Multiple modal constructions in Rural Chesapeake English and Mandarin Chinese*

Andrew Peters
York University

Multi-modal constructions, (also called ‘modal stacking’) appear in a number of natural languages; most notably in the literature are Southern (or Rural) American English, Scots English, Caribbean Creoles, German, Dutch, and the Chinese languages. Cross-linguistically there is considerable variation in the number of modals allowable in the syntax, from simple double-modal constructions in Southern American English (Labov, 1972; di Paolo, 1989; Hasty, 2012; Ellison, 2007 *inter alia*), to up to four projections in languages like Mandarin Chinese (Lin, 2012). The present study focusses on Southern English and Mandarin to answer long-standing theoretic questions about how multiple-modal constructions are derived, adopting a feature-geometric approach (Harley, 1994; Cowper, 2005 *inter alia*) to Distributed Morphology (Harley & Noyer, 2003). I posit that there are two distinct types of languages which exhibit multiple modality, exemplified by Mandarin and Southern English (respectively): those in which multiple-modality is a product of a recursive Modal Phrase, and those which are produced from a post-syntactic split of the INFL domain.

1 Introduction

1.1 Background

Multi-modal constructions, (also called ‘modal stacking’) of the type seen in Southern (or Rural) American English and Chinese languages have been a subject of curiosity for several decades. Research has focussed primarily around the structural derivations for these modals, as especially in Southern English their appearance runs counter to traditional analyses that posit that modality is a property of a single Tense head. Cross-linguistically among languages that permit multiple modality, there is considerable variation in the number allowable, from simple double-modal constructions in Southern (American) English (Labov, 1972; di Paolo, 1989; Hasty, 2012; Ellison, 2007 *inter alia*), to up to four projections in languages like Mandarin Chinese (Lin, 2012). Examples of these constructions include:

(1) He **might should** go into town.
   ‘He likely will have to go into town.’ [Southern English]

(2) Zhangsan **keneng bixu yuanyi hui** jiang Fayu.
   ‘Zhangsan likely (will) need to be willing to be able to speak French.’ [Mandarin]

The present study will begin with a brief exposition of the particular characteristics of verbal modality in languages like English and Mandarin, discussing their linguistic function and form. We will then discuss

---

*Thanks to all who have offered comments and constructive advice, especially Dr. Alboiu and Gavin Bembridge.

*Toronto Working Papers in Linguistics (TWPL), Volume 36
© 2016 Andrew Peters*
some of the relevant data and previous approaches to the topic. Next I will present the theoretical framework adopted here, as well as the rationale and benefits it will hopefully provide for a theory of not only multiple modality, but modality on the whole. An outline of certain characteristics that unify the behaviour cross-linguistically will provide a background and rationale for investigating Southern English and Mandarin within a single study. We will first examine the Mandarin-style multiple modal derivations, attempting to establish an adequate description of their structure in a fashion that is informative to the typology we are establishing. Finally we will discuss Southern English, the second member of our proposed typology, before formulating some conclusions and theoretical ramifications.

I posit that pursuing a Distributed Morphology (Harley & Noyer, 2003) framework, as well as a feature-geometric approach of Harley (1994), and Cowper and Hall (2007, 2013; Hall, 2001; Cowper, 2003 etc.) will permit us to understand the idiosyncratic behaviour of multiple-modals in Southern English, as well as to understand the differences between it and Mandarin. This approach will allow us to set up a typology of languages which permit stacked modality, divided among two primary types exemplified by Southern English and Mandarin. Furthermore, I argue for two distinct geometries of features involved with modality: one concerned primarily with epistemic modality, and a second involved with root modality.

1.2 The function and form of multiple modals

1.2.1 On the rise & function of verbal modality

First, a comment on the historical trajectory and linguistic motivation for the development of the current system of modality in English, and its parallel in Mandarin, may be fruitful for a better understanding of the semantics of multiple-modality. In short, modality expressed via discrete classes of modal auxiliaries is associated with ‘impoverishment’ of grammatical or morphological reflexes of these semantics. The generalisation may be that the more ‘impoverished’ these grammatical systems are, the more these languages may rely on verbal modality (Roberts, 1985).

The rise of modal verbs in English has been linked with their waxing morphological irregularities throughout the years of development from Old (OE) and Middle English (ME) to Present Day English (PDE). As Denison (1993: 292-337) notes, PDE modals are differentiated from their progenitors by the lack of a finite form, lack of agreement, irregular past tensed forms\(^1\), and their selectional criteria. These features, according to Denison, developed slowly and irregularly over the course of time from OE to PDE. Lightfoot’s (1974) analysis likewise treats OE modals as full verbs which underwent extensive syntactic changes in three stages to become their PDE descendants.

While Lightfoot’s treatment has been considered perhaps the most definitive explanation of the diachronic changes of English modals, an important analysis that takes this as a starting point is Roberts’ (1985) paper on agreement parameters and the same diachronic changes. Roberts posits that the largest changes involved are that modal verbs in PDE assign no theta-roles, and that their rise was conditioned by the decline of morphological expression of mood (namely, the subjunctive). Roberts differentiates between what he calls Morphological Agreement Systems and Syntactic Agreement Systems, the former being common in languages with rich agreement morphology, positing that this morphology involves the base-generation of affixes in INFL. Syntactic Agreement Systems, on the other hand, are the opposite, and occur in languages with impoverished agreement morphology. According to Roberts, English moved to a Syntactic Agreement System from the opposite in OE, and this encouraged the development of a unique class of verbs to handle the same semantic functions which had previously been relegated to morphology (and still are in other European languages especially).

