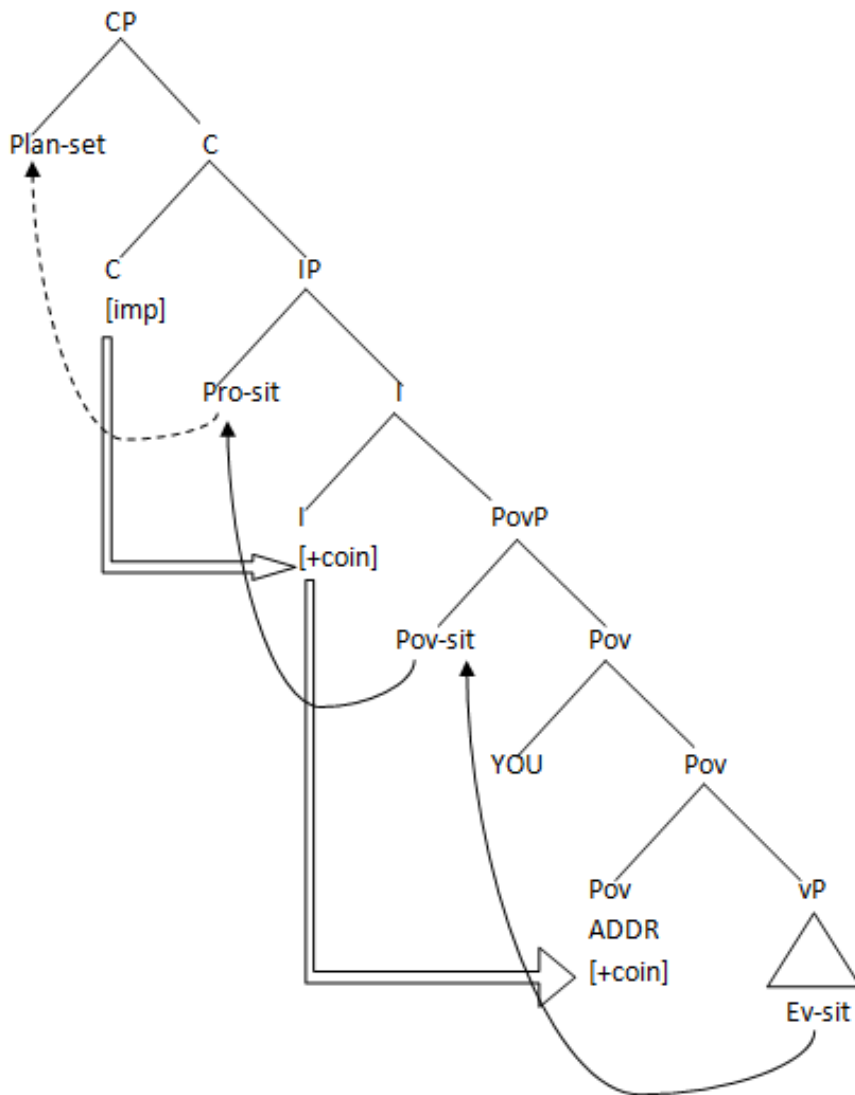
Figure 14: The structure of an affirmative imperative

This structure reflects the intended effect of a standard affirmative imperative: the event is added to the Plan-set from the addressee's point-of-view. This is because the Ev-sit is asserted to coincide with the addressee's point-of-view, which in turn coincides with the Plan-set.

5.3 Negative structure

Negative imperatives come in two types, depending on the structure of negation in the language. Here I draw on Wiltschko's (2014) notions of modification versus substantiation, corresponding approximately to the traditional notions of adjunction and substitution respectively – cf. Zeijlstra's (2004) phrasal verus head negation. In type 1 languages, the negator adjoins to PovP as a modifier, and does not project. In type 2 languages, the negator is a head, which substantiates part of the spine, projecting a phrasal label. The head negator is selected by I, and itself selects PovP.

In both types, the negator is able to value Pov as [-coin] (cf. Wiltschko 2014:170-2). In the modifying type 1, this mechanism is pred-valuation by lexical content; whereas in the substantiating type 2, it is f-valuation by a functional head. A crossed-through lined arrow signifies the non-coincidence of the two situation arguments, as asserted by [-coin]:
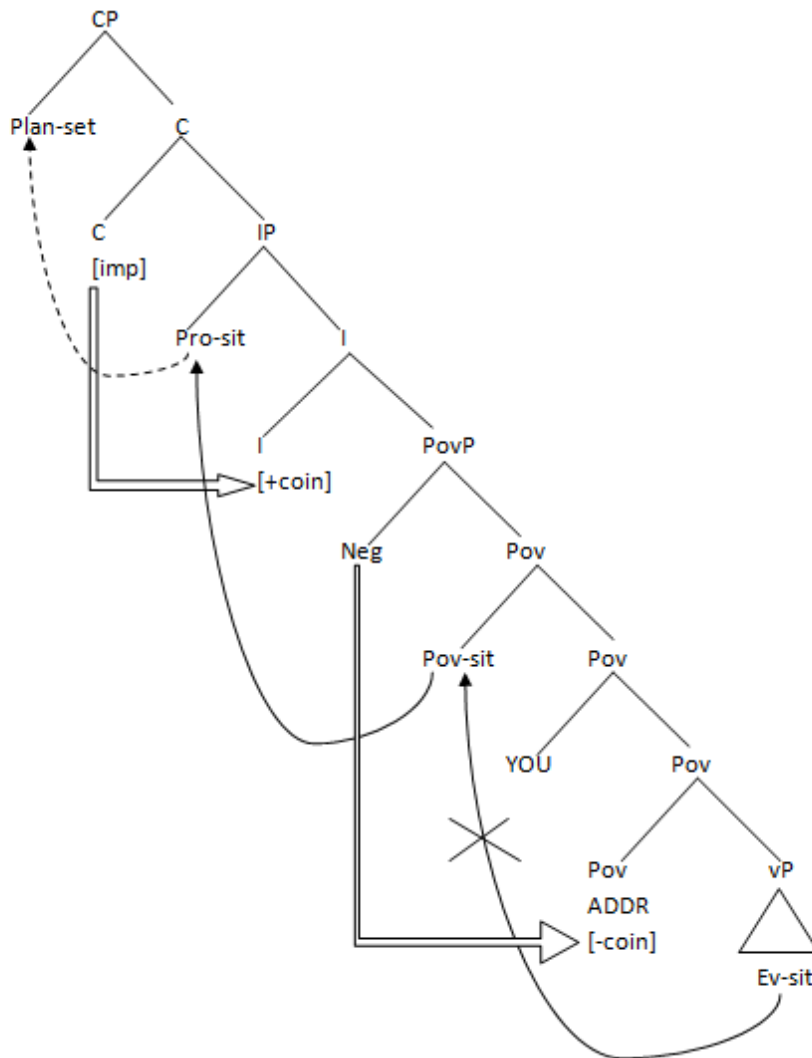
Figure 15: The structure of a negative imperative in a type 1 language, where the negator is a modifier that pred-values Pov [-coin]
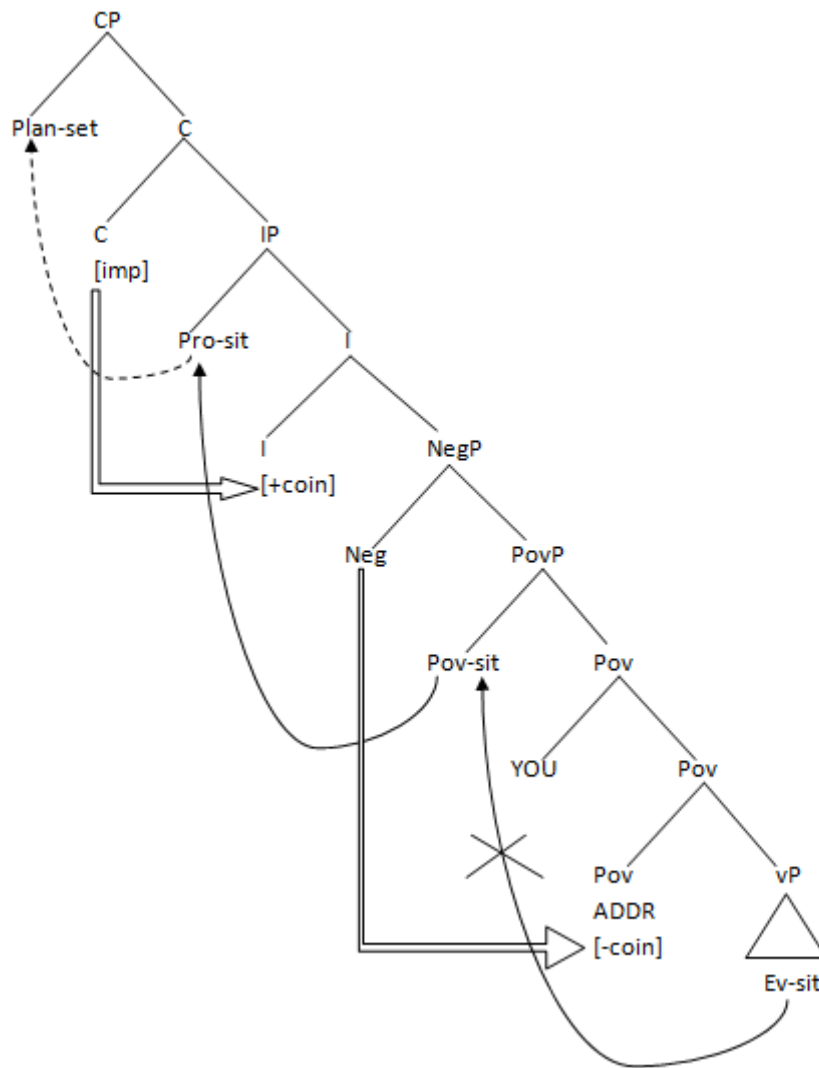
Figure 16: The structure of a negative imperative in a type 2 language, where the negator is a head that substantiates part of the clausal spine and f-values Pov [-coin]

In both types, the negator must be above Pov but below I. It must be above Pov in order to pred- or f-value it [-coin]. In addition, the negator must be below I due to an extension of the semantic incompatibility argument that negation cannot outscope imperative force §4.7.2. Negation cannot intervene between C and I, because it would block the valuation of I [+coin] by [imp] in C. Imperative force consists in the coincidence between the Plan-set in spec,CP and the addressee's point-of-view in spec,PovP, via Pro-sit in spec,IP. Blocking the valuation of I [+coin] would therefore amount to negating the imperative force, which would lead to semantic incoherence. From its lower position between I and Pov, negation asserts the non-coincidence of the addressee's point-of-view with the event situation. The structure thus models the intended effect of an imperative, which is for the addressee to plan not to do the described event.

