Relative Clauses that are in Tongan: Exploring syntax and prosody within the Tongan DP*

Byron Ahn (byron-[æt]-ucla-[dat]-edu)
UCLA Department of Linguistics

Abstract

Tongan, a Polynesian language, is almost entirely head-initial. There are some exceptions to this: Demonstratives and the so-called Definitive Accent (Churchward 1953). Also post-nominal in Tongan are relative clauses, and they can be variably ordered with respect to the Definitive Accent, but not Demonstratives. To derive his host of facts, I argue for a promotion analysis of relative clauses (Schachter 1973, Vergnaud 1974, Kayne 1994), and invoke three crosslinguistically supported movement operations. Moreover, each of these structures directly feeds the prosodic component, in which three Optimality Theory style constraints (Prince and Smolensky 1993) determine the available prosodic phrasings. Even when the movement that distinguishes these structures is string-vacuous, it still affects the possible prosodic phrasings, supporting a strong syntax-prosody connection.

1 Introduction

Tongan has post-nominal relative clauses that exhibit multiple word orders with regard to the Definitive Accent (DEFACC), a morpheme which Churchward (1953) defines as the "stressing of the final vowel for the sake of definiteness, of greater definiteness":[1]

(1) a. te u 'aka 'a e tangata -ná -a [na'e 'uma kia Mele 'aneafi]
FUT 1.SG kick ABS the man -DEM -DEFACC [PST kiss DAT Mary yesterday]

b. te u 'aka 'a e tangata -na [na'e 'uma kia Mele 'aneafí] -i
FUT 1.SG kick ABS the man -DEM [PST kiss DAT Mary yesterday] -DEFACC
'I will kick that man who kissed Mary yesterday'

Given this word order variability, two questions immediately arise. How can we explain these multiple word orders? And, do they correspond to different formal properties?

*I would like to thank the native speaker consultants, Piula Tonga and Saia Moala. Special thanks also go to Hilda Koopman and Kie Zuraw, for their clear guidance on this project. I would also like to thank all of the UCLA 2010 Field Methods course members, as well as Laura Kalin, Robyn Orfitelli, Norvin Richards, Matt Tucker, and the participants and organizers of AFLA XVIII, for their input and suggestions. Any remaining errors are of course my own.

This paper is an extended version of a proceedings paper for AFLA XVIII.

[1] Abbreviations used in this paper follow the Leipzig glossing conventions, with the exceptions of the following two: DEFACC: definitive accent; KO: pan-Polynesian predicate marker (Potsdam and Polinsky In Press).

In the spirit of Cinque 2005, Leu 2008 and Zamparelli 1995, I argue that there are multiple determiner projections in the DP-domain, and that they are hierarchically rank-ordered as (2):[2,3]

(2) *'e/'a* [**Case**] ≫ (*h*)*e* [**High D**] ≫ *ni/na* [**Demonstrative**] ≫ DEFACC [**Low D**] ≫ NP[4]

Following an analysis in which relative clauses are CPs introduced by the complement of the lowest D-head (the DEFACC, in the case of Tongan), I show that the positional variability of the relative clause arises from the interactions of three independently-motivated movement operations. The first of these is relative clause promotion (Schachter 1973, Vergnaud 1974, Kayne 1994, *inter alia*), the second is an optional movement of the relative clause itself (Kayne 1994, 2005), and the third is movement of the NP to a higher position within the DP.

The movements which derive the word order variability have observable effects on the prosodic phrasing of relative clauses. Under an OT-style constraint-based approach, only three rank-ordered constraints, typical of syntax-prosody interface work (as in Selkirk 1996, Truckenbrodt 1995, 1999, *inter alia*), are necessary to predict seven attested prosodic patterns with relative clauses, while also ruling out a number of unattested patterns.

The rest of this paper proceeds as follows. First, in Section 2, I introduce some of the functional elements in the DP, and I argue for a syntactic analysis of the word-order facts within the Tongan DP. Next, Section 3 introduces the question of the structural position of relative clauses in Tongan, and I provide a syntactic analysis. With an understanding of relative-clause syntax, Section 4 shows that the prosodic phrasing is directly fed by the syntactic structure. Finally, I present open questions in Section 5 and conclude in Section 6.

2 Functional Elements of the Tongan DP

Tongan, like other Polynesian languages, is typically a head-initial language: Ds precede NPs, Case precedes DPs, the language uses prepositions, and so on. However, certain functional categories appear to be head-final. For example, the demonstratives (henceforth Dems) *-ni* and *-na* are phrasal enclitics, obligatorily following the NP, including attributive adjectives (if there are any):

(3) a. 'oku lele 'a e kumaa 'i he [$_{NP}$ fale (fo'ou)] **-ni**
PRS run ABS the mouse LOC the [house (new)] -DEM
'The mouse is running in this (new) house.'
b. *'oku lele 'a e kumaa 'i he [$_{NP}$ fale **-ni** fo'ou]
PRS run ABS the mouse LOC the [house DEM new]

Note that Dem *-ni* co-occurs with the definite determiner (*h*)*e*, implicating that the two do not head the same XP. We will return to this shortly. In addition to the Dem, DEFACC is also a head-final morpheme that previous literature has treated as marking definiteness/specificity/uniqueness.[5]

[2]There is likely to be more functional material than is explicit in the hierarchy of (2).

[3]The definite determiner in Tongan has two morphologically conditioned allomorphs: *he* and *e*.

[4]Here, and throughout this paper, I use "NP" as a cover term that envelops a range of structure that may include adjectives, reduced relative clauses, (and perhaps more,) as well as the N's arguments and the N itself.

[5]The exact semantic contribution of the DEFACC is of some debate. See, for example, Churchward 1953, Chung 1978, Hendrick 2005 and Abner and Burnett 2010. Abner and Burnett's semantic analysis is briefly discussed in §2.1.

Before discussing the DEFACC as it relates to the syntactic structure, we must first have a basic understanding of Tongan stress.

Word-level primary stress is calculated based on right-aligned trochees – the prosodic word's primary stress falls on its penultimate vowel. However, stress appears to fall on its final vowel in contexts with DEFACC, so DEFACC has been treated as a stress-shift process (Churchward 1953):[6]

(4) a. he fàle fo'óu
the house new
'the new house'

b. he fàle fo'oú
the house new.**DEFACC**
'the new house'

Under this sort of analysis, the final vowels of (4a) and (4b) should be of similar lengths (with the exception of whatever effect on length stress has).