\(^1\)Although, it is not clear that in PDE modals traditionally treated as tensed forms (e.g. might from may, could from can etc.) are still relegated to exclusively past-tense uses.
MULTIPLE MODALS

The important conclusion to draw from Roberts is that the existence of modals as a distinct class in English is due to the fact that PDE is a language relatively impoverished in terms of morphological agreement, and thus must manifest mood and related semantics via individual lexical items. This is the case, as well, in languages like Mandarin, which is notable for its complete lack of tense, agreement and mood inflection (C. Li & Thompson, 1981). That both languages are similar in their impoverished morphological agreement is perhaps a unifying feature. It is then perhaps surprising not that Southern English exhibits multiple-modality, but rather that English in general does not have this property more widely. While this is no doubt a product of the syntactic structure of the English INFL domain, it nonetheless stands that the system of modality in Southern English is not surprising or unexpected.

1.2.2 The employment of multiple-modality

In the literature on double modal constructions in Southern English, the function of these is often discussed as serving the pragmatic purpose of hedging, reducing the impact of potentially face-threatening speech, or otherwise softening commitment (di Paolo, 1989; Mishoe & Montgomery, 1994; Hasty, 2012 etc.). Nevertheless, the employment of double modals adds distinct semantic contrast in comparison with canonical single-modal sentences. Consider the example in (3). The double modal form here provides not simply a meaning of “might drive you”, but involves additional dynamic semantics concerning Dawn’s capacity or volition, rather than simply epistemic possibility. An approximate gloss may be “If it were the case that you asked her to, Dawn would be able to drive you”.

(3) Dawn might could drive you to Selbyville tomorrow, (if you ask).

This is likewise the case in Mandarin, in which the semantic interpretation is compositional, based on further delimited sets of possible worlds\(^2\). As Li R. (2003) notes, modals in continuous succession in Mandarin do not form single grammatical units, but rather get their semantic interpretation by further delimiting upon succeeding and preceding elements individually. That is to say that each modal, as in Southern English, contributes its own semantics.

Examples in (4 & 5) from Lin (2012) demonstrate that Mandarin may have up to four modals, whose interpretation, while compositional, involves the independent interpretation of each modal in succession.

(4) Zhangsan keneng hui nenggou lai.
Zhangsan likely will can come
‘It is likely that Zhangsan will be able to come.’ (Lin, 2012)

(5) Zhangsan yinggai keneng hui nenggou lai.
Zhangsan should likely will can come
‘It should be the case that it is likely that Zhangsan will be able to come.’ (Lin, 2012)

In Mandarin, the employment of multiple modality is not limited solely to politeness, hedging and face-saving functions, but is employed frequently to express complex functions of necessity, possibility, permission, ability, volition etc.

1.3 Previous approaches

Southern English double modals have been diachronically related to those that appear in Scots and some varieties from Northern England. In a (1993) paper by Montgomery & Nagel, the development of these constructions was traced to Scottish immigration into the Southern United States. This unites the

\(^{2}\) Or delimited sets of propositions in the Consciousness-based modal semantics of Hall (2001).
attestation of the double modal constructions across several varieties of English both in North America and Britain, including that of African American English, and possibly even Caribbean varieties.

Perhaps the earliest formal mention of double modal constructions in North-American English was in Labov et al.’s (1968) description of non-standard English in the U.S., followed by a more detailed analysis in Labov’s (1972) account of urban African American vernacular. Labov analysed the first modal as adjectival in nature, relating to a supposed lack of syntactic tense. This study however proposed no structural derivation, and encounters difficulty when faced with evidence showing that adverbs such as "probably" are not in complementary distribution with the higher epistemic modals in Southern English.

Another influential analysis was conducted by Di Paolo (1989), and viewed double modal constructions as single lexical items. Di Paolo calls these "Modal-Modal" compounds, and equates them to idiomatic multi-word lexical items along the lines of similar V-V compounds like "drop-kick". This analysis, however, has been roundly rejected in the subsequent literature, as evidence for the compositional nature of multiple-modal constructions abounds.

Boertien (1986) on the other hand treats both modals as verbal (contra Labov), and as individual constituents (contra Di Paolo), positing two possible structures: one in which both modals project under a single T head, and a second in which the lower modal is a verb, taking the lexical verb as a complement, the higher a T head. This analysis is challenged by facts surrounding subject-auxiliary inversion as well as tag-questions, which suggest that the lower modal is similar to the singular T head in Standard English, while the higher modal resides above T.

Battistella (1995), in fact, was the first to propose that in English the first modal may reside somewhere above T/INFL, based on observations that the first modal seems "tenseless". Data from Subject-Aux inversions which puts the lower modal in the T/INFL head (rather than the higher one à la Boertien) support this claim as well. This theory posited that the higher modal was a modifier of the lower one at an X’ level in the syntax. While this is the progenitor of the present analysis, the X-Bar theory that Battistella’s study was couched in is outdated, and new explanations for the syntactic position of the first modal are required.

For Chinese Modals, authors such as Lin (2012), Li Renzhi (2003), Li & Thompson (1981) and Yip & Rimmington (1997) (among others) break with traditional Chinese grammarians and treat modals either as verbal auxiliaries or modal verbs, differentiating them from lexical verbs by their selection properties, semantics, interaction with negation and aspect, and other characteristics. Recent treatments such as X.-Y. Huang (2000) however, equate Chinese modals with Cinque’s (1999) Functional Projections, noting the ordering constraints in their distribution. A unique doctoral thesis by Li Renzhi (2003) investigates both English and Mandarin modality comparatively, and provides interesting discussion on the topic cross-linguistically.