The position of the negative head below I for type 2 languages is much lower than is usually assumed for, e.g., Romance (Poletto 2008). I argue below in §6.2 that this different position is what derives the ban on TNIs in some languages.

5.4 Hortatives

True hortatives can be analysed as standard imperatives, though care must be taken over those with an optative meaning, which have a different structure. To the extent that *let*-imperatives can be said to have the addressee mediating between the command and the performer, they can be assimilated to standard imperatives. Consider an English *let*-imperative, which can felicitously have an overt subject:

(43) Context: Mary is washing John's sports kit; observing this, Amy says to Mary:

(You) let John do it!

No extra causative verb position is necessary, contra Alcázar and Saltarelli (2014); instead, I claim that *let* realises the Pov position:
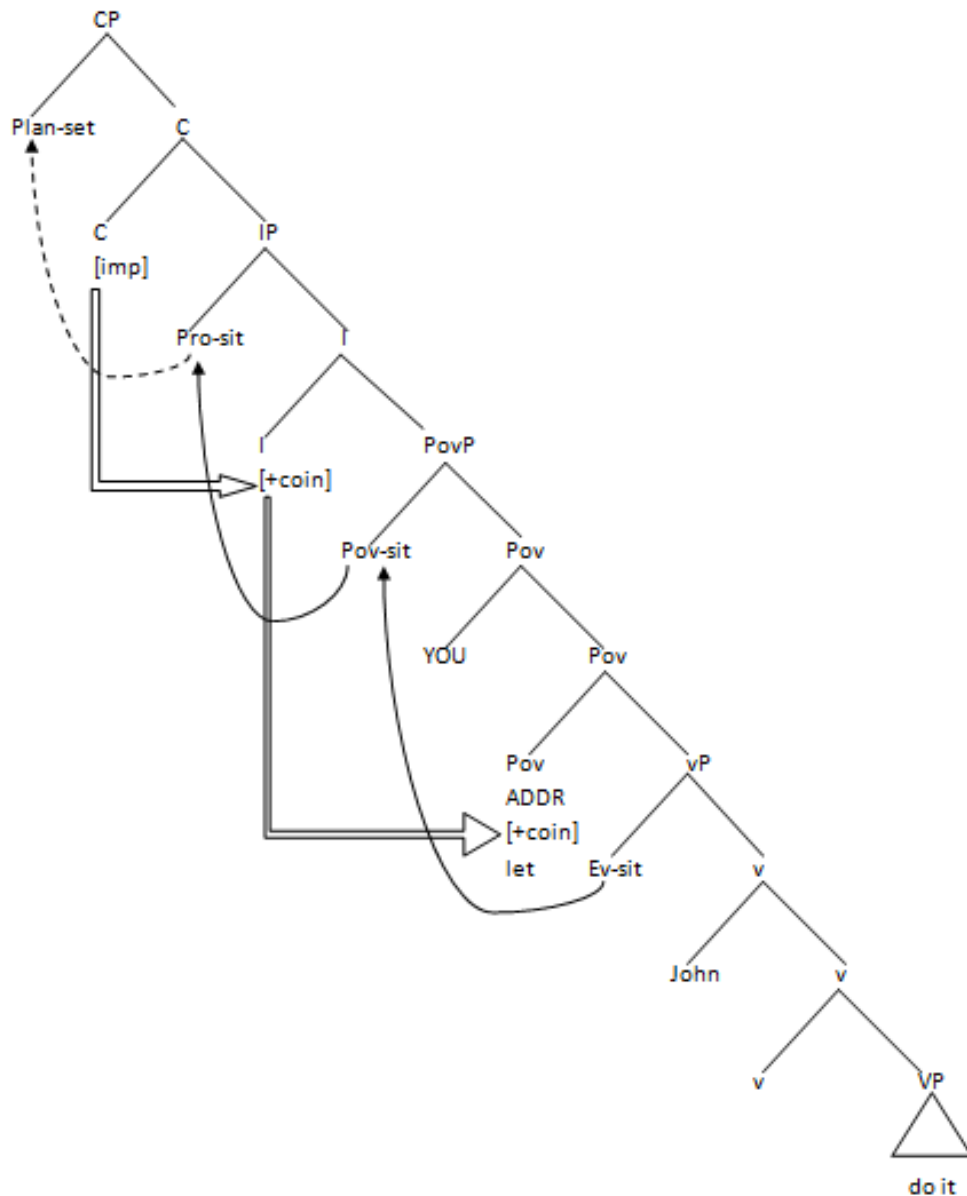
Figure 17: The structure of a hortative as a causative imperative

This simple collapsing of hortative into imperative above does not work for all so-called hortatives, however. Consider (43) in a different context, where it has an optative meaning and cannot have an overt subject:

(44) Context: Mary is watching John run a race, and hopes he wins; she says:

(*You) let John do it!

We saw in §4.4 that Zanuttini, Pak and Portner (2012:1252) note the possibility of this optative meaning for third person imperatives in Bhojpuri, but do not offer an analysis. I conjecture that optatives have the same syntactic structure as *let*-imperatives in Figure 17, except that Pov is substantiated by the speaker. Thus optatives align the Ev-sit with the Plan-set via the speaker's

point-of-view. The Plan-set need not be interpreted as attached to a particular speech act participant, but could be more abstractly predicated of the world. An optative would then express a wish from the speaker's point-of-view that the world make it such that the event occurs.

A final point to consider is how the imperative verb can come to be inflected with third person in hortatives in languages like Bhojpuri. I claim that there is parametric variation as to whether languages morphologically realise the inherent or syntactically derived features of the subject in these circumstances. Languages like Bhojpuri realise the subjects' inherent third person features, rather than the second person features it receives from Pov.

5.5   Gerund imperatives

I argue that Johannessen's (2015) prescriptive infinitives are gerunds. I incorporate generic imperatives in the same analysis, as found in instructions and on public signs:

> (45) No walking on the grass!

I term prescriptive infinitives and generic imperatives together gerund imperatives. I follow Abney's (1987) insight on gerunds in claiming that gerund imperatives are clausal structures up to vP, which are then embedded under D. The spec,v external argument position of gerund imperatives is occupied by PRO, which can receive a generic interpretation, or a contextually bound interpretation. Unlike in full imperative structures, the addressee is not syntactically represented, and PRO is not syntactically bound. Third person agreement is inserted by default (Preminger 2014). Likewise, the command meaning of gerund imperatives is not constructed syntactically. There is no C$_{[imp]}$, and no binding together of Ev-sits, points-of-view and Plan-sets. Instead, the command meaning of gerund imperatives is pragmatically determined. This lack of formal command semantics derives their underspecified, highly context-dependent meaning, which follows almost automatically from the scenario. The intimate yet directive meaning of prescriptive infinitives follows from their context of use by a parent to their child. Likewise, the command meaning of generic imperatives is also obvious in context: if you are reading (or hearing) a generic imperative, then it applies to you. I support this analysis of gerund imperatives in §6.3.

6   Further consequences

In this section I explore the consequences of my analysis of imperatives from a neo-emergentist acquisition perspective. First, imperatives and other clause-types provide evidence for a structural position Pov, in English and other languages, which suggests that acquirers generalise structural evidence from imperatives to other clause-types. Second, languages ban TNIs where their prohibitive negator is lower than its usual position above I, a difference which is signalled to the acquirer by a different verb form. Third, there is evidence to support my analysis of prescriptive infinitives and generic imperatives as gerund imperatives, with implications for the acquisition of nominal syntax. Fourth, acquirers generalise imperative syntax to realise other structurally related clause-types, especially counterfactuals. Fifth, particles have different meanings in imperatives than declaratives, signalling to the acquirer that each clause-type is characterised by different semantic content in C. Finally, I am not committed to a rich UG by couching my analysis in terms of USH, which can instead be derived from extralinguistic cognition.

6.1   Evidence for Pov

My analysis of imperatives argues that Pov is universally substantiated by the addressee participant in imperatives. This structural position, also proposed by Rupp's (2007) FP analysis, is not "unsupported elsewhere in English" (Potsdam 2007:268); rather, supporting evidence comes from English, in imperatives and beyond, as well as cross-linguistically.

6.1.1   English imperatives

Regarding imperatives, standard English evidence for Pov comes from the behaviour of *be* and *have*, and from the subject placement possibilities of negative imperatives of unaccusative verbs. Dialectal data regarding *let('s)* and subject positioning provide additional evidence.

First, as Pollock (1989) observed, *be* and *have* do not raise in English imperatives, in contrast to their behaviour in polar questions:

(46) Do(n't) be silly!

(47) Are you silly?

(48) Do(n't) have arrived!

(49) Have you arrived?

Whereas in polar questions *be* and *have* raise to C, they remain low in Pov in imperatives, with *do(n't)* in I blocking movement.

Second, consider the position of low overt subjects in negative imperatives of unaccusative verbs:

(50) Don't you go!

Unaccusatives do not have an external argument-introducing little v*. Instead, their subject raises from the internal argument position as sister to the lexical verb. Assuming *don't* is merged in I, the overt subject *you* must therefore be in a position below I but above V; here, spec,Pov:

(51) [$_{IP}$ [$_I$ Don't] [$_{PovP}$ you [$_{Pov}$ ∅] [$_{VP}$ [$_v$ go] [$_{DP}$ ~~you~~]]]]

Third, *let's* imperatives in some innovative dialects of English provide evidence for a Pov position. Collins (2004:301) presents the following data:

(52) let's you and me go for a walk down by the Snake   (Brown University Corpus, N13,2)

(53) oh Elli let's me sit opposite you   (Bergen Corpus of London Teenage English)

In (52), the inclusive first person plural semantics of the contracted subject *'s* are retained, but expanded into an overt subject. More strikingly, in (53), *let's* is used with a first person singular subject, showing a dissociation of *'s* from plural inclusive meaning. It is easy to see how this would come about by IG from the 'pseudo-participation' (Ervin-Tripp 1976:48) use of *let's* by caregivers to children, where the identity of the subject *'s* is obscured by the fact that the subject is pragmatically second person singular:

(54) Let's give you some more!