Closer investigation has shown that the length of vowels with DEFACC is like that of long vowels (which occur phonemically elsewhere in the language), and thus DEFACC is not a stress shift *process*, but a moraic vowel enclitic (Taumoefolau 2002, Anderson and Otsuka 2006, White 2010). This moraic vowel (which will be abbreviated -μ) gets its phonological feature values from the final vowel of the word it cliticizes to. Given this analysis, a more accurate representation of the DEFACC would be:

(5) a. he [$_{NP}$ falé] -e
the [house] -DEFACC
'the house'

b. he [$_{NP}$ fàle fò'oú] -u
the [house new] -**DEFACC**
'the new house'

As a phrasal enclitic, the DEFACC "shifts" the stress of whatever word is the at the right edge of the NP, by adding a mora to a prosodic word. This causes the final vowel of the NP – the [e] of *fale* in (5a), and the [u] of *fo'ou* in (5b) – to become the penultimate vowel of its prosodic word. This allows even words with the DEFACC to conform to the generalization that stress is always trochaic in Tongan.

In much the same way, Dem is also an enclitic that causes "stress-shift":

(6) a. he falé -e
the house -DEFACC
'the house'

b. he falé -ni
the house -DEM
'this house'

However, it cannot be that the DEFACC and the Dems *-ni* and *-na* are all heads of the same functional category: the DEFACC and a Dem can co-occur. When they do, the Dem obligatorily precedes DEFACC:

(7) a. he fàle fò'ou -ní -i
the house new -DEM -DEFACC
'this new house'

b.*he fàle fò'ou -ú -ni
the house new -DEFACC -DEM
Intended: 'this new house'

This strongly implicates syntactic structure as mediating these word orders, especially as these are *phrasal* enclitics.

[6]Throughout this paper, I use acute accents to indicate word-level primary stress, and grave accents to indicate secondary stress. These acute accents should not be confused with the Tongan orthographic representation the DEFACC.

2.1 Multiple Functional Layers of the DP

In an example like (7a), there appear to be three independent heads that would be classified as a D-like: (*h*)*e*, *-ni* and *-μ*. The first major component of my analysis is that (*h*)*e* and *-μ* in fact *are* both Ds – (*h*)*e* is a HighD and *-μ* is a LowD. Additionally, *-ni* is of category Dem which can co-occur with these Ds. There is cross-linguistic support for the idea of multiple D heads within a single "DP".[7] For example, many languages (e.g., Greek, Javanese, Welsh) express determiners and demonstratives in the same phrase (Leu 2008):[8]

(8) **afto to** vivlio (Greek)
this the book
'this book'

Additionally, Swedish marks certain DPs with two morphemes, each of which is associated with a distinct interpretation (LaCara 2011):

(9) **den** gamla häst **-en** (Swedish)
DEF old horse **-DEF**
'the old horse'

Similarly, other languages have two exponents in *demonstratives*, each with a different contribution to the interpretation. For example, French has a free word Dem, and an NP-enclitic Dem (Bernstein 1997):[9]

(10) **ce** livre jaune **-ci** (French)
DEM book yellow **-DEM**
'this (here) yellow book'

In each of these cases, both D-like morphemes make unique contributions to the interpretation, supporting the idea that they are each realizations of a distinct head. If they are distinct heads, we need multiple DP functional projections. If it is possible to have multiple DP projections in a single DP, what rules out English **this the book*, for example? Zamparelli (1995:126) proposes the following constraint on the usage of multiple determiners to explain the distribution of multiple Ds: "two determiners are possible only when each one adds something to the meaning of the other."[10] By this logic, if we are to believe (*h*)*e* and the DEFACC to each head their own DPs, we expect Tongan (*h*)*e* and *-μ* to have different semantic contributions.

Abner and Burnett (2010) reach this very conclusion, arguing that the DEFACC "anchor[s] the interpretation of the [DP] to the context of utterance." For that reason, the DEFACC is excluded in

[7] With an analysis whereby a DP has multiple D-like projections, a question might arise of what I mean by "DP". I mean this to refer to all D-projections, which I take to be sister of KP.

[8] Each of these languages behaves differently with their usage of these multiple Ds – for example, Swedish only uses two Ds under certain circumstances, e.g. when there is an adjective. Neither of the Tongan Ds, on the other hand, depend on modification of any kind, as exemplified in (5).

[9] English has a very similar pattern, with *this here yellow book*, which is formally distinct from *this yellow book here* (Bernstein 1997:91).

[10] Alternatively, it is possible that a string like *this the book* is ruled out for English because *this* is bimorphemic, and contains a definite D morpheme *th-*, which cannot occur twice in a DP (*this the book*) for structural reasons (e.g. *th-* instantiates a unique head in the DP structure).

cases like (11), because the speaker believes that devils don't exist.[11]

(11) ko Piúla, 'óku túli 'a e [tèevólo 'okú ne túi 'óku 'i tu'a] (*-a)
KO Piula, PRS chase ABS the [devil PRS 3.SG believe PRS LOC outside] (*DEFACC)
'Piula, she is chasing the devil that she believes is outside (but there is no devil).'

Moreover, the (*h*)*e* HighD *can* appear in (11), regardless of anyone's belief-state, providing support that DEFACC is not just a second realization of a single D.

Given these facts, it must be that the DEFACC is a head of a distinct functional projection in the DP, apart from the HighD (*h*)*e*.

2.2 A Syntactic Account of Word Order

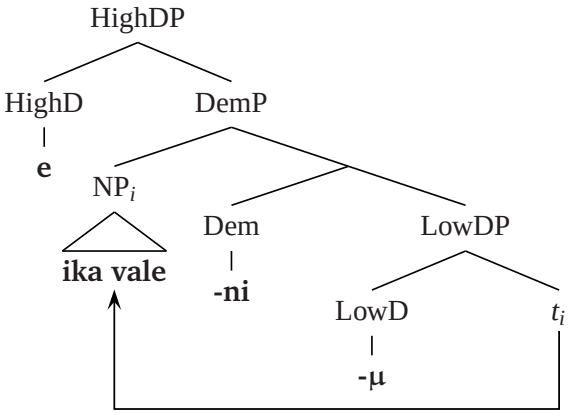
As in the sentential domain, variable word orders within the DP ought to be derived from the same underlying constituency. I pursue an analysis in the vein of Cinque 2005 (among others), in which movements applied to a universal hierarchy like (12) derive a given language's word order:

(12) Case (KP) ≫ Determiner (HighD) ≫ Demonstrative (Dem) ≫ Determiner (LowD) ≫ NP

Since Tongan NPs occur between HighD and Dem, it must be the case that there is movement. Specifically, I argue that the head-final enclitics here are derived by phrasal movement of the NP, as in (14), consistent with an Antisymmetric approach to syntax (Kayne 1994):[12,13]

(14) e íka vàle -ní -i
the fish stupid -DEM -DEFACC
'this stupid fish'

(15)



[11] Notably, DEFACC would be acceptable in (11) if the speaker believed there to be a devil outside.