1.4 Theoretical background

1.4.1 On modality

Before embarking on a discussion of multiple-modality, it is important to be clear on the definition of modality itself, as well as to arrive at a coherent description of the type of modality investigated here, and its functional manifestation in morphology and syntax. While Modality is a broad category applicable to a number of different linguistic purposes, a useful (if conservative) functional definition may be that modality involves the sorts of linguistic structures that permit discussion of non-real, non-immediate or non-actual events or states. The majority of semantic literature on the topic has discussed various classes of modality in terms of their logic concerning possible worlds. This form of modal logic, as outlined by the work of Kratzer (1981, 1991 inter alia), Bach (1989) (among many others), has been useful in delineating various types of modality, including Epistemic, Deontic, Volitional, Dynamic etc..

These models of modal logic are constructed primarily by positing that modality serves to restrict
the set of possible worlds under which a given proposition may be true. Nevertheless, as noted by Palmer (1986) there is an important subjective element to modality, especially with concern to epistemic modality – modality dealing with the speaker’s disposition towards knowledge they have of the world. This is to say that the employment of modality serves to indicate the speaker’s belief, commitment, opinion or disposition towards the proposition they are expressing.

Regardless of which semantic framework modality is discussed in, the most common trend in the literature is to delineate between what are called epistemic and root modalities, the former being associated with the speaker’s knowledge or disposition towards that knowledge, and the latter collecting a wide variety of other miscellaneous meanings.

Various criteria have been proposed to separate the two types, along semantic, syntactic and lexical grounds. Butler (2003) provides an account which marries the semantic proposals of Kratzer (1981, 1991) with syntactic Minimalism (Chomsky, 1995), and provides syntactic evidence that the two types are associated with unique syntactic positions, despite having a single form at PF. These two distinct types of modality, epistemic and root are taken as functionally independent by the present study, and as we will see in section (3.1), their independence may be provided for by distinct featural hierarchies.

The types of modality investigated in the present study are largely based on the typology established by Palmer (1990) in his descriptive account of English modality, with certain modifications gained especially from the recent literature on modality in Chinese languages. Palmer begins by distinguishing three types of modality: epistemic, deontic and dynamic. The distinction between epistemic and deontic modality has been common since Lyons’ (1977) comprehensive work, which posited semantics of speaker knowledge as the domain of epistemic modality, and "the necessity of possibility of acts performed by morally responsible agents" (Lyons, 1977: 823) as the domain of deontic modality. Palmer follows these definitions, however adding dynamic modality (borrowing von Wright’s (1951) terminology) to discuss capability.

The inclusion of dynamic modality is widely supported in the literature on Chinese languages. Li Renzhi (2003) considers ‘ability’ to be a unique class of modals in Mandarin, itself a sub-category of participant-internal modality along with a form of deontic necessity. Likewise, X.-Y. Huang (2009) considers a class of ‘ability’ modals related with the capacity of a subject. This class will be termed dynamic modality throughout the study to remain consistent with terminology employed in the English literature.

Similarly, a class of volitional modality has been suggested, at times to be a sub-form of deontic modality, but at other times in a class of its own. In Mandarin there is a set of modals like yuan yi, ken, gan etc. that express willingness, or daring (gan) of a subject to perform an action. This class is discussed by Li and Huang as well. We will call this form of modality volitional after the term used by Palmer (1990: 134).

Thus, the four main types of modality discussed in the present study are epistemic, deontic, dynamic and volitional modality. While these types are frequently subdivided into differing orientations in the literature, we will see in section (2) that our broad classification is sufficient to capture the distributional observations that are relevant to the investigation of multiple modality.

1.4.2 Syntactic frameworks

The present study takes a version of Distributed Morphology (DM) (Harley & Noyer, 2003) as a framework to the inquiry. Under DM the Late Insertion Hypothesis holds; that is to say that vocabulary items are inserted cyclically at spell-out only (rather than into syntax as in lexicalist approaches), as lexical items and their corresponding phonological information are drawn from an ‘Encyclopedia’ of the language in question at Phonological Form (PF). DM assumes only bundles of features are merged in the narrow syntax, and a new level called "Morphological Structure" (MS) is posited between syntax and PF in which this productive morphological work takes place. Anything occurring in MS may properly be called ‘post-
syntactic’, and there are a number of mechanisms posited in the literature that operate at this level, including Fission, Fusion, Morphological Merger, etc. As will be discussed in section (5.5), the function Fission will be important to the derivation of Southern English double modal structures.

The present study furthermore takes the Morphosyntactic Feature Hierarchy approach of Harley (1994) to organise the formal features in syntax. Harley proposed a hierarchical organisation of Noyer’s (1992) introduction of feature geometries (a staple of theoretical phonology) into syntactic theory. The driving force of a feature-geometric approach, especially a theory grounded in DM, is to (1) formalise subtrees that organise classes of features, and (2) define dependency relations therein. These feature geometries are dependency-structures: a more embedded node entails the de-linking of dominating nodes upon insertion. Furthermore, they are considered to be monovalent – that is, the absence of one feature entails the interpretation of a feature in the immediately dominating node.

This approach has seen growing popularity in the recent syntactic literature to provide systematic formalisation to the features involved in syntactic computation. Cowper (2005; 2003), Hall (2001) and Cowper & Hall (2007; 2013) employ Harley’s feature-geometric approach to posit the following feature geometry in figure (1) for [INFL] features in the syntax.