In view of this, I analyse *let's* as a unitary head in Pov. In opposition, Potsdam (1998) argues that *let's* is an inflectional head in I. However, as van der Wurff (2007:56f.) relates, this analysis requires several stipulations: first, that *let's* makes its subject in its specifier appear with accusative case; second, that *don't* is a particle when it co-occurs with *let's*; and third, that *let's* licenses *pro*. By contrast, the analysis here that *let's* is in Pov requires no such additional stipulations. Moreover, *let's* cannot be in I, because it is not mutually exclusive with *don't* in dialectal English:

(55) Don't let's go!

(56) [$_{IP}$ [$_I$ Don't] [$_{PovP}$ [$_{Pov}$ let's] [$_{VP}$ go]]]

Fourth, word order facts in imperatives in Belfast English (BE) (Henry 1995:ch.3,45-80) require a Pov position. BE imperatives have optional subject inversion, argued to be due to verb movement to I, and obligatory object shift of weak pronouns. Shifted weak object pronouns are argued to land beyond the left-edge of VP because they must appear to the left of adverbs:

(57) Henry (1995:72,(202)):

      a. Read it always you to me.

      b. Read you it always to me.

      c. *Read always you it to me.

      d. *Read you always it to me.

In (57b), the subject is above both the adverb and the weak object pronoun, and is therefore beyond VP. It is also below the verb in I, suggesting it is positioned in spec,Pov.

Overall, standard and dialectal English imperatives provide strong evidence for a Pov position.

6.1.2   Elsewhere in English

There is also occasional evidence for Pov in English beyond imperatives. Since this direct evidence is rare, it seems that the abundant imperative input is generalised to other clause-types.

Consider first quantifier float (Sportiche 1988):

(58) (All) the students (all) have [$_{AspP}$ (all) [$_{Asp}$ been]] (all) failing.

The quantifier can be stranded in spec,Asp, an instantiation of κ:point-of-view (Wiltschko 2014), here Pov. Moreover, a floated quantifier can appear in passives, where no overt aspectual marking is involved:

(59) The suspects were all beaten.

Similar to unaccusatives (50-51), passive verbs do not have an external argument-introducing little v*. There must therefore be another position between I and V to host *all*, namely spec,Pov:

(60) [$_{CP}$ [$_{IP}$ The suspects [$_I$ were] [$_{PovP}$ all ~~the suspects~~ [$_{Pov}$ ø] [$_{VP}$ [$_v$ beaten] [$_{QP}$ ~~all the suspects~~]]]]]

Consider in the same regard an example of negative inversion in African American Vernacular English (Sells, Rickford and Wasow 1996):

(61) I know ain't nobody leaving.

As in (51), the combination of negation with an unaccusative verb requires a spec,Pov position between I and V to host the subject:

(62) I know [$_{CP}$ [$_{IP}$ [$_I$ ain't] [$_{PovP}$ nobody [$_{Pov}$ ø] [$_{VP}$ [$_V$ leaving] [$_{DP}$ ~~nobody~~]]]]]

Such evidence for Pov in non-imperative structures in English suggests that acquirers are able to generalise structure from one clause-type to another. The resounding evidence for Pov in imperatives causes the acquirer to posit an analogous position in declaratives, generalising to a clause-type where evidence for Pov is less forthcoming in the input.

6.1.3   Other languages

Beyond English, imperatives and other clause-types provide a good deal of evidence for a Pov position.

Within imperatives, Kayne (1992) accounts for the pre-infinitive clitic position in negative imperatives in Italian by postulating a null modal auxiliary, to which the clitic procliticises:

(63) Non lo fare!

    not   it do-INF

    'Don't do it!'

(64) Non lo-AUX fare

This modal auxiliary seems to be overt in some Italian dialects: *sta* in Paduan and *scé* in Tarantino (Portner and Zanuttini 2003). I claim that this modal auxiliary manifests the Pov position.[38]

In addition, Alcázar and Saltarelli (2014:145f.) present data from doubled Panamanian Spanish imperatives:

(65) ¡oye-ve!

    hear-go

    'Hear!'

-ve is argued to require an additional structural position. For them, this position is their functional prescriptive light v, but I claim it as my Pov position. As Alcázar and Saltarelli (2014:145) note, this analysis could be extended to English double verb imperatives:[39]

---

[38] Cf. Koopman's (2001) analysis of past and imperative forms in Kisongo Maasai (eastern Nilotic), which she argues are introduced by a silent verb 'get'.

(66) [$_{CP}$ [$_{IP}$ [$_{PovP}$ [$_{Pov}$ Go] [$_{vP}$ [$_{VP}$ book it]]]]]]

Beyond imperatives, McCloskey (1996) shows that Irish declarative subjects occupy a position lower than the verb in I, but external to v – for me, Pov. In (67) the subject appears after the verb, but before the sentential adverb (McCloskey 1997:219,(50b)):

(67) Chuala Róisie go minic roimhe   an t-amhrán sin
     heard Róisie often   before-it   that-song
    'Róisie had often heard that song before.'

Finally, McCloskey (1997:216ff.) presents evidence for a subject position between T and v, my Pov, from transitive-expletive constructions. Given that the target of Scandinavian Object Shift is external to VP (Holmberg 1986, Vikner 1991), we find that both the expletive and the associate are outside VP in transitive-expletive constructions in Icelandic (Jonas and Bobaljik 1993, Bobaljik and Jonas 1996):[40]

(68) Það   borðað margir strákar bjúgur   ekki
     there ate   many boys   the-sausages NEG
    'Many boys did not eat the sausages.'

There must therefore be two subject positions between C and V: one for the expletive *það* in spec,I, and a second for the associate *margir strákar* in, on my analysis, spec,Pov.

6.1.4 Summary of evidence for Pov

Overall, imperatives and other structures provide evidence for a structural position Pov, in English and other languages. Acquirers must be able to generalise evidence for the structural position Pov, encountered most readily in imperatives, to other clause-types, where the relevant input data for a Pov position may be scarce.

6.2 The (in)effability of negative imperatives

This subsection aims to draw out the typological predictive consequences of the type 1 and type 2 negative imperative structures sketched in §5.3. I argue that any surface differences between standard clausal negation and negative imperatives highlight to an acquirer important underlying structural aspects of negation and imperatives.

---

[39] Which could be part of a general pattern of grammaticalization of *come* and *go*; see the contributions in Devos and van der Wal (2014).
[40] Cf. Holmberg and Nikanne (1994) for Finnish, and Zwart (1992) for Dutch.

Recall that in a language of type 1 structure, negation is a modifier, while in type 2 languages negation is a head, projecting a label in the clausal spine. We noted in §5.3 that the position of the negative head below I in type 2 languages is much lower than is usually assumed (Poletto 2008). Take, for example, an Italian declarative:

(69) Non ho   letto il   libro.

     not   have.1SG read the book

     'I haven't read the book.'

(70) [$_{CP}$ [$_{NegP}$ Non [$_{IP}$ ho [$_{VP}$ letto il libro]]]]

I argue that the lower position of negation in imperatives is what derives the ban on TNIs in type 2 languages.

The typology of negative imperatives is summarised in van der Auwera and Lejeune's (2013b) WALS chapter on prohibitives.[41] Languages can use their normal imperative with their normal clausal negator (A), their normal imperative with a different negator (B), a different verb form with their normal negator (C), or both a different verb form and a different negator (D):

Table 7: The four-way typology of normal/different negators/verb forms in negative imperatives in van der Auwera and Lejeune's (2013b) WALS survey of 495 languages, with frequencies (and percentages)

		Verb form	
		normal	**different**
Negator	**normal**	113 (23%) (A)	55 (11%) (C)
	different	182 (37%) (B)	145 (29%) (D)

Examples of each language type are:

---

[41] Cf. Sadock and Zwicky (1985:175-7), whose survey of 23 languages found that ¾ had either a special verb form, negator, or both in negative imperatives.

Table 8: Examples of each language type in the four-way typology of negative imperatives, with a summary of their differences from affirmative imperatives

Language		Negative imperative differences		
		Negator		Verb form
Type	Example	normal	different	different
A	German	-	-	-
B	Tagalog	*hindi*	*huwag*	-
C	Spanish[42]	-	-	subjunctive[43]
D	Hebrew	*'al*	*lo'*	future

My central claim is that a special verb form in negative imperatives signals that the verb does not enter into head-to-head relations with I. In addition, while the negator in all languages must be able to value Pov [-coin] in negative imperatives, some languages have a special morphological form for their Pov-valuing negator in imperatives, as opposed to their standard negator elsewhere.

Type A and type B languages therefore have a type 1 clause structure, with a modifying, non-projecting negator. This negator is not a head in the clausal spine, so it does not block head-to-head relations. As such, Agree and head movement can operate just as if the negator was not present. Thus type A languages have their usual imperative verb form and negator in prohibitives. In type B languages, the special negator additionally signals that it can value Pov [-coin].

I claim that language types C and D are of structural type 2. The negator is a head, but is merged lower than the usual position for clausal negation above I; otherwise, it would block the crucial valuing relationship between C and I, leading to semantic incoherence. As a head, the negator blocks head-to-head relations of Agree and movement between the verb and I. The special form of the verb signals that it has not entered into a relationship with I. In Spanish, this can be realised by the subjunctive, where the word order reflects the hierarchy in my type 2 structure:

(71) ¡No (te)   comas   el pan!
      NEG (you.REFL) eat.2.SG.SBJV the bread
      'Don't (you) eat the bread!'

In type D languages, the special form of the negator additionally signals the negator's ability to value [-coin], as well as, perhaps, its unusually low position below I.

---

[42] Worth noting is that the Romance languages, which have spawned the vast literature on the ban on TNIs (cf. §4.7.1), are of the least common type C, highlighting the Euro-centricity of much generative research.
[43] Or, more rarely, the infinitive.

There is, however, a complicating factor: some languages of types C and D will not have clausal negative imperatives, but instead a prohibitive verb plus a gerund. In these languages, the negative imperative is expressed by embedding a gerund structure under an inherently negative or prohibitive verb. For example, Welsh[44] expresses the prohibitive by embedding a gerund under the verb 'stop':

(72) Paid   (â)   mynd yn rhy bell!

    Stop.IMP.2SG (with) going in too far

    'Don't go too far!'

Such structures are essentially affirmative imperatives, with the prohibition expressed lexically. This conflation of languages with different structures into the same groupings in typological surveys highlights the inherent tension between typological and generative work (Baker and McCloskey 2007).

Regarding acquisition, I claim that if a language has a different position for negation in imperatives than in standard clauses, then it overtly signals this different position with a different verb form; such a difference would otherwise be unacquirable.