[12] One might want to propose a left-branching structure whereby the DemP and LowDP are head-final. Under such an account, Dem would need to be lower than HighD and LowD, for Dem to be an NP enclitic:

(13) [$_{\text{HighD}}$ he [$_{\text{LowDP}}$ [$_{\text{DemP}}$ [$_{\text{NP}}$ ika vale] -ni] -μ]]

However, this would go against the findings in Ishizuka 2007, which finds evidence for Dem≫D, based on data from Javanese. Moreover, this would require directionality parameters for each XP; see, for example, Kayne (2010) for arguments against these directionality parameters.

[13] For ease of exposition, I represent this NP movement as a single movement from the complement of LowD to the Specifier of DemP. It is likely theoretically desirable that such movement is impossible, and the NP must instead 'stop in' the Specifier of the LowDP 'on its way' to the DemP (Koopman and Szabolcsi 2000, Cinque 2005).

Since nothing is able to intervene between the NP and the Dem, it is reasonable to assume that the NP moves to Spec,DemP because NP has an extremely local relationship with the Dem.[14] Such an NP-movement operation within the DP has been independently motivated in, for example, French. Consider the derivation for *ce … -ci/là* in (17) (Bernstein 1997):

(17) $[_{DemP}$ ce $[_{DemP}$ $[_{NP}$ livre jaune $]$ -ci t_{NP} $]$ $]$ (French)
DEM book yellow -DEM
'this (here) yellow book'

2.3 Against a "Morphological" Account

In spite of these motivations for a syntactic analysis, it may seem to some that what I refer to as a LowD, DEFACC (and perhaps the affixal Swedish D and/or French Dem), is optionally inserted by some morpho-phonological process which is a reflex of being in the context of what I call the HighD, (*h*)*e*. However, since the DEFACC makes its own contribution to the interpretation, it would need to be present at LF, and must not be inserted anywhere in PF (where morphological insertion processes are thought to occur; Embick and Noyer 2001). Thus, in order to contribute to the meaning and have a pronounced form, it must be that DEFACC is a head in the narrow syntax.

Moreover, if its placement as a head were achieved by a post-syntactic morphological operation, it would seem that the appropriate candidates for such an operation would be Lowering or Local Dislocation, as defined in Embick and Noyer 2001. However, as a *phrasal* enclitic, the DEFACC's placement would be problematic for each of these operations. Lowering has been typically defined targeting *morphological heads* as the landing site for movement – not syntactic phrases. Since DEFACC cliticizes to NPs, and not Ns or As, a Lowering account would seem untenable. On the other hand, Local Dislocation – which applies after spell-out – would not be provided with the necessary information about syntactic phrases to be able to have the DEFACC cliticize to the NP.[15] As a result, a morphological approach to the DEFACC in a post-syntactic domain would fail to capture key structural facts.

[14] Alternatively, the DEFACC may be higher than the HighD. If so, the constituency would need to be as follows:

(16) $[_{HighD}$ $[_{DemP}$ $[_{LowD}$ he ika vale $]$ -ni $]$ $[_{HighD'}$ -i t_{DemP} $]$ $]$

While this may work for simple cases like (16), this would require a far more complex structure to account for word orders with relative clauses. See Appendix B.

[15] An analysis involving Local Dislocation might be successful if we make the appropriate assumptions about spell-out domains. Namely, if we assume that DEFACC's phrasal host is a spelled-out phrase which the DEFACC immediately precedes at linearization, Local Dislocation might be able to produce the correct ordering, along the lines of Kramer 2010. However, assuming that a spelled-out phrase is complete in terms of stress calculation (e.g. Kratzer and Selkirk 2007), such a solution is problematic in that the DEFACC would be trying to shift the stress of an immutable, spelled-out phrase. In other words, the location of primary stress in the DEFACC's host must be determined *after* the DEFACC has cliticized to it, in order to achieve the "stress shift" phenomena seen in (5).

3 Syntactic Properties of Tongan Relative Clauses

3.1 Where are Relative Clauses?

As we have already seen, relative clauses (henceforth RCs) in Tongan are post-nominal. Having established a clear idea of the basic structure of the Tongan DP and its post-nominal functional material, consider the data in (18), which are representative of the available word orders for RCs:

(18) a. ’oku ma’a ’a e sote (-na) (-a) [na’a ku foo]
PRS clean ABS the shirt (-DEM) (-DEFACC) [PST 1.SG wash]
‘That/the shirt that I washed is clean.’
b. ’oku ma’a ’a e sote (-na) [na’a ku foo] (-o)
PRS clean ABS the shirt (-DEM) [PST 1.SG wash] (-DEFACC)
c. *’oku ma’a ’a e sote [na’a ku foo] -na
PRS clean ABS the shirt [PST 1.SG wash] -DEM
d. *’oku ma’a ’a e sote [na’a ku foo] -na -a
PRS clean ABS the shirt [PST 1.SG wash] -DEM -DEFACC

As we saw in (3), Dems obligatorily follow adjectives; on the other hand, (18) shows that Dems obligatorily precede RCs. This indicates that RCs and adjectives are not in the same syntactic relationship with the NP, contrary to NP-adjunct approaches to RCs (e.g., Ross 1967). If the RC were an NP adjunct, we would predict (19b) to be grammatical in the same way as (19a):

(19) a. [$_{\text{HighDP}}$ e [$_{\text{DemP}}$ [$_{\text{NP}}$ sóte ’ulí] -na t_{NP}]]
the shirt dirty -DEM
‘that dirty shirt’
b. *[$_{\text{HighDP}}$ e [$_{\text{DemP}}$ [$_{\text{NP}}$ sóte na’á ku foó] -na t_{NP}]]
the shirt PST 1.SG wash -DEM
‘that shirt that I washed’

In fact, RCs and adjectives have completely different distributions with regard to the Dem and DEFACC:[16]

(21)

	Adjective	Relative Clause
a. N ___ Dem DEFACC	✓	*
b. N Dem ___ DEFACC	*	✓
c. N Dem DEFACC ___	*	✓

This strongly argues against an NP-adjunct analysis of RCs.