![Figure 1: INFL feature geometry](image)

I take this approach to the organisation of formal features in syntax to be instructive and important to understanding the relations between the varying semantics of modality as well as multiple-modality. As we will see in the following sections, the syntactic locus of these feature geometries, as well as properties of the system itself (e.g. uniqueness when associated with functional projections) will account neatly for the various types of multiple-modality observed cross-linguistically.

2 Cross-linguistic observations & relevant data

There are some unifying characteristics cross-linguistically to suggest it is useful to analyse multiple-modality as a single related phenomenon. These are primarily concerned with ordering constraints, and uniqueness of modals of each semantic class within a single clause.

2.1 Observation: There are ordering constraints on multiple modals

When multiple modals are present in a sentence, their ordering adheres to a hierarchy similar to the hierarchy Cinque (1999) observed for functional projections and adverbs, simplified in (6) to include just the relevant modals.

(6) \[\text{Mod}_{\text{epistemic}}] > \text{Mod}_{\text{irrealis}} > \text{Mod}_{\text{necessity}} > \text{Mod}_{\text{possibility}} > \text{Mod}_{\text{volitional}}\]

[Simplified subsection of Cinque’s hierarchy]
As discussed in section (1.4.1), the present study limits its scope to discussing a subset of all proposed types of modality. We adapt Cinque’s hierarchy in (7), adding our volitional and dynamic modalities.

(7) \[ \text{Mod}_{\text{Epistemic}} > \text{Mod}_{\text{Deontic}} > \text{Mod}_{\text{Volitional}} / \text{Mod}_{\text{Dynamic}} \]

[Modified modal hierarchy]

In fact, we will see as we approach Mandarin and Southern English individually that the hierarchy in (7) is not only the product of two individual feature geometries, but may not contain all of the same types of modality. There appears to be evidence, for example from Mandarin, that each of these categories may be further subdivided. Epistemicity, for example, is neatly divided into epistemic necessity and possibility, and deontic modality may be divided along internal and external obligation. Treatments such as those of Cinque (1999) have divided modality up in this fashion, although linguists concerned with Chinese modality have introduced categories such as ‘future’ into the inventory of modality for those languages specifically (X.-Y. K. Huang, 2000). English, on the other hand may not have a separate category for volitional modality.

While the list in (7) may not be exhaustive or consider fully the rich gamut of semantics available in all languages, we will see that it is nevertheless sufficiently broad to account for the ordering and uniqueness constraints of stacked modals cross-linguistically. This is to say that the organisation in this fashion is broadly descriptive of ordering constraints observed in natural language: as observed by several Chinese linguistics (X.-Y. K. Huang, 2000; Lin, 2012) as well as researchers concerned with Southern English modality, the order in which modals appear is highly constrained.

In (8) & (9) we are provided with excellent examples from Mandarin:

(8) Ta bixu neng chuxi
He must can present
‘He is required to be present.’
**Deontic > Dynamic** (X.-Y. K. Huang, 2000)

(9) * Ta neng bixu chuxi
He can must present
**Dynamic > Deontic** (X.-Y. K. Huang, 2000)

In example (8), a deontic modal precedes a lower dynamic modal, and the sentence is felicitous. In (9), however, when the order is reversed, the result is ungrammaticality. Southern English as well shows the same restriction, as shown in (10), where a dynamic modal precedes an epistemic one.

(10) * He can might bring the car over.

In section (5) we will see that Southern English *de-facto* adheres to this hierarchy, as it falls out from the limited set of features available for English modal derivations.

2.2 Observation: There is a uniqueness constraint on modals of the same class

A second important observation is that only one modal of each type established in our proposed inventory (epistemic, deontic, volitional, dynamic) may appear per clause. In example (11), two deontic modals are included in a single Mandarin clause, rendering it infelicitous. Likewise in (12), two dynamic modals render the sentence ungrammatical. The same observation holds in Southern English\(^4\), as seen in (13) where two epistemic modals are incompatible in a single clause.

\(^4\)There have been suggestion from some, for example Ellision (2007), who have speculated that in some cases ‘double epistemic’ readings are possible in Southern English. Ellision cites examples from a (1975) dissertation by Coleman such as "John might may go to the party" (Ellison 2007: 58 from Coleman, 1975). However in the opinion of the
(11) * Zhangsan **kending yiding** yao lai.
    Zhangsan surely surely will come

Deontic x 2

(12) * Zhangsan **nenggou hui** xie shi.
    Zhangsan can can write poem

Dynamic x 2

(13) * John **might must** go to work.

Epistemic x 2

From these observations, I will posit a **Uniqueness Constraint** which I argue will fall out from a requirement that the modal features must be unique in the syntax. The syntactic source of this constraint will become apparent when we describe the proposed formal feature-geometry of modality in section (1.4). Essentially, any feature merged into syntax under a feature hierarchy in INFL (and, as we will posit from an additional root modal hierarchy) is required to be unique. This observation is proposed and supported by Cowper (2003) in her discussion of the Tense Aspect and Modality (TAM) system cross-linguistically, wherein the merging of additional INFL feature geometries is only possible in multi-clausal syntax. As we have no evidence to suggest that modality in either English or Mandarin inherently involves a bi-clausal structure, it follows that the Uniqueness Constraint is related directly to the required uniqueness of features from INFL.

