

Lexicalist vs. exoskeletal approaches to language mixing

Maren Berg Grimstad¹, Brita Ramsevik Riksem¹, Terje Lohndal^{1,2} and Tor A. Åfarli¹

Norwegian University of Science and Technology¹, UiT The Arctic University of Norway²

1. Introduction¹

In this article, we argue that a lexicalist analysis relying on features being an inherent property of lexical items, as in mainstream analyses within the Minimalist Program, is ill-prepared to explain the phenomenon of language mixing, i.e., intrasentential mixing where linguistic strings contain elements from both a language A and a language B. The main reason is that the feature matching/agreement typically adopted by a lexicalist feature-driven syntax naturally requires matching between elements of the same language, viz. the probe and the goal need the same features. Contrary to that requirement, we will show that items drawn from different lexicons often do not have the matching/agreeing features that are required for convergence given such an analysis, yet they frequently co-occur in language mixing. We argue that lexicalist-type feature matching theories generally predict that language mixing will be extremely restricted, contrary to fact.

The focus of this article is critical since we will concentrate on the shortcomings of lexicalist feature-driven syntax as a tool for the analysis of language mixing. This relates to the debate between MacSwan (2000, 2005) and Jake, Myers-Scotton, & Gross (2002, 2005) concerning the utility of a minimalist approach to language mixing. However, the current analysis will improve on both approaches in that it develops an exoskeletal analysis which provides a formal and more descriptively adequate generative analysis of the data. Given the programmatic nature of the Minimalist Program, we consider our exoskeletal analysis a variety of minimalism, albeit a non-lexicalist variety.

We will argue that an exoskeletal approach to language mixing can account for the data that we claim are problematic for the lexicalist approach, and further that it correctly predicts that language mixing is ubiquitous in language. The latter is not surprising, given the overwhelming evidence that both grammars are active simultaneously in the mind of

bilinguals (see Kroll & Grollan 2014 for an overview). Particular exoskeletal analyses have been defended in detail in our other work (see, e.g., Grimstad, Lohndal & Åfarli 2014, Alexiadou et al. 2015b). In this article, we will present a general outline of how an exoskeletal analysis works in order to demonstrate that it provides a better alternative to the analysis of language mixing.²

The article is organized as follows. We start out in section 2 by discussing different types of theories of language mixing. In section 3, we discuss the main characteristics of the standard version of the Minimalist Program and in particular its possible relevance to mixing phenomena. Section 4 provides a detailed criticism of one particular minimalist lexicalist analysis of mixing phenomena that has been proposed, namely Moro (2014), which seeks to explain English-Spanish mixing within the DP. In section 5, we investigate a possible minimalist lexicalist approach to mixing phenomena within the DP as found in American Norwegian. Section 6 provides an exoskeletal analysis which is argued to be superior both descriptively and theoretically. In section 7, we consider whether the lexicalist approach works better for mixing in the verbal domain, concluding that it does not, and we suggest an exoskeletal analysis of that as well. Section 8 concludes the article.

2. Theories of language mixing

There are essentially two types of language mixing theories: those that posit special machinery to handle mixing data and those that do not.³ The latter are the so-called null theories or constraint-free theories, see, e.g., Mahootian (1993), MacSwan (1999, 2014), González-Vilbazo & López (2011, 2012), Pierantozzi (2012), Bandi-Rao & den Dikken (2014), Grimstad, Lohndal, & Åfarli (2014), Åfarli (2015a), Merchant (2015), and Alexiadou (2017).

According to Mahootian (1993: 3), a null theory of what we label language mixing asserts that mixing is not constrained by any special mechanisms or principles specific to mixing, and that “exactly the same principles which apply to monolingual speech apply to codeswitching”. That a theory of language mixing should be a null theory is important simply because the internalized grammars that we postulate should be able to account for all sorts of natural language outcomes, including language mixing. If we are forced to postulate special mechanisms to account for language mixing outcomes, that would mean that neither the internalized grammar that we assume nor the special mechanisms that we adopt are on the right track. This is simply a question of theoretical parsimony. Muysken (2000: 3) states that

[t]he challenge is to account for the patterns found in terms of general properties of grammar. Notice that only in this way can the phenomena of code-mixing help refine our perspective on general grammatical theory. If there were a special and separate theory of code-mixing, it might well be less relevant to general theoretical concerns.