Moreover, as argued in Chung 1978, we can conclude that RCs in Tongan are indeed a constituent within the DP, since they can appear between the noun and the DEFACC, ruling out DP

[16] Note that data like (20) would seem to indicate that reduced relative clauses *can* appear between the N and the Dem:

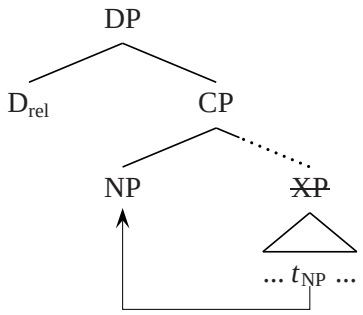
(20) he [**ta’u** [kuo ’osi]] **-na**
DET **year** PERF finish **-DEM**
‘**That year** (which is) just finished.’ (Lit. *that year having finished*)

adjunction for cases like (21b). Moreover, systematic investigation has revealed that word order has no correlation to interpretation (e.g. restrictivity).[17] For this reason, the Tongan RC must always originate within the DP, even when it appears to be outside of it, as in (21c). My analysis therefore relies on a different theory of RCs, which predicts these behaviors: the promotion analysis of RCs.

3.2 A Promotion Analysis of Relative Clauses

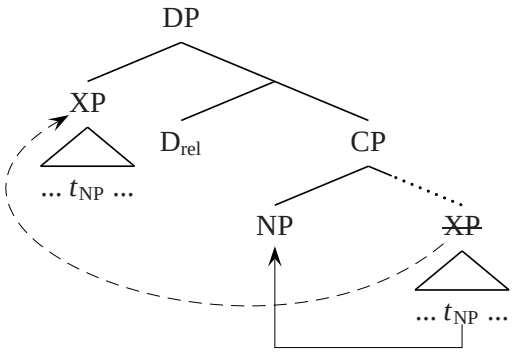
Under the promotion analysis of relative clauses, a relative clause is a CP introduced by a relativizer D. The relativized NP, which is base-generated in its argument position within the CP and undergoes movement to the CP edge (Schachter 1973, Vergnaud 1974, Kayne 1994, among many others). As a result, post-nominal RC languages (e.g., English) look like (22):

(22)



Thus, under an Antisymmetric approach, languages with pre-nominal RCs (e.g., Japanese) involve an additional movement to front the RC is necessary. This movement fronts a sub-constituent of the CP (for arguments, see e.g. Kayne 1994, Kornfilt 2000, Kayne 2005, Ishizuka 2008), which I label XP.[18] In this way, a pre-nominal RC is derived as in (23):

(23)



[17] To test for restrictivity, native speakers were given, for example, the following situations in Tongan, and asked to translate the underlined sentence from English '*Manu received a postcard from her friend living in Samoa and several postcards from her relatives living in Hawaii. She lost the postcard that her friend living in Samoa sent.*' and '*Yesterday Manu received a postcard. She lost the postcard, which her friend living in Samoa sent.*' There was no correlation found between restrictivity and word order or prosodic possibilities.

[18] In these previous analyses, XP is assumed to be TP/IP, though nothing seems to crucially rely on this (only that Tense/Infl is inside the pre-nominal RCs). I do not use the TP/IP label, in order to avoid a commitment to the location of the Tongan Tense/Aspect/Mood morpheme.

3.3 Syntactic Derivations

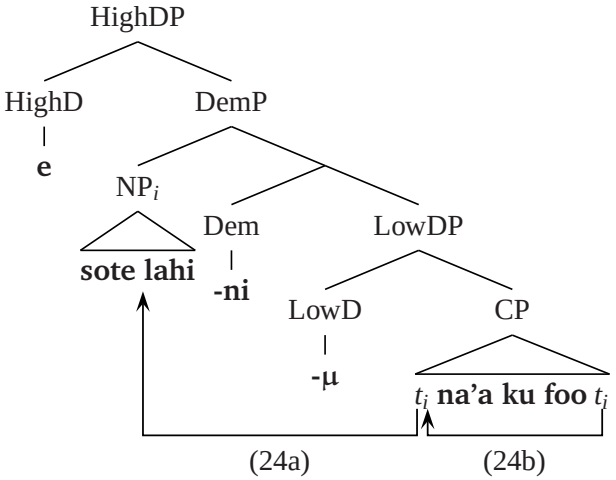
At this point we have seen three DP-internal movements, laid out in (24):

(24) a. **NP fronting:** deriving French *ce...ci*, as in (17)
b. **NP Relativization:** at the heart of the promotion analysis of RCs, as in (22)
c. **RC fronting:** deriving Japanese-like relative clauses, as in (23)

Assuming that Tongan DPs always involve (24a) and that all RCs use (24b), we predict straightforwardly the word order and structure in ((25)-(26)), in which the relativization feeds the NP fronting:[19]

(25) e sóte làhi -ní -i na'á ku fóo
the shirt large -DEM -DEFACC PST 1.SG wash
'this large shirt that I washed'

(26)



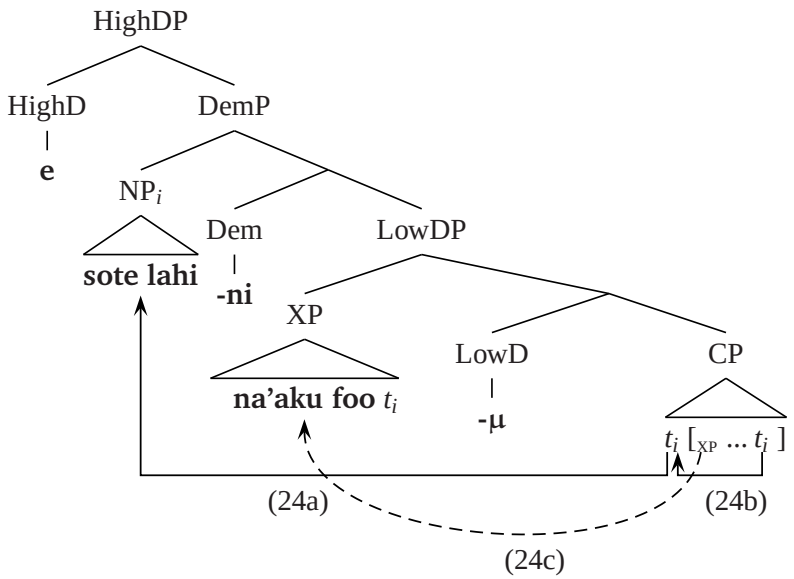
Additionally, if the RC fronting that occurs in some languages is optional in Tongan,[20] we derive (27) – a minimal pair with (25) – with the structure and movements in (28):

(27) e sóte lahí -ni na'á ku foó -o
the shirt large -DEM PST 1.SG wash -DEFACC
'this large shirt that I washed'

[19] Strikingly similarly, in French, the NP must front to the pre-Dem position, out of the RC (Bernstein 1997).

[20] I have not found any evidence that this movement has any interpretational consequences, even though such a consequence would be desirable.