2.3 Observation: Double modals are independent, not a single lexical item

Di Paolo (1989) posits that double modals are idiomatic, single lexical items, rather than formed via syntactic derivation. She argues that the somewhat limited set of double-modal pairs available in Southern English (see section (5.1.1)) and their variant acceptability is evidence for their appearance being simply a lexical quirk of certain varieties. Di Paolo conducted ranking tests on acceptability of double-modal constructions with consultants to determine their inventory, preference tests to determine available senses, and naturalistic observation. Di Paolo concluded that as systematic preferences of interpretation were difficult to arrive at in the surveyed population, and furthermore that the tense of the double-modals seemed to behave similarly to single modals, that modal pairs must be single lexical items.

However, as we will show in section (5.1.4), the tense effects di Paolo and others have observed result from the fact that only one of the modals is involved with the Tense head in English. Furthermore, it is argued that the idiosyncratic interpretations di Paolo observed are a product of the highly ambiguous semantics of English modals. Most troubling for di Paolo’s analysis however, is the presence of intervening adverbials and negation between double modals in Southern English, the fact that they may undergo subject-aux inversion but not do so together as a single unit, and the fact that only one of the modals may be used to form tag questions. It has never been proposed that multiple-modality is idiomatic in Chinese languages.

First, double modals in Southern English may produce subject-aux inversion, as noted by Hasty (2012), Ellison (2007), Battistella (1995) and many others. However, only the second modal may be inverted with the subject, as observed in examples (14 & 15). This would not be the expected behaviour should the two be a single lexical unit. Furthermore, as noted by Ellison (2007), cases in which both modals are inverted together are highly marginal. The author in fact finds these completely ungrammatical. It is hypothesised that when acceptable, they may be the...
(14) **Should I might** not go to Jack’s tonight?
(15) * **Might I should** not go to Jack’s tonight?

Furthermore, modals may be split by negation or adverbials:

(16) **He might** probably **could** wash the car.
(17) **He might** not **can** run that fast.

The distributional facts of double modals in Southern English are incompatible with a single-lexical-item analysis. Idiomatic explanations of this sort have not been proposed in the literature on Chinese, so it seems there is no motivation to seriously consider multiple-modality as anything other than the multiple employment of discrete modal items.

3 Proposal
3.1 Two types of modal feature hierarchies

I propose that there are two types of feature hierarchies in natural language which conspire to produce the morphology of modality. The first is the product of the INFL feature geometry as proposed by Cowper (2005). This feature geometry has simply the feature [IRREALIS] as the lowest dependency of a hierarchy including [PROPOSITION], and features concerned with finiteness and deixis. I argue that this feature geometry may only produce epistemic modality, which may be assumed to be a sort of unmarked case, and is highly grammaticalised.

However, I also propose a distinct root modal feature hierarchy in (18) in order to account for various additional semantics exhibited by root modality. The independence of this hierarchy from the formal organisation of INFL features is related to Butler’s (2003) suggestion that while epistemic modality is manifested at or above the T° or I°, deontic/root modality is below this projection (perhaps originating in the vP domain in Butler’s proposal).

(18) **[Dynamic] > [Volitional] > [Deontic] > ([Epistemic])**

[Root modal feature hierarchy]

Note that the order here is the reverse of the hierarchy in (7), modified from Cinque’s (1999) hierarchy. This ordering is determined by dependency relations and the idea that multiple-modals will be merged in independent projections from the bottom up.

With these two modal hierarchies established, along with our uniqueness constraint, we can account for several of the cross-linguistic observations made above. First, as we cyclically merge from bottom to top, if a traditional "higher" modal is selected, it de-links all dominating nodes, preventing higher modals from occurring lower in the structure. If, however, a "lower" modal projects first, so long as the language has a place to merge them, less-embedded modals may merge above it until the hierarchy is exhausted.

3.2 Two types of multi-modal languages

Much as there are two types of featural hierarchies which interact with modality, I propose that there are two types of languages which permit multiple modals, Southern English-style and Mandarin-style. These two languages are differentiated by the way in which they produce modality. English, as discussed by product of late acquisition. An anonymous reviewer of another version of the present study presented the possibility as described by Labov (1972) that these non-standard forms are often acquired idiomatically by non-native speakers of such varieties. 
Andrew Peters

Cowper (2003, 2005, 2012) derives its modality from the manifestation of modal / irrealis features in the INFL domain. I argue that modals in Standard English involve the spell-out of both the modality features proposed by Cowper, as well as the optional specification of additional modal semantics via the projection of features from our root modal hierarchy. Mandarin, however, under this paradigm would involve only the projection of root modality. When we approach the individual derivational strategies of both of these languages, we will see that this approach is in line with Butler’s observation that root modality involves structures below T, while epistemic modality is constrained to remain in an INFL domain.

There are, theoretically, four possible combinations in the way languages may produce modality, the most plausible being the two outlined above. However, a language might be considered which produces modality solely via the INFL feature geometry, which would result in a system which produces only epistemic modality. A less plausible fourth option would be a language that did not overtly manifest modality, in which neither feature geometry is employed. The two paradigms we are discussing however, are hypothesised to be representative of the majority of natural languages, and are here typified by Southern English and Mandarin.

Multiple modality in these two types of languages may then have the following distinguishing characteristics: Mandarin-style multiple modals will result from the recursion of a verbal or modal phrase below T. This phrase projects only root modal features, and is not dependant on an INFL feature geometry as in English. Southern English-style double modals result from the post-syntactic split of the INFL domain. The lower modal in double-modal pairs will initially appear just as in Standard English, bearing features from both the INFL feature geometry, as well as the lower root modal geometry. The higher modal is then be a product of Fission, and may only be epistemic and inherits features from only the INFL feature geometry.