6.3   Evidence for gerund imperatives

To Johannessen's (2015) data presented in §4.5, I add four pieces of evidence in support of my analysis in §5.5 of prescriptive infinitives and generic imperatives as gerund imperatives: (i) the widespread nature of prescriptive infinitives, beyond Nordic; (ii) the possibility of embedding certain imperatives in some languages; (iii) the overtly nominal nature of generic imperatives in some languages; and (iv) clitic person and placement restrictions in Italian generic imperatives. I finish by considering the usefulness of gerund imperatives as a 'way-in' to nominal(ization) syntax.

First, the bifurcation of the PLD into finite imperatives and prescriptive infinitives extends beyond the Nordic survey presented by Johannessen (2015). Mills (1985:153) reports infinitival imperatives in German as part of "syntactic baby talk" (73-74), which are particularly common in negative commands (75-76):

(73) Jetzt auf-stehen!   (German)   (Mills 1985:160)

    now up stand-INF

    'Stand up!'

---

[44] See the next subsection for a similar example from Korean in (84).

(74) Stehe   jetzt auf!

   stand-2SG.IMP now up

   'Stand up!'

(75) Nicht beiss-en!

   NEG   bite-INF

   'Don't bite!'

(76) Beiss-e   nicht!

   bite-2SG.IMP NEG

   'Don't bite!'

Beyond Indo-European too, Berman (1985:288) observes infinitival imperatives in child input and output in Modern Hebrew, using the general negator *lo'* plus the infinitive, as opposed to the special negator *'al* plus a future verb form:

(77) Lo'   le-cayer al ha kirot   (Modern Hebrew)   (Berman 1985:288)

   NEG INF-draw to the walls

   'Don't draw on the walls!'

(78) 'Al   te-cayer   al ha   kirot

   NEG FUT-draw to the walls

   'Don't draw on the walls!'

Beyond CDS, negative prescriptive infinitives are used as negative generic imperatives in, e.g., Russian, German and Italian:

(79) Pri avarii   razbit'   steklo   (Birjulin and Xrakovskij 2001:44)

   in emergency-PREP.SG break-INF.PFV glass-ACC.SG

   'Break the glass in an emergency.'

(80) Bitte   nicht auf den rasen treten.   (King's College, Cambridge)

   please not   on the lawn tread-INF

   'Please don't walk on the grass.'

(81) Per favore non calpestare il   prato.

   for favour not tread-INF the lawn

   'Please don't walk on the grass.'

Second, gerund imperatives could be invoked to account for the few instances where imperatives can be embedded. The usual reasoning for why imperatives must be matrix clauses is their [imp] in C (see §4.1 and note 17). However, as gerunds that lack C, gerund imperatives should be able to embed. And indeed, embedded neutral imperatives are grammatical in Korean (Sohn 1999:272) (cf. Zanuttini, Pak and Portner 2012:1268):

(82) Yeki tto   o-la   ko hay-yo

　　 here again come-IMP(neutral) QT say-POL[45]

　　 '(They) tell me to come back here again.'

Third, generic imperatives are overtly nominal in some languages. For example, negative prescriptions in English use a gerund *-ing* form:

(83) No walking on the grass!

Stronger evidence still comes from Korean, where negative prescriptive infinitives are marked by a nominalizer as complements of the verb 'stop' (Sohn 1994:350):[46]

(84) canti-ey tul-e   ka-ci   ma-l   kes

　　 lawn-to   enter-INF go-NOMZ stop-PRS fact[47]

　　 'Keep off the grass'

Fourth, person and placement restrictions on clitics in generic imperatives in Italian support my analysis of gerund imperatives by showing that the Pov position is absent. Italian generic imperatives use the infinitive.[48] For example, on bottles one finds:

(85) Non disperdere   nell'   ambiente.

　　 not   disperse.INF in-the environment

　　 'Don't discard in the environment'

Just as with prescriptive infinitives (Table 5), there is a ban on second person in generic imperatives. Subject clitics can appear in generic imperatives, but they must be third person:

(86) Mettersi   nella corsia di destra.   (Maiden and Robustelli 2000:248)

　　 place.INF-3.SG.REFL in-the lane   of right

　　 'Drive on the right.'

---

[45] QT = quotative particle; POL = politeness marker
[46] Cf. the Welsh example (72) in §6.2.
[47] NOM = nominalizer; PRS = prospective suffix
[48] Cf. (81)

(87) *Metterti   nella   corsia di destra.

   place.INF-2.SG.REFL in-the lane   of right

This person restriction contrasts with standard negative imperatives in Italian, where the infinitive can be used with a second person clitic:

(88) Non alzarti!

   not get-up.INF-2.SG.REFL

   'Don't you get up!'

This restriction to third person suggests that generic imperatives can be equated with prescriptive infinitives, and supports my analysis of prescriptive infinitives as gerunds, with third person agreement by default.

In addition to the person restriction, the placement of the clitic is restricted to following the infinitive in Italian generic imperatives:

(89) Non mettersi   nella   corsia di sinistra.

   not place.INF-3.SG.REFL in-the lane   of left

   'Don't drive on the left!'

(90) *Non si   mettere   nella   corsia di sinistra.

   not 3.SG.REFL place-INF in-the lane   of left

Our first thought might be to ascribe this restriction to the general fact that negation blocks clitic climbing in Italian (Zanuttini 1997). However, the same restriction is not found in standard negative imperatives, which allow the clitic to appear on either side of the infinitive:

(91) Non alzarti!

   not get-up.INF-2.SG.REFL

   'Don't you get up!'

(92) Non ti   alzare!

   not 2.SG.REFL get-up.INF

Moreover, clitics are also restricted to appearing after the infinitive in affirmative generic imperatives:

(93) Mettersi   nella   corsia di destra.

      place.INF-3.SG.REFL in-the lane   of right

      'Drive on the right.'

(94) *Si   mettere nella corsia di destra

       3.SG.REFL place.INF in-the lane   of right

Recall from §6.1.3 Kayne's (1992) argument that pre-infinitive clitics in negative imperatives are supported by a null modal auxiliary:

(95) Non ti-AUX alzare

I claim that Kayne's null auxiliary is in Pov. Given this, the clitic placement restriction in generic imperatives is due to their structural difference from standard imperatives. Generic imperatives lack a Pov position, and indeed any clausal structure above vP, because they are gerunds. Therefore, there is no Pov position for clitics to attach to in generic imperatives, accounting for the ungrammaticality of pre-infinitive clitics in generic imperatives.

In sum, the restrictions on clitic person and placement in Italian generic imperatives derive from their reduced clausal structure as gerunds.

Overall, I conclude that prescriptive infinitives and generic imperatives pattern together as gerund imperatives. From an acquisition perspective, gerund imperatives could offer a 'way-in' to nominal(isation) syntax. Regarding nominals, the bifurcation of the PLD into gerund imperatives and finite imperatives could usefully signal the nominal/verbal divide. More specifically, gerund imperatives could offer a 'way-in' to nominalization syntax. Despite their vanishing rarity, speakers of English readily accept sentences with, e.g., gerund subjects:

(96) His dating her will end in tears.

I claim that the acceptability of such sentences is related to the intake of gerund imperatives from the PLD.[49]

---

[49] An outstanding issue is the ban on pronouns in prescriptive infinitives (§4.5). As (96) shows, pronouns are perfectly acceptable in gerunds. I speculate that the claimed ungrammaticality could in fact be a felicity issue, relative to the register of motherese.

6.4 Input generalisation to counterfactuals

In very many languages, acquirers generalise the imperative by IG to realise counterfactuals. This generalisation reflects the structural parallel that I is f-valued from C in both imperatives and counterfactuals.

This application of IG is readily apparent in English. The protasis of counterfactuals in English can be realised by an imperative:

(97) Steal my bike and I will call the police.

      Meaning: If you steal my bike, I will call the police.

This generalisation can be interpreted in the combined spirit of Ritter and Wiltschko (2014) and ReCoS. Ritter and Wiltschko (2014) (§4.3.4) analyse imperatives and counterfactuals as syntactically similar: both involve f-valuation of [$u$coin] by the semantic content of C. On this view, the use of the imperative in counterfactuals is a case of multifunctionality: the imperative can realise either value of f-valuation, [+/-coin]. On the ReCoS-ian view, this multifunctionality would have arisen diachronically from successive generations of children applying IG in acquisition. On intaking an imperative, and analysing it correctly, imperative syntax would be generalised to all instances of f-valuation by IG. Without sufficient evidence to force a retreat from this position, the child maintains imperative syntax as applicable to any instance of f-valuation, including counterfactuals.

This counterfactual conditional use of imperatives is not peculiar to English, but is common across Eurasia (Aikhenvald 210:237). The commonness of this pattern supports the universality of the structure of imperatives and counterfactuals, and the IG acquisition mechanism by which the multifunctionality comes about. Below I present supporting data from German (cf. Platzack and Rosengren 1998:195,fn.27), Tagalog, and Russian (Aikhenvald 2010:237):

(98) Stiehl   mein Fahrrad und ich werde die Polizei rufen!

     steal.IMP my   bike   and I   will   the police call.INF

     'Steal my bike and I will call the police!'

(99) Nakaw-in   mo   ang bike ko   at ta-tawag   ako   ng pulis

     steal-TRANS.NONFIN 2SG.ERG ABS bike 1SG.GEN and ANTIPASS -call 1SG.ABS OBL police[50]

     'Steal my bike and I will call the police!'

---

[50] The glosses follow the ergative analysis of Aldridge (2012).

(100) Pishi   (by)   uchenik, uchitel' ne   delal   by   emu

write:2SG.IMP (COND) pupil   teacher NEG do.PAST.MASC.SG COND to.him

zamechanji

remarks:ACC.PL

'If the pupil had been writing, the teacher would not be making remarks to him.'

Only the singular imperative verb form can be used in conditionals in Russian: *pishi-te* 'write (you plural)' would be ungrammatical even with *ucheniki* 'pupils (plural)' in (100). This departure from Suassurean arbitrariness signals the workings of syntax as a formal system.

Further evidence for IG in this context comes from Cypriot Greek (Christodoulou and Wiltschko 2012), where the subjunctive form *na* marks every instance where external valuation has occurred. The use of *na* thus spans pred-valuation into complement clauses to aspectual verbs [+coin] and future-oriented verbs [-coin], as well as f-valuation in imperatives [+coin] and counterfactuals [-coin]. Compared to the generalisation of the imperative across f-valuation in English (see Table 4), in Greek IG has applied to the max.