(28)



In this way, NPs move to Spec,DemP for the same reason that RCs cannot intervene between NP and the Dem: there is simply no space for the RC to move to. An NP (not a CP) occupies Spec,DemP just like it does when there is no RC.

Thus, based on what has been independently motivated for RCs (crosslinguistically) and for DPs (in Tongan), we straightforwardly derive the (im)possibility of the RC word-orders in (18).

4 Prosodic Breaks

This analysis involving two syntactic derivations finds extra evidence in the distribution of strong Intonation Phrase (IP) level phonological breaks[21] that sometimes separate the NP and the RC. Importantly, there is no relationship between the necessity of a strong prosodic break and a restrictive/non-restrictive interpretation (unlike languages such as English). The relevant range of data is given in the table below, which also indicates that the different phrasings correspond to the two derivations we have seen:

(29)

	Prosodic Phrasing	Derivation in (26)	Derivation in (28)
a.	[$_{IP}$ 'oku ma'a 'ae soté e] [$_{IP}$ na'a ku fóo]	✓	
b.	[$_{IP}$ 'oku ma'a 'ae sòte ní i] [$_{IP}$ na'a ku fóo]	✓	
c.	[$_{IP}$ 'oku ma'a 'ae sóte] [$_{IP}$ na'a ku fóo]	✓	
d.	[$_{IP}$ 'oku ma'a 'ae sóte na'a ku fóo]		✓
e.	[$_{IP}$ 'oku ma'a 'ae sóte] [$_{IP}$ na'a ku foó o]		✓
f.	[$_{IP}$ 'oku ma'a 'ae soté ni] [$_{IP}$ na'a ku foó o]		✓
g.	[$_{IP}$ 'oku ma'a 'ae soté ni] [$_{IP}$ na'a ku fóo]	✓	

[21] Native speaker consultants referred to this kind of break as a 'comma', as opposed to a 'full stop', saying that they consider these utterances to be one sentence and not two. However, the only way found to distinguish the two was native speaker intuition, as the 'comma' does not seem to be measurably different from a 'full stop' – that is, both inter- and intra-sentence prosodic groupings (above the Accentual Phrase, which is irrelevant here) may involve a long pause, an entire pitch reset, and final lengthening (Vicenik and Kuo 2010).

4.1 A Constraint-Based Analysis

To derive the phrasings in table above, I assume three Optimality Theory-style constraints (Prince and Smolensky 1993). Using only these three constraints, the appropriate phrasing can be determined by providing the appropriate structure from (26) and (28) as input:

(30) a. **AFFIXSUPPORT** An affix must not be prosodically separated from its morpho-phonological host.
b. **ALIGN(IP,L;CP,L)** Align the left edge of an Intonation Phrase (IP) to the left edge of a CP.
c. **ALIGN(IP,L;LowDP*,L)** Align the left edge of an Intonation Phrase (IP) to the left edge of a lexically filled LowDP.

AFFIXSUPPORT (as defined in Richards 2010) crucially outranks the latter of the two ALIGN (in the spirit of McCarthy and Prince 1993, Prince and Smolensky 1993, Truckenbrodt 1995, Selkirk 1996, *inter alia*) constraints; and the ALIGN constraint for the CP crucially outranks the ALIGN constraint for the LowDP, evidence for which we will see shortly:

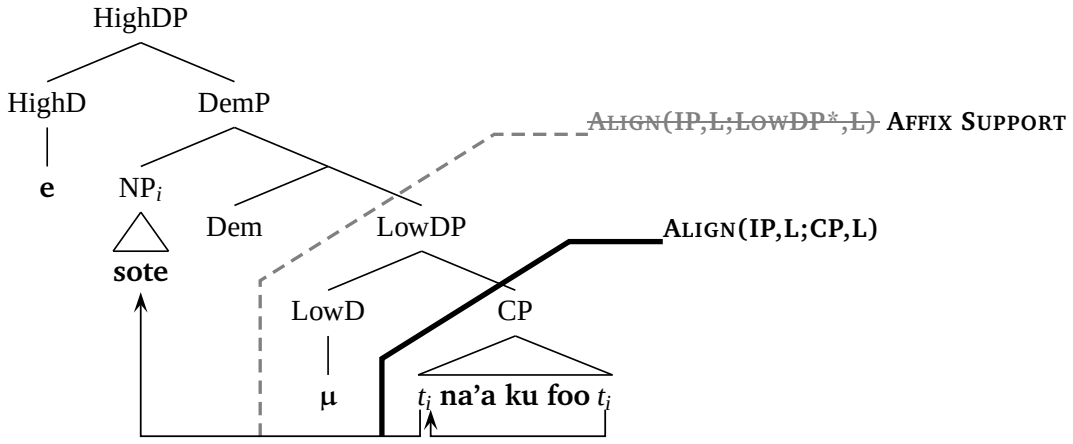
(31) AFFIXSUPPORT ≫ ALIGN(IP,L;CP,L) ≫ ALIGN(IP,L;LowDP*,L)

Using some given structure as input, this system dictates how to prosodically phrase the utterance. Consider the sentence in (29a), which must be in two Intonation Phrases:

(32) a. [$_{IP}$ 'oku ma'a 'a e sote e] [$_{IP}$ na'a ku foo]
PRES clean ABS the shirt -DEFACC PAST 1.SG wash
'The shirt that I washed is clean.'
b.* [$_{IP}$ 'oku ma'a 'a e sote e na'a ku foo]
c.* [$_{IP}$ 'oku ma'a 'a e sote] [$_{IP}$ e na'a ku foo]

To derive the word order in (32), we need a structure like (26), in which the RC has stayed within the CP. To avoid a violation of ALIGN(IP,L;CP,L), a prosodic break just before the RC is required. To avoid a violation of ALIGN(IP,L;LOWDP*,L), a prosodic break between DEFACC and NP would be required (since LowDP is lexically headed by μ). However, this would violate the more highly ranked AFFIXSUPPORT, and as such, no break is inserted between DEFACC and NP:

(33)



The phrasing in (32a) is the optimal phrasing, as shown in the tableau below (which also demonstrates the constraint ranking in (31)):

(34)

		AFF.SUPP.	ALIGN-CP	ALIGN-LOWDEM
a.	☞ phrasing in (32a)			*
b.	phrasing in (32b)		*!	*
c.	phrasing in (32c)	*!		

The prosodic derivation for (29b), which must also be in two IPs in the same way, proceeds identically with the what we have seen here.