4 Mandarin-style multiple modals

4.1 Characteristics of Mandarin modals

4.1.1 Classification of Mandarin modals

Chinese modals have been categorised in a number of different ways in the traditional literature. Some treat them much like verbs, others as adverbial in nature, but more recently Huang et al. (2009), Lin (2012), Li Renzhi (2003), Li & Thompson (1981) and Yip & Rimmington (1997) among other linguists break with traditional Chinese grammar to posit a distinct class of modal verbs. Li Renzhi (2003) discusses possible criteria for Mandarin modals, which include a ban on reduplication, incompatibility with aspect markers, resistance to passivisation with bei, incompatibility with the attributive de, and resistance to nominalisation etc., but argues that four distinguishing characteristics are the most salient for the category. These are abbreviated as ‘NORA’, after Palmer’s (1965) ‘NICE’ criteria for English modals. NORA stands for:

N Negation with bu

O Occurrence with full verbs

R Reduplication banned

A Aspect markers not taken

Mandarin modals are negated with bu, the generic negation, but not via other negation mechanisms. Furthermore, just as in English, their occurrence is limited to clauses which contain a full verb, and they may not stand alone. Where reduplication is permitted in Mandarin for a number of purposes, including adverbial modification, question formation, etc. (C.-T. J. Huang et al., 2009), modal verbs do not permit
reduplication. Finally, Mandarin modal verbs may not appear with aspeocal markers such as the perfective le, experiential guo etc. modifying them.

4.1.2 On finiteness, and the distribution of Mandarin modals

There has always been an ongoing debate on whether Chinese languages have a finite-nonfinite distinction (generally a property used to differentiate modals from lexical verbs). The positive view of Huang (1984 inter alia) and Tang (2000) among others use evidence from the distribution of Empty Categories (ec’s) and aspeocal markers to posit finiteness. Lin (2012) uses various types of modality in Mandarin to determine the finiteness of clauses, however noting that some modals appear only in supposed finite contexts, and others in non-finite contexts.

The present analysis is agnostic as to whether there is in fact a finite distinction in Mandarin, however the important distributional fact is that Mandarin modals may occur in any clause, finite or not. Keeping the featural hierarchy in figure (1) in mind, this would suggest that Mandarin modals employ solely root modal features and do not interact with INFL modality features. Further semantic tests may reveal that Mandarin modality is always solely root, however the discussion on this topic is not definitive. For discussion and one potential test, see Appendix (A).

4.1.3 The form of multiple-modality in Mandarin

As previously noted, Mandarin may have three or even four modals in a single clause:

(19) Zhangsan keneng hui nenggou lai.
   Zhangsan likely will can come
   ‘It is likely that Zhangsan will be able to come’ (Lin, 2012)

(20) Zhangsan yinggai keneng hui nenggou lai.
   Zhangsan should likely will can come
   ‘It should be the case that it is likely that Zhangsan will be able to come.’ (Lin, 2012)

Lin (2012) in fact notes that, while acceptability generally declines after three modals, it is not necessarily the case that there is an upper bound of modals in Mandarin, and perhaps it is processing difficulty which constrains their multiple employment.

Lin remarks that sentences with two modals occur very frequently in Mandarin speech, unlike their counterparts in Southern English. Mandarin multiple modals are subject to the ordering constraints discussed in section (2.1). Epistemic modals always precede other types, and while authors such as Lin consider as many as five additional types of modality in Mandarin, we nevertheless find that the broad categories established in the present study account for these ordering constraints.

4.2 Deriving the Mandarin multiple modal structure

Under the current theory, multiple modals in Mandarin are derived by recursive merge of modalP (or V in Hall (2001) etc.) which projects a subset of the Lexical Modal Hierarchy. The INFL domain and the VP/ModP complements are shown in figure (2) to demonstrate the proposed derivation for Mandarin multiple modal constructions.

In figure (2), ModII, the lower modal, merges first, projecting a feature from the root modal hierarchy, and all dominating nodes from this geometry. ModI or higher modals, merge above this lower projection, with an additional subset of the root modal hierarchy. So long as the features merged here do not conflict via the Uniqueness Constraint with the set of features merged in the previous modal, the derivation will
converge. In this fashion Mandarin may theoretically merge as many modals as there are unique features for.

5 Southern English-style multiple modals

We have already seen one possible derivational strategy which produces multiple modality in Mandarin. However, Mandarin-style derivation of multiple modals is clearly not feasible for Southern English: it would over-generate, as there is nothing in either our feature geometries or syntax to restrict this derivational strategy to only double modals. Furthermore, the fact that English modals are restricted to finite clauses (and thus are dependant on the finite feature from INFL) alone is enough to rule out a recursive structure below T.

5.1 The nature of English double modals

I will discuss certain properties of English modals which demonstrate not only that the derivational strategy for Mandarin modals does not account for the distribution seen in Southern English, but furthermore that the two modals in Southern English are not multiple merges of the same sort of syntactic element. We will see that the first modal in Southern English is always epistemic, that there is distributional evidence requiring at least one of the Southern English modals to be a T head (as in Standard English), and that the two modals interact with tense in different ways.

5.1.1 Southern English double modal inventory

Hasty (2012) surveys the literature on Southern English double modals and finds that the first modal is always either might, must or may, all epistemic modals. The modal pairs attested in Southern English include:

might could might can might will might should might would
may could may can may will may should
must could must can

In our theoretical paradigm, we can thus assume that first modal must project features from the INFL feature geometry, while the second modal requires features from the root hierarchy. However, we will soon
see that the second modal requires interaction with the INFL domain as well to derive certain tense effects.