Table 9: The range of formal contexts realised by subjunctive *na* in Greek

		coincidence	
		[+ coin]	[-coin]
external valuation	pred-valuation	*na*	
	f-valuation		

The above generalisations are based on the mechanism of valuation, i.e. from f-valuation in imperatives to f-valuation in counterfactuals, or, in the Greek case, over all methods of external valuation. It also seems possible to apply IG from the [+coin] value of imperatives. Imperatives are often recycled into discourse markers (Aikhenvald 2010:246): e.g. Italian *Guarda*, English *Look*, (Irish) English *Listen*. These attention-getting devices could plausibly form part of the C domain, enjoining coincidence [+coin] between the speech act and the addressee.

Overall, the patterns of generalisation between imperatives and other structurally related clauses is symptomatic of the fact the children actively construct formal features in acquisition, and attempt to extend the application of postulated features (FE) as far as possible (IG).

6.5 Particles

Particles can have different meanings in imperatives than in other clause-types. This difference signals to acquirers that different clause-types have different semantic content in C. The same particle interacts differently with each type of semantic content, yielding different meanings.

Aikhenvald (2010:97-9) presents examples of sentential particles whose meaning differs between declaratives and imperatives. In Lao (Tai-Kadai) and German, a particle with a minimal extent meaning in declaratives serves to mark politeness in imperatives. In Lao, the particle *dèèl* attenuates the strength of the proposition in declaratives, but marks politeness in imperatives (Enfield 2007:67):

(101) Jaak5 kham1   mùùt4 dèèl   lèèw4[51]

      tend evening dark   a.little PRF

    'It was already getting a little dark'

(102) qaw3 kùa3 haj5 khòòj5   dèèl

      take   salt   give 1SG.POL IMP.SOFT

    'Please give me the salt.'

Similarly, German *mal*, reduces the specificity of the proposition in declaratives, whereas in imperatives it weakens the strength of the command as a relatively intimate politeness marker:

(103) Ich war mal   in München.

      I   was MAL in Munich

    'I've been to Munich before.'

(104) Gib   mir   mal   das Salz!

      give.IMP me-DAT MAL the salt

    'Pass the salt, would you?'

Adopting a USH perspective, these meaning differences could result from how the same UoL combines with different semantic content in C (=κ:discourse-linking). The particle weakens the assertive force in declaratives, and the directive force in imperatives.[52]

---

[51] Numbers indicate tones.

[52] These particles arguably have a subjective speaker's meaning (cf. note 37). However, my analysis does not involve a projection for the speaker's perspective; instead, the particles attach to and affect the semantic content of C. In a way, everything in a sentence is speaker-subjective: the speaker chooses which words to merge into the sentence, after all.

That the particle relates to C seems straightforward in Lao due to its sentence-peripheral position. This is not so clear in German, where *mal* and other particles (e.g. *doch*, *ja*) appear sentence-internally. However, Struckmeier (2014) convincingly argues that while German modal particles are intrasentential Mittelfeld elements between T and v, they are still licensed by C as periphrastic spellouts of some of C's features. Despite their positioning, therefore, German modal particles still interact with the semantic content of C.

Overall, the contrasting interpretive effects of particles between imperatives and declaratives signal to the acquirer that those clause-types are characterised by different semantic content.

6.6   One step beyond

My analysis of imperatives seems to rely heavily on the backbone of Wiltschko's (2014) USH. However, I do not have to follow the line that this backbone is an innate part of UG.

By USH, UG still has a substantial innate core. UG defines the universal spine (11), and provides the ability to categorise UoLs via the operation Associate (12). From our neo-emergentist perspective, we should consider whether such content really needs to be written into the human genome, or whether it can be derived from elsewhere. In fact, Wiltschko (2014:323) shares this concern, identifying the next step in her research program as investigating whether κs really are primitive elements in the language faculty.

Ramchand and Svenonius (2014) confront this question. Like USH, they seek to mediate between the two prominent positions in generative linguistics: Minimal UG (MUG) (Chomsky 1995 et seq.) (cf. NBH), versus Cartography (Cinque 1999 et seq.) and its UG-given Rich Functional Hierarchies (RFH) (cf. UBH). Ramchand and Svenonius (2014) advance an empirical argument against RFHs, showing with reference to the English auxiliary system that an overly specific innate endowment would overstate the universality of word orders observed in Cartographic research. This argument could be taken as a case study, mirroring Wiltschko's (2014) argument against RFHs in terms of the division between function and substantive content. In addition, Ramchand and Svenonius (2014) attack RFHs from the perspective of evolutionary implausibility: such rich language-specific content could not plausibly have entered the human genome in the order of a hundred thousand years – an evolutionary blink of an eye. Though Wiltschko's USH is substantially less rich, this argument could still be levelled it.

Ramchand and Svenonius (2014) meet this challenge by arguing that hierarchies could emerge in a way highly constrained by domain-general, extralinguistic cognition. The C-T-v hierarchy could

derive from a cognitive proclivity to perceive the world in terms of propositions, situations and events, with each layer built out of the next. These basic ontological semantic notions, universal to humans, provide cognitive biases that are reflected in the universal hierarchical structuring of language.

Of course, Ramchand and Svenonius' (2014) conjecture does not completely resolve the question: it remains to explain, at a level higher up, how (and why) these general cognitive biases are manifested in the human genome. Still, this line of thought suggests it is not madness to suppose that Wiltschko's universal spine may not be UG-given, but could be derived from domain-general cognition. Thus Wiltschko's (2014) universal template for the emergence of syntactic features could itself be emergent.

Wiltschko (2014:317) highlights a further issue regarding UG content, namely the proper characterisation of the coincidence feature. USH does not manage to divorce function entirely from substantive content, because [*u*coin] is a residue of content associated with all instances of κ. Hale (1986:238) argues that the notion of coincidence must be part of UG, as it would be difficult to glean from the data in acquisition – a poverty of the stimulus argument (Chomsky 1965). But following the reasoning in Ramchand and Svenonius (2014), the abstract notion of coincidence could be derived from extra-linguistic cognition.

In the other direction, the syntactic reality of [coin] could be tied more closely to the data. Ritter and Wiltschko (2014:1343) argue that substantive content values [*u*coin] directly. Wiltschko (2014:141) terms this m(orphological)-valuation. Ritter and Wiltschko (2014:1343,fn.20) and Wiltschko (2014:141,fn.31;319-30) note that this differs from the mechanism of valuation standardly assumed in minimalist syntax. In the Probe-Goal Agree system (Chomsky 2000, 2001), an interpretable feature [iF] on the goal is probed by a corresponding uninterpretable feature [uF] on a higher head, establishing an Agree relation. Ritter and Wiltschko claim that their mechanism could be recast in these terms, but there is no reason for doing so. On the contrary, the minimalist concern for a parsimonious theory should methodologically oppose positing an additional valuation operation. Moreover, recasting the valuation mechanism in standard terms affords the PLD a more direct role in establishing the syntax of [*u*coin]. By Ritter and Wiltschko's valuation mechanism, in a tense language I is m-valued [+coin] directly by present tense morphology, a UoL. Present tense morphology has no formal specification for [*u*coin]. Alternatively, present tense morphology could come to be associated with a [+coin] feature, which values I [+coin] by standard Agree. This formalisation in standard terms expresses more directly the interplay between morphology, for which evidence is aplenty in the PLD, and syntactic valuation. Thus, the relation between the PLD

and [*u*coin] can be concretised. This makes the syntax of [*u*coin] more plausibly acquirable, meaning it need not be written into UG.

Overall, both the universal spine and the notion of coincidence do not have to be written into UG, but can be derived from domain-general cognition. Meanwhile, the syntax of [*u*coin] is more plausibly acquirable when recast in standard Probe-Goal Agree terms.

6.7   Summary of further consequences

In this section we have seen some of the consequences of my analysis of imperatives from the perspective of neo-emergentist acquisition. Acquirers apply IG in generalising the structural position Pov from imperatives to other clause-types, and in generalising imperative syntax to other structurally related clause-types, while there is empirical support for my analysis of gerund imperatives. Languages ban TNIs where the negator is forced to be lower than usual in prohibitives, because this difference must be signalled to the acquirer with a different verb form. The different meanings of particles in imperatives than declaratives demonstrate to acquirers that each is typed by different semantic content in C. Lastly, my adoption of a USH framework does not require a rich UG, as the spine and its characteristics can be argued to emerge from domain-general cognition.

7   Conclusion

This thesis has considered imperatives from the perspective of a theory of emergent syntax, as advanced by some ReCoS research. Neo-emergentism aims to reduce the innate content in UG by emphasising instead the roles of a domain-general Minimax acquisition bias and the PLD. My analysis of imperatives accords with these aims in assuming little innate machinery. The overarching structure can arguably be derived from domain-general cognition, and the centrality of the addressee could derive from the acquirer's own centrality in the context of commands from caregivers to children. Furthermore, neo-emergentism is clearly implicated in how acquirers generalise from the information they intake from imperatives in strikingly similar ways across languages.

This thesis has examined the particular structural aspects of imperatives, in contrast with other clause-types. These other clause-types, interrogatives and declaratives, await investigation from a neo-emergentist perspective.[53] So does the syntax of nominals, which were considered here in the

---

[53] See again Biberauer and Roberts (2015:7), who "take the acquirer to be sensitive to particular aspects of PLD such as movement, agreement, etc., readily encountered in simple declaratives, questions and imperatives."

discussion of gerund imperatives. Interesting insights could follow from the issue of how syntactically valued second person features are reconciled with the inherent person features of a nominal in imperatives. In addition, neo-emergentism could offer a fruitful account of the nominal/verbal split, and the parallels between them.

References

Abney, S. P. 1987. *The English Noun Phrase in its sentential aspect*. Ph.D. dissertation, MIT.

Aikhenvald, A. Y. 2010. *Imperatives and commands*. Oxford: Oxford University Press.

Akalu, A. 1985. *Beyond morals? Experiences of living the life of the Ethiopian Nuer*. Lund: Gleerup.

Alcázar, A., and M. Saltarelli. 2014. *The syntax of imperatives*. Cambridge: Cambridge University Press.

Aldridge, E. 2012. Antipassive and ergativity in Tagalog. *Lingua* 122: 192–203.

Auwera, J. van der, N. Dobrushina and V. Goussev. 2013. Imperative-Hortative Systems. In M. Dryer and M. Haspelmath (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/72, Accessed on 2015-06-02.)