4.2 Accounting for Other Data

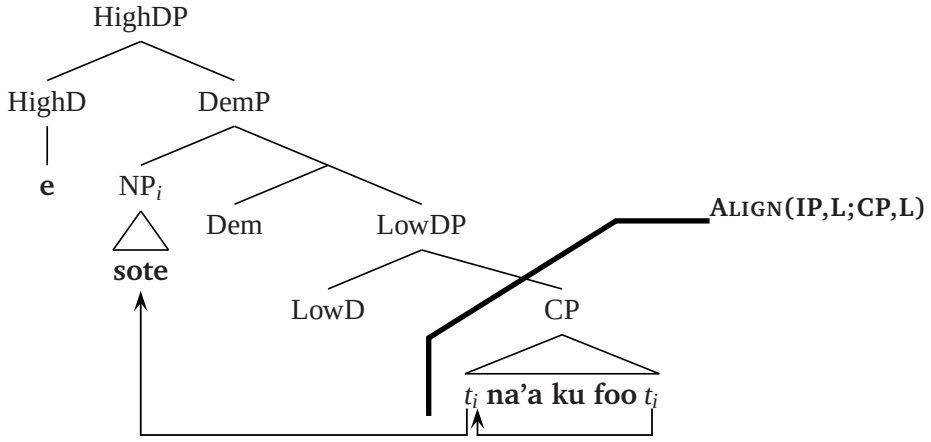
Let us also look at derivations for the prosodic minimal pair in (35):

(35) a. [$_{IP}$ 'oku ma'a 'a e sote] [$_{IP}$ na'a ku foo]
PRES clean ABS the shirt PAST 1.SG wash

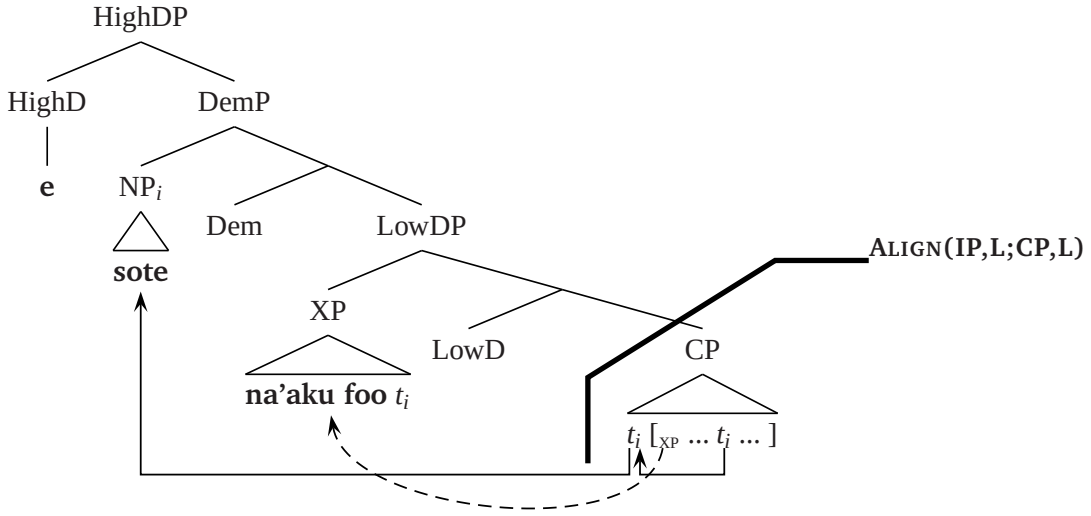
b. [$_{IP}$ 'oku ma'a 'a e sote na'a ku foo]
PRES clean ABS the shirt PAST 1.SG wash

This prosodic difference is accounted for directly by the two different structures we have seen:

(36) a.



b.



In both cases, ALIGN(IP,L;CP,L) correlates to a prosodic break just before the CP – but only in (36a) does this make a difference, since there is no pronounced material following the break that gets inserted in (36b). Though (36a) and (b) might seem indistinguishable since the additional movement in (36b) is string-vacuous, there is an *empirically measurable effect* on the prosody as a result. This adds clear support to both the syntactic and prosodic analyses promoted here.

The prosodic derivation for (29g) proceeds identically with (36a), and prosodic derivations for (29e-f) proceed identically with (36b). Therefore, these two structures and three rank-ordered constraints straightforwardly account for the phrasing possibilities in (29).

Under this prosodic analysis, the possibilities in (29) are accounted for by having the prosodic component take two different syntactic structures – which are necessary to account for word-order data – as input. This analysis also rules out several unattested phrasings, such as (32b) and (32c), among others.[22] As such, both the syntactic and prosodic structures in the Tongan DP are directly related – a strongly desirable result under modern approaches to the syntax-prosody interface.

5 Open Questions

5.1 Spell-Out and Metrical Stress

There is a theoretical problem with the DEFACC's ability to "shift stress". Assuming that metrical stress is calculated when spell-out occurs (e.g., Kratzer and Selkirk 2007), it seems that the DEFACC would have to be within the same spell-out domain as its host. While this is often rather straightforward, RCs present a vexing case:

(37) mé'a [$_{CP}$ ná'e ínu 'e [$_{DP}$ Sìoné]] -e
thing PAST drink ERG John -DEFACC
'thing that John drank'

The DEFACC enclitic shifts the stress of '*Sione*', after the DP and CP (and perhaps other) phases have been sent to PF, and have been spelled-out with metrical structure calculated. Yet, the Definitive Accent is somehow still able to manipulate the previously calculated stress. It must thus be the case that either *Sione* and the DEFACC are indeed in the same spell-out domain – thus requiring major revisions to this analysis – or stress can be modified after spell-out, "counter-cyclically". The syntactic proposal made here would require the latter being true, raising the question of when counter-cyclic operations are allowed to apply.

5.2 Multiple Definitive Accents

According to native speaker informants, a DP like (38), in which multiple DEFACCs occur with only one obvious NP, is possible. However, it should be noted that such a DP was never produced without direct elicitation.

(38) e soté **-e** na'a ku foó **-o**
the shirt **-DEFACC** PST 1.SG wash **-DEFACC**
'the shirt that I washed'

[22] There remain some apparent issues with this analysis. See Appendix A.

This is unpredicted under this analysis, unless it is possible for a DEFACC to be realized in the DP out of which the NP has relativized:

(39) $[_{DemP}\ [_{NP}\ sote_i\ ]\ [_{LowDP}\ \text{-}\mu\ [_{CP}\ t_i$ na'a ku foo $[_{LowD}\ \text{-}\mu\ t_i]]]]$

In this way, the DEFACC can appear multiple times within a single noun's functional structure, without any obvious second DP for a DEFACC. Similarly, this analysis resolves the theoretical issue of how to attach a DEFACC to a word at the right edge of a RC – the DEFACC is actually found within the RC.