5.1.2 T-to-C movement in Southern English

In Southern English, it is the second modal which inverts with the subject to form Yes/No questions. It is ungrammatical to raise the first modal. The examples in (21 & 22) demonstrate that where T to C movement occurs, it is importantly the lower modal that inverts with the subject, and never the first modal. The first modal must importantly not be of the same syntactic class as the lower modal, otherwise it would be attracted to C, or it must otherwise not be present when C probes the INFL domain for a T head.

(21) He might could wash the car.
(22) Could he might wash the car?
(23) *Might he could wash the car?

This evidence rules out several possible derivations proposed in the literature, including di Paolo’s (1989) single-lexical-item hypothesis, and Boertian’s (1986) approach which treated the lower modal as verbal.

5.1.3 Negation with Southern English double modals

In Southern English, negation may occur in one of two positions, with three possible scopes:

(24) He might could not wash his car.
(25) He might not could wash his car.
(26) *He not might could wash his car.

In (24) the negation may either take clausal scope, or constituent scope over the verb. In (25), however, only constituent negation of "could" is permitted. This is evidenced by the fact that NPIs are not licensed in this configuration, seen in (27):

(27) *He might not could ever wash his car.

The clausal negation in (24) however, easily licenses NPIs as in (28):

(28) He might could not ever wash his car.

Hasty (2012) posits that the evidence from (25) rules out Battistella’s (1995) analysis. Battistella had proposed that the lower modal was a proper T head as in Standard English (due to evidence from subject-auxiliary inversion), and that the upper modal was somewhat ‘spurious’ and involved adjunction or modification of the lower T head. Thus, negation as in (25) would require a PolP breaking up the T domain (on the clausal spine). Hasty, however, did not consider the scope of the negation in (25), which suggests that a Neg advP likely adjoins to whatever phrase the second modal heads, and does not live on the clausal spine, thus deriving only the requisite narrow scope.

5.1.4 Tense and double modals

Battistella (1995) and Hasty (2012) rely on an observation that the first modal seems tenseless to support its location above the T head. Hasty examines Sequence of Tense (SoT) effects which suggest that tense is assigned by the lower modal: "John said it might can snow" can only be felicitous for a reading where the event occurs in the future, and shows that Mod\_I does not affect the tense reading.

As is well known in English, sentences like that in (29) may have two interpretations, either in which the event in the subordinate clause is complete at the time of speech, or in which it is not.
John said he might cook dinner

Thus, the sentence may be completed in two ways and remain felicitous:

John said he might cook dinner, and he did.

John said he might cook dinner, but he didn’t yet.

The two examples in (30 & 31) are illustrated graphically in figure (3).

Figure 3: Temporality and tense effects with embedded modals

However, both interpretations are not available for every modal in English, suggesting that some modals are inherently tensed. In (32b.), only a reading in which the event is complete prior to $t_{now}$ is felicitous.

John said he can cook dinner,

a. and he did. 

b. * but he didn’t yet.

Thus, when we employ one of the inherently tensed modals with a modal that does not show these SoT effects, we should be able to determine which is performing temporal deixis. If the modal which is ambiguous is involved with tense and aspect, we should expect both interpretations to be felicitous. If however, only one interpretation is available, we will have evidence that one modal is solely involved with temporal deixis. Consider the examples in (33a. & 33b.), where we see the same results as above with (32b.).

John said he might can cook dinner,

a. and he did. 

b. * but he didn’t yet.

This suggests, as Battistella (1995), Hasty (2012), Labov (1972) and others note, that the upper modal is not involved in temporal deixis. I posit this is further evidence that the lower modal must reside in a position in INFL that is involved with tense.

5.2 Possible structures for English double modals

The literature has proposed several structural configurations for Southern English double modals. We have already falsified di Paolo’s (1989) Single-Lexical Item Hypothesis, but two major options remain.

Hasty (2012) posits that the first modal projects in a ModP which immediately dominates T/INFL, putting it outside the domain of Tense features, with the second modal base-generating in T. Battistella (1995) on the other hand puts the first modal as an adjoined modifier to T, the second base-generated in T. These two approaches are somewhat related, and both capture the requirement that the second modal must be in T in order for Subject-Aux inversion to select it in the following example (repeated from (22) above):
Could he might wash the car?

* Might he could wash the car?

However, Battistella’s analysis was couched in X’ theory; since that time it has come to be accepted that the specifier positions in T are case-assigning positions and reserved for arguments (Chomsky, 1995). Furthermore, Hasty (2012) raises the point that there is little stopping the over-generation of high modals in Battistella’s adjunction theory.

Hasty’s analysis is equally unappealing under our current syntactic paradigm which assumes a distributed morphology whose features are organised by hierarchical feature geometries. Under this paradigm, we have a requirement that the featural semantics of the first modal are dependant on the INFL features of finiteness and proposition, a position strongly defended by Cowper (2005).

5.3 What we know so far: A conundrum

So far, we know several things about Southern English double modal structure, which present something of a problem for a straightforward structural explanation. First, epistemic modals always occur in first position, suggesting that they inherit features from INFL. However, the first modal may not be in the T head as this would interfere with subject-auxiliary inversion. The second modal on the other hand must be in the T head, in order to raise to C in questions, as well as to be involved with temporal deixis as shown by SoT effects. Both the upper and lower modals interact with the INFL feature geometry, as they both are restricted to finite clauses. However, there can be only one INFL feature geometry per clause, as we require a uniqueness constraint.