Auwera, J. van der, and L. Lejeune (with U. Pappuswamy and V. Goussev). 2013a. The Morphological Imperative. In M. Dryer and M. Haspelmath (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/70, Accessed on 2015-06-02.)

Auwera, J. van der, and L. Lejeune (with V. Goussev). 2013b. The Prohibitive. In M. Dryer and M. Haspelmath (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/71, Accessed on 2015-06-02.)

Baker, C. L. 1979. Syntactic theory and the projection problem. *Linguistic Inquiry* 10: 533–581.

Baker, C. L., and McCarthy, J. J. 1981. *The logical problem of language acquisition*. Cambridge, MA: MIT Press.

Baker, M., and J. McCloskey. 2007. On the relationship of typology to theoretical syntax. *Linguistic Typology* 11(1): 273–284.

Baker, M. 2008. The macroparameter in a microparametric world. In T. Biberauer (ed.), *The Limits of Syntactic Variation*. Amsterdam: John Benjamins. 351-74.

Bavin, E. 1992. The acquisition of Warlpiri. In D. Slobin (ed.) *The cross-linguistic study of language acquisition, vol. 3*. Hillsdale, NJ: Lawrence Erlbaum Associates. 309-371.

Bazalgette, T. 2015. Algorithmic acquisition of focus parameters. Ph.D dissertation. University of Cambridge.

Berman, R. 1985. The acquisition of Hebrew. In D. Slobin (ed.) *The cross-linguistic study of language acquisition, vol. 1: The data*. Hillsdale, NJ: Lawrence Erlbaum Associates. 255-372.

Berwick, R. C. 1985. *The acquisition of syntactic knowledge*. Cambridge, MA: MIT Press.

Bever, T. 1982. Regression in the Service of Development. In T. Bever (ed.), *Regression in Mental Development: Basic Properties and Mechanisms*. Hillsdale, NJ.: Lawrence Erlbaum. 153-88.

Beukema, F. and P. Coopmans. 1989. A Government-Binding perspective on the imperative in English. *Journal of Linguistics* 25:417-36.

Bianchi, V. 2003. On finiteness as logophoric anchoring. In J. Guéron and L. Tasmovski (eds.). *Temps et Point de Vue*/Tense and Point of View. Paris: Université Paris X Nanterre. 213–46.

Biberauer, T. 2013a. *Emergent parameters: hierarchies and phylogenies*. Invited talk presented at the 'Advances in Phylogenetic Linguistics' conference. Ragusa, 17 July.

Biberauer, T. 2013b. *The case for emergent parameters: empirical and conceptual arguments*. Invited talk as part of the University of Kent Leverhulme Seminar Series. Canterbury, 29 July.

Biberauer, T. 2014a. *Acquiring syntactic categories with minimal (but crucial) help from UG*. Talk at the University of Cambridge. 30 June.

Biberauer, T. 2014b. *The universal component of emergent categories*. Paper presented at the Deutsche Gesellschaft für Sprachwissenschaft (DGfS) Annual Conference. Marburg, 6 March.

Biberauer, T. 2014c. *Syntactic and phonological features: is there a worthwhile comparison?* Presented to the MPhil syntax seminar, University of Cambridge. 11 February.

Biberauer, T. 2015a. *The limits of syntactic variation: an emergentist generative perspective*. Invited talk given at the Workshop on Language Variation and Change and Cultural Evolution. Centre for Linguistics History and Diversity, York University. 13 February.

Biberauer, T. 2015b. *From hard-wired to emergent parameters: parameters and acquisition revisited*. Presented to the MPhil syntax seminar, University of Cambridge. 27 January.

Biberauer, T. and I. Roberts. 2009. The return of the subset principle. In P. Crisma and G. Longobardi (eds.), *Historical syntax and linguistic theory.* Oxford: Oxford University Press. 8-74.

Biberauer, T. and I. Roberts. 2012. Towards a parameter hierarchy for auxiliaries: diachronic considerations. In J. Chancharu, X. Hu and M. Mitrović (eds.) *Cambridge Occasional Papers in Linguistics, Vol. 6*. University of Cambridge. 267-294.

Biberauer, T. and I. Roberts. 2014*. Contrastive Hierarchies in Phonology and Syntax: The Role of Third Factors (or: Why Phonology is not Different)*. Unpublished ms: University of Cambridge.

Biberauer, T. and I. Roberts. 2015. Rethinking formal hierarchies: a proposed unification. In J. Chancharu, X. Hu and M. Mitrović (eds.) *Cambridge Occasional Papers in Linguistics, Vol. 7*. University of Cambridge. 1-31.

Biberauer, T. and I. Roberts. to appear a. Parameter setting. In A. Ledgeway and I. Roberts (eds). *The Cambridge Handbook of Historical Syntax.* Cambridge: Cambridge University Press.

Biberauer, T. and I. Roberts. to appear b. Conditional Inversion and types of parametric change. In B. Los and P. de Haan (eds.) *Verb-Second Languages: Essays in honour of Ans van Kemenade.* Amsterdam: John Benjamins.

Biberauer, T., A. Holmberg, I. Roberts and M. Sheehan. 2014*.* Complexity in comparative syntax: the view from modern parametric theory. In F. Newmeyer and L. Preston (eds.) *Measuring Linguistic Complexity.* Oxford: Oxford University Press. 103-127.

Biberauer, T., and S. Vikner. 2015. *Pseudo-coordination in Danish and Afrikaans*. Talk at SyntaxLab, University of Cambridge. 19 May.

Biberauer, T. and H. Zeijlstra. 2012. Negative changes: three factors and the diachrony of Afrikaans negation. In: C. Galves, S. Cyrino, R. Lopes, F. Sândalo and J. Avelar (eds.) *Parameter Theory and dynamics of change*. Oxford: Oxford University Press. 238-264.

Birjulin, L. A., and V. S. Xrakovskij. 2001. Imperative sentences: theoretical problems. In V. S. Xrakovskij (ed.) *Typology of Imperative Constructions*. Munich: Lincom Europa. 3–50.

Bobaljik, J. and D. Jonas. 1996. Subject positions and the roles of TP. *Linguistic Inquiry* 27.2: 195-236.

Boeckx, C. *Elementary syntactic structures*. 2014. Cambridge: Cambridge University Press.

Borer, H. 1984. *Parametric syntax*. Dordrecht: Foris.

Brown R., and C. Hanlon. 1970*.* Derivational complexity and order of acquisition in child speech. In J. R. Hayes (ed.), *Cognition and the development of language*. New York: Wiley. 11–53.

Branigan, P. 2012. *Macroparameter Learnability: An Algonquian Case Study*. Ms. Memorial University. Available at <http://www.ucs.mun.ca/~branigan/branigan-amp.pdf>.

Chomsky, N. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. 1970. Remarks on nominalization. In R. Jacobs and P. Rosenbaum (eds.), *Readings in English Transformational Grammar*. Waltham, MA: Blaisdell. 184-221.

Chomsky, N. 1981. *Lectures on government and binding*. Dordrecht: Foris.

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

Chomsky, N. 2000. Minimalist Inquiries. In R. Martin, D. Michaels & J. Uriagereka (eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, MA: MIT Press. 89-155.

Chomsky, N. 2001. Derivation by Phase. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language*. Cambridge, MA: MIT Press. 1-52.

Chomsky, N. 2004. Beyond explanatory adequacy. In A. Belletti (ed.), *Structures and Beyond*. Oxford: Oxford University Press.

Chomsky, N. 2005. Three factors in language design. *Linguistic Inquiry* 36: 1-22.

Christodolou, C., and M. Wiltschko. 2012. Function without content: Evidence from Greek Subjunctive na. In A.-M. Di Sciullo (ed.), *Towards a Biolinguistic Understanding of Grammar: Essays on Interfaces*. Amsterdam: John Benjamins. 117-40.

Cinque, G. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. Oxford Studies in Comparative Syntax. Oxford: Oxford University Press.

Clancy, P. M. 1985. The acquisition of Japanese. In D. Slobin (ed.), *The cross-linguistic study of language acquisition, vol. 1: The data*. Hillsdale, NJ: Lawrence Erlbaum Associates. 373-524.

Clark, R. 1992. The selection of syntactic knowledge. *Language Acquisition* 2: 83–149.

Clark, R. and I. Roberts. 1993. A computational approach to language, learnability and language change. *Linguistic Inquiry* 24: 299-345.

Clements, N. 2003. Feature economy in sound systems. *Phonology*, 20: 287-333.

Clements, N. 2009. The Role of Features in Phonological Inventories. In E. Raimy and C. E. Cairns (eds.) *Contemporary Views on Architecture and Representations in Phonology*. Cambridge, MA: MIT Press. 19–68.

Crazzolara, Father J. P. 1933. *Outlines of a Nuer grammar*. Vienna: Verlag der internationalen Zeitschrift 'Anthropos'.

Dahl, Ö. and V. Velupillai. 2013. Perfective/Imperfective Aspect. In M. Dryer and M. Haspelmath (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/65, Accessed on 2015-06-03.)

Davies, E. 1986. *The English Imperative*. London: Croom Helm.

Déchaine, R.-M. and M. Wiltschko. 2010. *Interface syntax*. Unpublished ms. University of British Columbia, Vancouver.

Dell, F. 1981. On the learnability of optional phonological rules. *Linguistic Inquiry* 12: 31-37.

Devos, M., and J. van der Wal (eds.). 2014. *'COME' and 'GO' off the Beaten Grammaticalization Path*. Berlin, Boston: De Gruyter Mouton.

Dresher, E. 1998. Child phonology, learnability and phonological theory. In T. Bhatia and W. C. Ritchie (eds.), *Handbook of Language Acquisition*. New York: Academic Press. 299-346.

Dresher, E. 1999. Charting the learning path: cues to parameter setting. *Linguistic Inquiry* 30: 27-67.

Dresher, E. 2003. Contrasts and asymmetries in inventories. In A.-M. di Sciullo (ed.), *Asymmetry in Grammar II: Phonology, Morphology, Acquisition*. Amsterdam: John Benjamins. 239-257.

Dresher, E. 2008. The contrastive hierarchy in phonology. In P. Avery, E. Dresher and K. Rice (eds.), *Contrast in Phonology: Theory, Perception, Acquisition*. Berlin: Mouton de Gruyter. 11-33.

Dresher, E. 2009. *The Contrastive Hierarchy in Phonology*. Cambridge: Cambridge University Press.

Enfield, N. J. 2007. *A grammar of Lao*. Berlin: Mouton de Gruyter.