However, this analysis seems to create more problems than it solves. First and foremost, if the DP remnant is the source of the DEFACC on the RC, we have to explain why other DP functional material besides the DEFACC (Determiners, Case, Demonstrative enclitics) are obligatorily unpronounced. The lack of Demonstrative enclitics is especially problematic since, as DP functional enclitics, they seem to be most similar to DEFACCs.

Furthermore, there would not be any reason to rule out data like (40), which differs from (38) in that there is additionally a Dem (also cf. (18a)):

(40) *e sòte -ná **-a** [na'á ku foó] **-o**
D shirt -DEM **-DEFACC** [PAST 1.SG wash] **-DEFACC**
Intended: 'That shirt that I washed'

This alternative approach has more issues than advantages, and is taken to be inferior to the analysis promoted elsewhere in this paper. We are thus left to wonder how to account for data like (38), to the extent that they are truly grammatical in natural Tongan speech.

6 Conclusion

In this paper, I have demonstrated that the Tongan DP contains multiple functional projections, including HighDP, DemP, and LowDP. Though Tongan relative clauses exhibit two possible word orders with regard to the Definitive Accent LowD, I have shown this variation can be entirely predicted by independently motivated movements on a single underlying structure. Crucially, the data cannot be derived under an adjunct analysis of relative clauses – such an analysis would incorrectly predict adjectives and relative clauses to have the same distribution.

Moreover, this structural analysis of word order facts leads directly to an account of prosodic phrasings for relative clauses in Tongan, whereby the syntactic structures directly feed the prosody. Though linear word order at times conceals the two surface constituencies (as in (36)), each structure maps onto different prosody, directly manifesting the syntax-prosody interface.

Finally, the syntactic and prosodic data lead to an underlying hierarchical structure within the DP, namely: HighD≫Dem≫LowD. This will have implications for our approach to DPs that are multiply marked for definiteness, and prompts a second look at the DP structure in languages with multiple overt Ds, such as Swedish or Greek.

References

Abner, Natasha, and Heather Burnett. 2010. On the so-called “determiners” of Tongan. UCLA.

Anderson, Victoria, and Yuko Otsuka. 2006. The phonetics and phonology of "defnitive accent" in Tongan. *Oceanic Linguistics* 45:21–42.

Bernstein, Judy B. 1997. Demonstratives and reinforcers in Romance and Germanic languages. *Lingua* 102:87–113.

Chung, Sandra. 1978. *Case marking and grammatical relations in Polynesian*. University of Texas Press.

Churchward, C. Maxwell. 1953. *Tongan grammar*. Oxford: Oxford University Press.

Cinque, Guglielmo. 2005. Deriving Greenberg’s universal 20 and its exceptions. *Linguistic Inquiry* 36:315–332.

Embick, David, and Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32:555–595.

Hendrick, Randall. 2005. Resumptive and bound variable pronouns in tongan. In *Proceedings of the Twelfth Annual Conference of the Austronesian Formal Linguistics Association*, ed. Jeffrey Heinz and Dimitris Ntelitheos.

Ishizuka, Tomoko. 2007. Internal structure of the DP in Javanese. Paper presented at the 14th Austronesian Formal Linguistic Association.

Ishizuka, Tomoko. 2008. Restrictive and non-restrictive relative clauses in Japanese: Antisymmetric approach. http://ling.auf.net/lingBuzz/000808.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Kayne, Richard. 2005. Antisymmetry and Japanese. In *Movement and silence*, chapter 9. Oxford: Oxford University Press.

Kayne, Richard. 2010. Why are there no directionality parameters? Presented at WCCFL XXVIII, University of Southern California.

Koopman, Hilda, and Anna Szabolcsi. 2000. *Verbal complexes*, volume 34 of *Current Studies in Linguistics*. Cambridge, MA: MIT Press.

Kornfilt, Jaklin. 2000. Locating relative agreement in Turkish and Turkic. In *Studies on Turkish and Turkic languages*, ed. Aslı Göksel and Celia Kerslake, 189–196. Harrassowitz.

Kramer, Ruth. 2010. The Amharic definite marker and the syntax-morphology interface. *Syntax* 13:196–240.

Kratzer, Angelika, and Elisabeth Selkirk. 2007. Phase theory and prosodic spellout: The case of verbs. *The Linguistic Review* 24:93–105.

LaCara, Nicholas. 2011. A definite problem: The morphosyntax of double definiteness in swedish. In *Morphology at Santa Cruz: Papers in Honor of Jorge Hankamer*, ed. Nicholas LaCara, Anie Thompson, and Matthew A. Tucker, volume 55–83.

Leu, Thomas. 2008. The internal syntax of determiners. Doctoral Dissertation, NYU.

McCarthy, John, and Alan Prince. 1993. Prosodic morphology i: constraint interaction and satisfaction. University of Massachusetts Amherst, and Rutgers University.

Potsdam, Eric, and Maria Polinsky. In Press. Questions and word order in polynesian. In *Morphological and syntactic aspects of oceanic languages*.

Prince, Alan, and Paul Smolensky. 1993. Optimality theory: Constraint interaction in generative grammar. Rutgers University and University of Colorado, Boulder.

Richards, Norvin. 2010. Prosody and syntax. Presented at Syntax Fest 2010.
Ross, John R. 1967. Constraints on variables in syntax. Doctoral Dissertation, MIT.
Schachter, Paul. 1973. Focus and relativization. *Language* 49:19–46.
Selkirk, Elisabeth. 1996. The prosodic structure of function words. In *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*, ed. James Morgan and Katherine Demuth, 187–213. Mahwah, NJ: Lawrence Erlbaum Associates.
Selkirk, Elisabeth. 2011. The syntax-phonology interface. In *The handbook of phonological theory*, ed. John A. Goldsmith, Jason Riggle, and Alan Yu. Blackwell Publishing, 2nd edition edition.
Taumoefolau, Melenaite. 2002. Stress in Tongan. *MIT Working Papers in Linguistics* 44:341–54.
Truckenbrodt, Hubert. 1995. Phonological phrases: Their relation to syntax, focus, and prominence. Doctoral Dissertation, MIT.
Truckenbrodt, Hubert. 1999. On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30:219–255.
Vergnaud, Jean Roger. 1974. French relative clauses. Doctoral Dissertation, MIT.
Vicenik, Chad, and Grace Kuo. 2010. The intonation of tongan. UCLA.
White, James. 2010. The duration of long and short vowels in Tongan. UCLA.
Zamparelli, Roberto. 1995. Layers in the determiner phrase. Doctoral Dissertation, University of Rochester.