There are several contradictions produced by these observations which are problematic for any explanation that posits that multiple modality is derived via first merge, in the narrow syntax, into any sort of canonical Standard English-style INFL domain. These issues are furthermore problematic for a cartographic structure as per Cinque (1999). We must therefore posit that some post-syntactic mechanism produces the structures we see in Southern English double modality.

5.4 The first modal is the product of post-syntactic operations

A preliminary approach considered the possibility that Mod₁ was merged in the syntax in a split INFL domain as is seen in other Germanic languages (Wurmbrand, 1998). This explanation is rejected however because of the ungrammaticality of tag questions employing the first modal:

(36) You might should read Syntactic Structures, shouldn’t you?
(37) *You might will take a philosophy class, mightn’t you?
(38) *You may will want to refrigerate the leftovers, may won’t you?

That the first modal is unavailable for use in a tag structure either by itself or with the lower modal suggests that it is not present in the narrow syntax. Battistella confirms this observation with judgements from other informants, supporting his classification of the first modal as "spurious" (Battistella, 1995). What Battistella had observed as the spurious nature of the first modal may in fact be indicative of its post-syntactic origins.

5.5 Deriving the English double modal structure

Knowing that the first modal is likely a product of post-syntactic operations, the exact mechanism of its origin then remains an open question. It may either be that the item is inserted on whole post-syntactically, or that it is produced via an operation on a previously existing item in the syntax. The latter approach is
favoured, as it does not involve the post-syntactic insertion of an additional functional category or layer. Furthermore, there is precedent in the literature for post-syntactic operations which involve elements merged in the narrow syntax.

Under DM, a function called fission exists, posited by Noyer (1992) and discussed in detail by Halle (1996). Fission involves the insertion of more than one vocabulary item for a single exponent, by splitting the terminal node. If we assume that the INFL domain may be split post-syntactically between a lower T head, and an upper Mood head, Southern English INFL domains may now resemble those of German, as Wurmbrand (1998) posited for multiple-modal structures there. The result of Fission thus produces a Mod° and a T°, splitting the INFL domain.

Our approach solves several of the structural problems presented above. Mod_I will not intervene with T-to-C movement required for subject-aux inversion because it is not of the type attracted by C due to lack of T features (cf. Roberts, 2001). Furthermore, we are not required to merge Mod_I as an adjunct to T/INFL as per Battistella, while still maintaining the fact that the first modal is structurally above T°, but still within INFL. We can account for the fact that with one single INFL feature hierarchy, we now have two individual heads which inherit its features at different times in the derivation.

The upper modal will always be epistemic, as Fission leaves it with only the grammatical modality features. This is perhaps one of the largest points gained by the present hypothesis, as it neatly accounts for the limited inventory of double modals in Southern English. Furthermore, tense features are not transferred to Mod_I, capturing the observation that it might be "tenseless". These features will remain on the proper T head. Finally, we cannot over-generate Mod_I as in Battistella’s structure, because Fission operates only on the limited set of features in syntax.

I present a possible structure in figure (4), including our expanded INFL domain with the split T° and Mod°, and a possible position for an adverbial adjunct, as well as a lower location for negation.

Figure 4: Southern English double modal derivation
6 Conclusions

Accounting for the occurrence of multiple-modality cross-linguistically, I have proposed that there are two modal feature hierarchies in natural language, one grammatical, dependant on the features of INFL, and a second set of root modality features. These two hierarchies are represented in figure (5).

![Figure 5: Two feature hierarchies (*: Recursive)](image)

Furthermore, I have proposed that there there are two types of languages which permit multiple modals, typified by Mandarin-style languages which project a recursive modal verb phrase, expressing only root modality, and Southern English-style languages which involve Fission of an otherwise syncretic I°. This typology should be sufficiently descriptively-adequate to subsume all natural languages which permit the multiple employment of syntactic modality.

However, many questions remain for future work, most importantly the examination of other languages that permit multiple modality, such as Caribbean creoles (Durrleman-Tame, 2008), German (Wurmbrand, 1998) etc. Furthermore, the question remains as to whether a more detailed hierarchy for root modality may be important or necessary, especially in light of data from Mandarin.

References


Hasty, J. D. (2012). We might should oughta take a second look at this: A syntactic re-analysis of double modals in Southern United States English. Lingua, 122, 1716-1738.


MULTIPLE MODALS


Appendix

A Potential semantic test for Mandarin root modals

Tests from Hall (2001), used to differentiate epistemic vs. root Modality may suggest that even epistemic modals in Mandarin pattern like root modals. Epistemic modals are seen as taking a proposition as their complement, while root modals do not. The sentences in (39) and (40) with epistemic modals are both truth-conditionally synonymous and the modal semantics apply to the same proposition, even with passivisation. However, in (41) and (42) containing deontic (root) modals, the truth values are the same, but the desiderata is realised by Mary in (41), while the agent responsible for realising the desiderata in (42) is Alexander.

(39) Jack might cut the carrots.
(40) The carrots might be cut by Jack.
(41) Mary shouldn’t confuse Alexander.
(42) Alexander shouldn’t be confused by Mary.

Might Mandarin epistemic modals display a similar asymmetry, where the modals in (43) and (44) are calculated on two different propositions?

(43) Zhangsan keneng piping le Meili
   Zhangsan might criticise ASP.PERF Meili
   “Zhangsan might have criticised Meili”

(44) Meili keneng bei Zhangsan piping le.
   Meili might by Zhangsan criticise ASP.PERF
   “Meili might have been criticised by Zhangsan.”