Evers, A. and J. van Kampen. 2008. Parameter setting and input reduction. In T. Biberauer (ed.), *The Limits of Syntactic Variation.* Amsterdam: John Benjamins. 483-14.

Fodor, J. 1975. *The language of thought*. Sussex: Harvester Press.

Fodor, J. 1998. Unambiguous Triggers. *Linguistic Inquiry* 29: 1-36.

Fodor, J. D. and W. G. Sakas. 2005. The Subset Principle in syntax: Costs of compliance. *Journal of Linguistics* 41.3: 513-69.

Fortescue, M., and L. L. Olsen. 1992. The acquisition of West Greenlandic. In D. Slobin (ed.) *The cross-linguistic study of language acquisition, vol. 3*. Hillsdale, NJ: Lawrence Erlbaum Associates. 111-219.

Frege, G. 1892. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik* 100: 25–50. [Translated as: On Sinn and Bedeutung. In M. Beaney (ed.), 1997, *The Frege Reader*. Blackwell Publishers. 151–171.]

Fukui, N. 1986. *A theory of category projection and its applications*. Ph.D. thesis, MIT. [Revised version published as Fukui, N. 1995. *Theory of projection in syntax*. Stanford, CA: CSLI.]

Gardner, A. 1957. Probability-learning with two and three choices. *The American Journal of Psychology* 70: 174-85.

Gianollo, C., C. Guardiano and G. Longobardi. 2008. Three fundamental issues in parametric linguistics. In T. Biberauer (ed.), *The limits of syntactic variation.* Amsterdam: Benjamins. 109-142.

Gibson, E., and K. Wexler, K. 1994. Triggers. *Linguistic Inquiry* 25: 407-454.

Gillespie, R. 1991. *Manufacturing knowledge : a history of the Hawthorne experiments*. Cambridge: Cambridge University Press.

Gilligan, G. 1987. *A cross linguistic approach to the pro-drop parameter.* Ph.D. dissertation, University of Southern California.

Giorgi, A. 2010. *About the speaker: towards a syntax of indexicality*. Oxford University Press.

Goldberg, A. 1999. The Emergence of the Semantics of Argument Structure Constructions. In B. MacWhinney (ed.), *The Emergence of Language*. Mahwah, NJ: Erlbaum. 197–212.

Greenberg, J. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Greenberg, J. (ed.), Universals of human language. Cambridge, MA: MIT Press. 73-113.

Grice, P. 1989. *Studies in the Way of Words.* Cambridge, MA: Harvard University Press*.*

Guasti, M.-T. 2002. *Language Acquisition*. Cambridge, MA: MIT Press.

Haegeman, Liliane. 2014. West Flemish verb-based discourse markers and the articulation of the speech act layer. Studia Linguistica 68.1: 116-139.

Haegeman, L, and V. Hill. 2013. The syntactization of discourse. In R. Folly, C. Sevdalli, and R. Truswell (eds.), *Syntax and its limits*. Oxford: Oxford University Press. 370-390.

Hale, K. 1986. Notes on world view and semantic categories: Some Walpiri examples. In P. Muysken and H. van Riemsdijk (eds.), *Features and Projections*. Dordrecht: Foris. 233–254.

Hale, M. and C. Reiss. 2003. The subset principle in phonology: why the tabula can't be rasa. *Journal of Linguistics* 39: 219–244.

Han, C.-H. 2001. Force, negation and imperatives. *Linguistic Review* 18(4): 289–325.

Haspelmath, M. 2010. Comparative concepts and descriptive categories in cross-linguistic studies. *Language* 86(3): 663-687.

Hawkins, J. 1994. *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.

Hawkins, J. 2004. *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.

Hegarty, M. 2005. *A feature-based syntax of functional categories: The structure, acquisition and specific language impairment of functional systems*. Berlin: Mouton de Gruyter.

Heim, J., H. Keupdjio , Z. Lam, A. Osa-Gómez and M. Wiltschko. 2014. *How to do things with particles*. to appear in Proceedings of the Canadian Linguistic Association.

Henry, A. 1995. *Belfast English and Standard English*. Oxford University Press.

Hill, V. 2013. *Vocatives: How syntax meets with pragmatics*. Leiden: Brill.

Holmberg, A. 1986. *Word order and syntactic features in the Scandinavian languages and English*. Ph.D dissertation, University of Stockholm.

Holmberg, A. and U. Nikanne. 1994. Expletives and subject positions in Finnish. *NELS 24, Priceedings of the Twenty-Fourth Annual Meeting of the Northeastern Linguistics Society, GLSA*. Department of Linguistics, University of Massachusetts, Amherst. 173-187.

Hornstein, N. 1995. *Logical Form: From GB to Minimalism*. Oxford: Blackwell.

Hornstein, N. 2009. *A theory of syntax: minimal operations and universal grammar*. Cambridge: Cambridge University Press.

Humboldt, W. von. 1829. *Über die Verschiedenheiten des menschlichen Sprachbaues*. In A. Flitner and K. Giel (eds.), 1963, Schriften zur Sprachphilosophie, pp. 144–367. Darmstadt: Wissenschaftlich Buchgesellschaft.

Hyams, N. *1983. The acquisition of parametrized grammars*. Ph.D. dissertation, City University of New York.

Hyams, N. M. 1986. *Language acquisition and the theory of parameters*. Dordrecht: Reidel.

Imedadze, N., and K. Tuite. 1992. The acquisition of Georgian. In D. Slobin (ed.) *The cross-linguistic study of language acquisition, vol. 3*. Hillsdale, NJ: Lawrence Erlbaum Associates. 39-109.

Jackendoff, R. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Jackendoff, R. 1977. *X-Bar Syntax: A Study of Phrase Structure*. Cambridge, MA: MIT Press.

Jackendoff, R. 1990. *Semantic structures*. Cambridge, MA: MIT Press.

Jaspers, D. 2005. *Operators in the Lexicon: on the negative logic of natural language.* LOT Dissertation Series, vol. 117. Utrecht Institute of Linguistics/LOT Netherlands Graduate School in Linguistics.

Jaspers, D. 2012. Logic and colour. *Logica Universalis* 6: 227-248.

Jaspers, D. 2013. *Constraints on concept formation*. Talk given at 36th GLOW Colloquium, Lund, Sweden.

Johannessen, J. B. 2015. Child-directed prescriptive infinitives. Talk at the University of Cambridge. 16 February.

Jonas, D., and J. Bobaljik. 1993. Specs for subjects: the role of TP in Icelandic. In J. Bobaljik and C. Phillips (eds.) *Papers on Case and Agreement I, MIT Working Papers in Linguistics* 18: 59-98.

Kaplan, D. 1989. Demonstratives. In J. Almog , J. Perry and H. Wettstein (eds.), *Themes from Kaplan*. Oxford University Press. 481–563.

Katz, J. J. and P. M. Postal 1964. An Integrated Theory of Linguistic Descriptions. Cambridge, MA: MIT Press.

Kayne, R. 1989. Facets of Romance past participle agreement. In P. Benincà (ed.). *Dialect Variation and the Theory of Grammar*. Dordrecht: Foris. 85–103.

Kayne, R. 1992. Italian negative infinitival imperatives and clitic climbing. In L. Tasmowsky and A. Zribi-Hertz (eds.), *Hommages à Nicolas Ruwet*. Ghent: Communication and Cognition. 300–12.

Kiss, K. 1996. Two subject positions in English. *The Linguistic Review* 13(2): 119-142.

Kohl, K. 1999. *An Analysis of Finite Parameter Learning in Linguistic Space*. MA thesis, MIT.

Koopman, H. *1984*. *The Syntax of Verbs: From Verb Movement Rules in the Kru Languages to Universal Grammar*. Dordrecht: Foris Publications.

Koopman, H. 2001. On the homophony of past tense and imperatives in Kisongo Maasai. In H. Torrence (ed.), *UCLA Working Papers in Linguistics 6, Papers in African Linguistics 1*. 1–13.

Laka, I. 1990. *Negation in Syntax: on the Nature of Functional Categories and Projections*. Ph.D dissertation, MIT.

Lidz, J. and A. Gagliardi. (2015) How Nature Meets Nurture: Universal Grammar and Statistical Learning. *Annual Review of Linguistics* (1): 333-353.

Lightfoot, D. 1991. *How to Set Parameters: Arguments from Language Change.* Cambridge, MA: MIT Press.

Lyons, J. 1977. *Semantics*. Vols. 1–2. Cambridge University Press.

Maiden, M., and C. Robustelli. 2000. *A Reference Grammar of Modern Italian*. UK: Arnold.

Mandler, J. M. 1992. How to build a baby II: Conceptual primitives. *Psychological Review* 99: 587–604.

Manzini, M. R., and K. Wexler. 1987. Parameters, binding theory, and learnability. *Linguistic Inquiry* 18: 413–444.

May, R. 1977. *The Grammar of Quantification*. Ph.D dissertation, MIT.

May, R. 1985. *Logical Form: Its Structure and Derivation*. Cambridge, MA: MIT Press.

McCloskey, J. 1996. Subjects and subject-positions in Irish. In Borsley and Roberts (eds.), *The Syntax of the Celtic Languages, a Comparative Perspective*. Cambridge: Cambridge University Press. 241-283.

McCloskey, J. 1997. Subjecthood and subject positions. In L. Haegeman (ed.), *Elements of grammar: A handbook of generative syntax*. Dordrecht: Kluwer. 197-235.

Mielke, J. 2008. *The emergence of distinctive features.* Oxford: Oxford University Press.

Mills, A. E. 1985. The acquisition of German. In D. Slobin (ed.) *The cross-linguistic study of language acquisition, vol. 1: The data*. Hillsdale, NJ: Lawrence Erlbaum Associates. 141-254.

Miyagawa, S. 2010. *Why Agree? Why Move?* Cambridge, MA: MIT Press.

Mobbs, I. 2015. *Minimalism and the design of the language faculty*. Ph.D dissertation, University of Cambridge.

Newport, E., H. Gleitman and E. Gleitman. 1977. Mother I'd rather do it myself: Some effects and non-effects of maternal speech style. In C. Snow and C. Ferguson (eds.), *Talking to Children: Language Input and Acquisition.* Cambridge: Cambridge University Press. 109-149.

O'Grady, W. 2005. *Syntactic Carpentry: An Emergentist Approach to Syntax*. Mahwah, NJ: Erlbaum.

O'Grady, W. 2008. Innateness, universal grammar, and emergentism. *Lingua* 118: 620-631.