Appendices

Appendix A. Remaining Prosodic Issues

In (41), the relative clause XP moves to Spec,LowDP causing it to precede the DEFACC: (41a) is grammatical, but (41b) would not predict it to be possible.[23]

(41) a. [$_{IP}$ ... e sote na'a ku foo o]
the shirt PST 1.SG wash -DEFACC
b.☹ ... [$_{HighDP}$ e [$_{DemP}$ [$_{NP}$ sote] [$_{LowDP}$ [$_{XP}$ na'a ku foo] -μ [$_{CP}$ t_{NP} t_{XP}]]]]
[$_{IP}$ ... [$_{IP}$ [$_{IP}$

In order to produce (41a), you would have to violate ALIGN(IP,L;LOWDP*,L). Since (41b) does not violate this constraint, it is incorrectly predicted to be the optimal phrasing.

A second, perhaps related issue arises in (42). (42a) is in fact ungrammatical, but this analysis predicts to be the winning candidate. Namely, if the RC moves to Spec,LowDP whose head is not lexically filled, ALIGN(IP,L;LOWDP*,L) will not apply. On the other hand, ALIGN(IP,L;CP,L) will insert a break, but vacuously, as nothing is left in the CP:

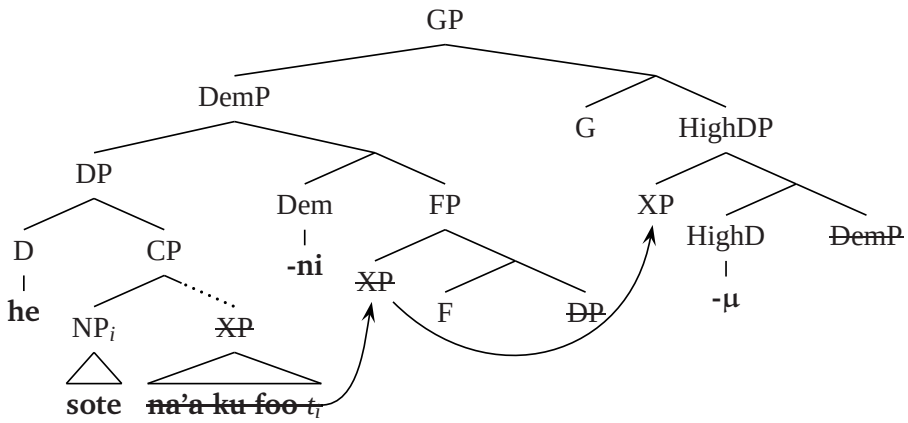
(42) a. *[$_{IP}$ ... e sote -ni na'a ku foo]
the shirt DEM PST 1.SG wash
b.💣 ... [$_{HighDP}$ e [$_{DemP}$ [$_{NP}$ sote] -ni [$_{LowDP}$ [$_{XP}$ na'a ku foo] [$_{CP}$ t_{NP} t_{XP}]]]]
[$_{IP}$ ... [$_{IP}$

Thus, the derivation in (42b) predicts the ungrammatical phrasing in (42a).

Appendix B. DEFACC as HighD

Alternatively, the DEFACC could the HighD, and (*h*)*e* could be the LowD. Under such an analysis, at least an additional three syntactic phrases and an additional two movements of the RC would be necessary:

(43)

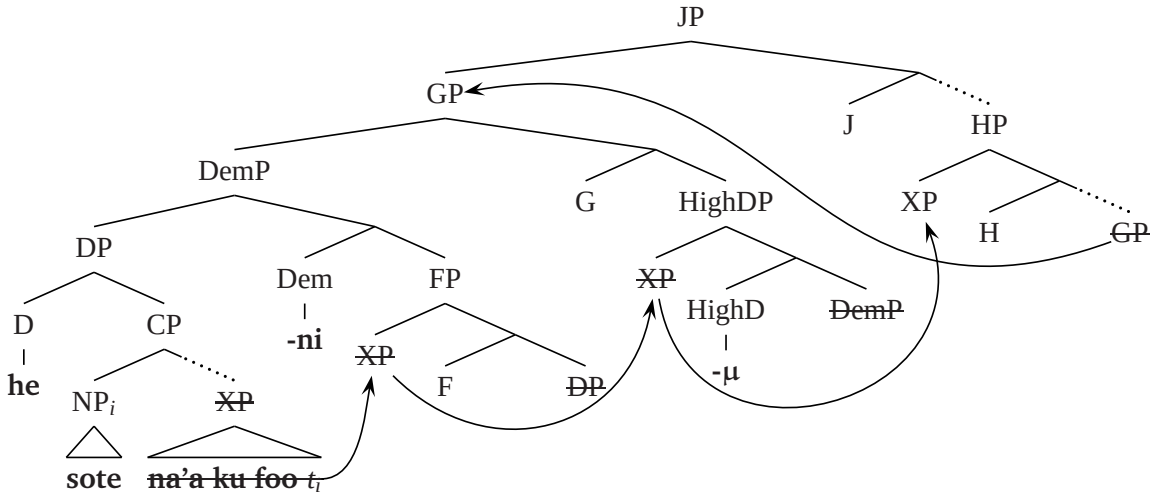


[23]The ☹ is the notation used for attested forms that the system doesn't predict; and the 💣 is used for unattested forms that the system (incorrectly) predicts.

Here, the relative clause XP moves to FP, like the RC movement in (23). After this, the Dem attracts a DP to its specifier, not unlike the NP movement in (15). This DemP is the complement of the HighD, which will attract the relative clause XP to its specifier before the DemP gets attracted to a higher projection, GP. These movements are necessary so that the relative clause XP can form a constituent with the DEFACC. This constituency is required if the syntax governs the RC and DEFACC being able prosodically phrased as in (29e-f). Moreover, this also allows the DemP to be a constituent, allowing for it to be prosodically phrased as a unit as well.

In addition, there needs to be a way for the relative clause to *follow* the DEFACC. An optional extraposition-like movement would allow for this. To achieve extraposition under an Antisymmetric approach, two more projections are required. First, one is needed for the right-extraposed material to move to: HP. Second, the rest of the DP must move to the specifier of a higher phrase, JP, in order to linearly precede the extraposed RC.

(44)



The syntactic models in (43) and (44) are powerful in that they are compatible with a more direct mapping from syntactic phrases to prosodic phrases, such as Selkirk 2011's MATCH theory. Moreover, this structure is perhaps more compatible with the semantic fact that the DEFACC has indexical properties, which are merged outside of the definiteness properties of the DP (as proposed by Abner and Burnett 2010).

However, the nature of the syntactic phrases needed for these models (i.e. FP, GP, HP and JP) remains unclear, as do the motivations for the additional movements. As such, these models would lose the independent support found for the analysis ultimately promoted in the body of this paper. Until the appropriate evidence for this alternative is found, I set these models aside for reasons of parsimony.