O'Grady, W. 2010. Emergentism. In P. Hogan (ed.) *The Cambridge Encyclopaedia of the Language Sciences*. Cambridge: Cambridge University Press. 274-76.

Ouhalla, J. 1991. *Functional Categories and Parametric Variation*. London: Routledge.

Pinker, S. 1984. *Language Learnability and Language Development*. Cambridge, MA: Harvard University Press.

Platzack, C., and I. Rosengren. 1998. On the subject of imperatives: A minimalist account of the imperative clause. *The Journal of Comparative Germanic Linguistics* 1: 177–224.

Poletto , C. 2008. On negative doubling. *Quadierni di lavoro dell'ASIt 8*. 57-84.

Pollock, J.-Y. 1989. Verb Movement, Universal Grammar, and the Structure of IP, *Linguistic Inquiry* 20: 365-424.

Portner, P. 2004. The semantics of imperatives within a theory of clause types. In K. Watanabe and R. B. Young (eds.), *Proceedings of Semantics and Linguistic Theory 14*. Ithaca, NY: CLC Publications. 235–52.

Portner, P., and R. Zanuttini. 2003. Decomposing imperatives. Paper presented at *IX Giornata di Dialettologia*, Padua. 26 June.

Postma, G., and W. van der Wurff. 2007. How to say *no* and *don't*. In W. van der Wurff (ed.), *Imperatives in Generative Grammar*. Amsterdam: John Benjamins. 205-249.

Potsdam, E. 1998. *Syntactic Issues in the English Imperative*. New York: Garland.

Potsdam, E. 2007. Analysing word order in the English imperative. In W. van der Wurff (ed.), *Imperatives in Generative Grammar*. Amsterdam: John Benjamins. 111–28.

Preminger, O. 2014. *Agreement and its failures. Linguistic Inquiry Monograph 68*. Cambridge, MA: MIT Press.

Pylyshyn, Z. 1973. The role of competence theories in cognitive psychology. *Journal of Psycholinguistic Research* 2: 21–50.

Ramchand, G., and P. Svenonius. 2014. Deriving the functional hierarchy. *Language Sciences* 46: 152-174.

Reichenbach, H. 1947. *Elements of Symbolic Logic*. New York: Macmillan & Co.

Ritter, E., and M. Wiltschko. 2005. Anchoring events to utterances without tense. *Proceedings of the 24th West Coast Conference on Formal Linguistics*. 343–351.

Ritter, E., and M. Wiltschko. 2009. Varieties of infl: tense, location, and person. In J. Craenenbroeck (ed.), *Alternatives to Cartography*. Berlin: Mouton de Gruyter. 153–201.

Ritter, E., and M. Wiltschko. 2014. The composition of INFL: An exploration of tense, tenseless languages and tenseless constructions. *Natural Language and Linguistic Theory* 32: 1331-1386.

Rivero, M.-L. and A. Terzi. 1995. Imperatives, V-Movement and Logical Mood. *Journal of Linguistics* 31: 301-332.

Rizzi, L. 1982. *Issues in Italian syntax.* Dordrecht: Foris.

Rizzi, L. 1990. *Relativized Minimality*. Cambridge, MA: MIT Press.

Roberts, I. 2007. *Diachronic Syntax.* Oxford: Oxford University Press.

Roberts, I. 2012. Macroparameters and Minimalism: A Programme for Comparative Research. In: C. Galves, S. Cyrino, R. Lopes, F. Sândalo and J. Avelar (eds), *Parameter Theory and Linguistic Change*. Oxford: Oxford University Press. 319-334.

Roberts, I. and A. Roussou. 2003. *Syntactic Change: A Minimalist Approach to Grammaticalization.* Cambridge: Cambridge University Press.

Ross, J. R. 1970. On declarative sentences. In R. A. Jacobs and P. S. Rosenbaum (eds.), *Readings in English Transformational Grammar*. Waltham, MS: Ginn and Co. 222–72.

Rupp, L. 1999. *Aspects of the Syntax of English Imperatives*. Ph.D dissertation, University of Essex.

Rupp, L. 2003. *The Syntax of Imperatives in English and Germanic: Word Order Variation in the Minimalist Framework*. Basingstoke: Palgrave Macmillan.

Rupp, L. 2007. 'Inverted' imperatives. In W. van der Wurff (ed.), *Imperatives in Generative Grammar*. Amsterdam: John Benjamins. 297-322.

Sadock, J. and A. Zwicky. 1985. Speech act distinctions in syntax. In T. Shopen (ed.), *Language Typology and Syntactic Description. Vol. I: Clause Structure*. Cambridge: Cambridge University Press. 155–96.

Salustri, M. and N. Hyams. 2006. Looking for the universal core of the RI stage. In V. Torrens and L. Escobar (eds.), *The Acquisition of Syntax in Romance Languages*. Amsterdam: John Benjamins. 159-182.

Saussure, Ferdinand de. 1916. *Cours de linguistique générale*. Paris: Payot.

Schieffelin, B. 1985. The acquisition of Kaluli. In D. Slobin (ed.), *The cross-linguistic study of language acquisition, vol. 1: The data*. Hillsdale, NJ: Lawrence Erlbaum Associates. 525-593.

Schmerling, S. 1982. How imperatives are special, and how they aren't. In R. Schneider, K. Tuite, and R. Chametzky (eds.), *Papers from the Parasession on Nondeclaratives*. University of Chicago: Chicago Linguistics Society. 202-18.

Sells, P., J. Rickford and T. Wasow. 1996. Negative inversion in African American Vernacular English. *Natural Language and Linguistic Theory* 14.3: 591-627.

Sheehan, M. to appear. Parameterising ergativity: an inherent case approach. In J. Coon, D. Massam and L. Travis (eds.), *The Oxford Handbook of Ergativity*. Oxford: Oxford University Press.

Sigurðsson, H. Á. 2004. The syntax of person and speech features. In V. Bianchi and K. Safir (eds.), *The Syntax and Interpretation of Person Features, Rivista di Linguistica* 16(1): 219–51.

Skinner, B. F. 1957. *Verbal behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Sohn, H.-M. 1994. *Korean*. London: Routledge.

Sohn, H.-M. 1999. *The Korean language*. Cambridge: Cambridge University Press.

Sperber, D. and D. Wilson. 1986. *Relevance: Communication and Cognition*. Oxford: Blackwell.

Sperber, D. and D. Wilson. 1995. *Relevance: Communication and Cognition. Second edition*. Oxford: Blackwell.

Speas, M. and C. Tenny. 2003. Configurational properties of point of view roles. In A. DiSciullo (ed.), *Asymmetry in Grammar*. Amsterdam: John Benjamins. 315–43.

Sportiche, D. 1988. A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry* 19: 425-449.

Tomasello, M. 2003. *Constructing a Language: a Usage-based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.

Teleman, U., S. Hellberg and E. Andersson. 1999. *Svenska Akademiens Grammatik*. Stockholm: Norstedts.

Tenny, C. 1987. *Grammatical Aspect and Affectedness*. Ph.D dissertation, MIT.

Travis, L. 2010. *Inner Aspect: The Articulation of VP*. Dordrecht: Springer.

Vikner, S. 1991. Verb movement and the licensing of NP-positions in the Germanic languages. Ph.D dissertation, University of Geneva; revised version, University of Stuttgart.

Weir, M. 1964. Developmental changes in problem-solving strategies. *Psychological Review* 71: 473-90.

Wexler, K. and R. Manzini. 1987. Parameters and learnability in binding theory. In T. Roeper and E. Williams (eds.), *Parameter Setting*. Dordrecht: Reidel. 41-76.

Wierzbicka, A. 1996. Semantics: Primes and Universals. Oxford: Oxford University Press.

Wiltschko, M. 2006. Inlocatives in Halkomelem Salish. In M. Kiyota, J. Thompson, and N. Yamane-Tanaka (eds.), *Papers for the 41st Conference on Salish and Neighbouring Languages, University of British Columbia Working Papers in Linguistics, vol. 18*. 286–310.

Wiltschko, M. 2014. *The Universal Structure of Categories: Towards a Formal Typology*. Cambridge: Cambridge University Press.

Wiltschko, M. 2015. *The (not so silent) syntax of discourse*. Talk given at Cambridge Comparative Syntax (CamCoS) 4, University of Cambridge. 9 May.

Wurff, W. van der. 2007. Imperative clauses in generative grammar: an introduction. In W. van der (ed.), *Imperatives in Generative Grammar*. Amsterdam: John Benjamins. 1–94.

Xrakovskij, V. S. (ed.). 2001. *Typology of Imperative Constructions*. Munich: Lincom Europa.

Yang, C. 2013. *Tipping points*. Paper presented at GLOW 36, Lund.

Zanuttini, R. 1991. *Syntactic Properties of Sentential Negation: A Comparative Study of Romance Languages*. Ph.D dissertation, University of Pennsylvania.

Zanuttini, R. 1994. Speculations on negative imperatives. *Rivista di Linguistica* 6(1): 67–89.

Zanuttini, R. 1996. On the relevance of tense for sentential negation. In A. Belletti and L. Rizzi (eds.), *Parameters and Functional Heads: Essays in Comparative Syntax*. Oxford: Oxford University Press. 181–207.

Zanuttini, R. 1997. *Negation and Clausal Structure: A Comparative Study of Romance Languages*. Oxford: Oxford University Press.

Zanuttini, R. 2008. Encoding the Addressee in the syntax: evidence from English imperative subjects. *Natural Language and Linguistic Theory* 26(1): 185–218.

Zanuttini, R., M. Pak and P. Portner. 2012. A Syntactic Analysis of Interpretive Restrictions on Imperative, Promissive, and Exhortative Subjects. *Natural Language and Linguistic Theory* 30(4): 1231-1274.

Zeijlstra, H. 2004. *Sentential Negation and Negative Concord*. Ph.D dissertation, University of Amsterdam. Utrecht: LOT Publications.

Zeijlstra, H. 2008. On the syntactic flexibility of formal features. In T. Biberauer (ed.), *The Limits of Syntactic Variation*. Amsterdam: Benjamins. 143-174.

Zeijlstra, H. 2013. Not in the first place. *Natural Language and Linguistic Theory* 31: 865-900.

Zhang, S. 1990. *The Status of Imperatives in Theories of Grammar*. Ph.D dissertation, University of Arizona.

Zwicky, A. M. and G. K. Pullum. 1983. Cliticization versus inflection: English n't. *Language* 59: 502–513.