MacSwan (2014: 2-3) claims that whereas many language mixing theorists have considered the attainment of a null theory or constraint-free theory of language mixing to be the ideal, in practice, theories and analyses still have resorted (explicitly or implicitly) to special mechanisms for language mixing, i.e. constraint-based mechanisms in MacSwan’s terms. This has often been based on postulating constraints on where language mixing would occur (see Pfaff 1979, Poplack 1980, and Sankoff & Poplack 1981 for important early work), constraints that were unique to mixing as such. See MacSwan (2014: 2 ff.) for extensive discussion of how particular analyses in the history of language mixing theory have fared in this respect. Thus, according to MacSwan, there is an unfulfilled quest in language mixing analyses/theories for constraint-free or null theory solutions, something which he claims is fulfilled by the standard lexicalist version of the Minimalist Program. This version ensures

that “Nothing constrains [code-switching] apart from the requirements of the mixed grammars” (MacSwan 1999, 2014: 18).⁴

We agree that it is important that the analysis of language mixing is based on a null theory. However, we are not convinced that a minimalist lexicalist approach or other lexicalist feature-driven approaches are ideally suited to account for language mixing phenomena, a claim that we will try to substantiate in the remainder of this article. In what follows, we will introduce the standard lexicalist version of the Minimalist Program before turning to how and to what extent both that approach and the exoskeletal approach can explain certain patterns of language mixing.

3. The Minimalist Program

The standard version of the Minimalist Program, or just minimalism, adopts a lexicalist feature-driven model of grammar (Chomsky 1995; see a textbook version in Adger 2003).⁵ This is the version of lexicalism that we will discuss in the present paper, even though lexicalism comes in many different guises (see Ackerman et al. 2011). For reasons of space, we limit our focus to this particular version and set aside how other versions of lexicalism potentially could deal with the data in the present paper.

Within the minimalist lexicalist approach, phrase structures are generated or projected based on formal features of lexical items. Thus, the features of these lexical items determine in part the syntactic structure. A mechanism called *Select* in Chomsky (1995) provides a selection of items from the lexicon. These items constitute a Numeration or a Lexical Array (the difference need not concern us here). The computational system then generates a structure based on the numeration/lexical array. The resulting structure is further altered by way of agreement and movement, which in turn is partly driven by feature

matching/checking/valuation, that is, the need to value unvalued grammatical features that are driving the derivation. Put differently, functional features are typically unvalued and need to be valued by valued lexical features. To give one example, consider subject-verb agreement in English. The assumption is that the structure looks like in (1), where *U* denotes an unvalued feature.

(1) [TP T_[NUM:U, PERS:U, CASE:NOM] [VP [DP she_[NUM:SG, PERS:3, CASE:U]] [v [VP ...]]]]

Movement and agreement ensure that the correct grammatical representation is arrived at, where strikethrough denotes a copy/trace.

(2) [TP [DP she_[NUM:SG, PERS:3, CASE:NOM]] T_[NUM:SG, PERS:3, CASE:NOM] [VP ~~[DP she_[NUM:SG, PERS:3, CASE:U]]~~ [v [VP ...]]]]

Crucially, feature valuation takes place through Agree, which is an abstract agreement operation that connects a probe to a goal. The Agree mechanism ensures that the same feature values occur in two different places. In Chomsky (1995), a specifier-head relationship, as in (2), had to be established in order to trigger Agree (see also Koopman 2006 for a later defense of the same idea). In later versions, Agree was argued to take place long-distance (Chomsky 2000), so that movement had to be captured through additional movement-triggering features, such as a ‘generalized EPP’ feature (Chomsky 2001). Crucially, no feature can be sent off to the interfaces without being valued. An unvalued feature causes a crash at the interface.

There are several approaches to language mixing within the Minimalist Program. MacSwan (1999, 2000, 2005, 2009, 2014), Chan (2008), González-Vilbazo & López (2011, 2012), Shim (2013), Bandi-Rao & den Dikken (2014) all pursue different versions, for

example. In the present paper, we will focus on the approach most closely associated with MacSwan since this is the version that is most clearly associated with lexicalism.

MacSwan (1999, 2000, 2005, 2009, 2014) relies on the technical approach in Chomsky (1995) in developing his minimalist lexicalist approach to language mixing. Given the assumption that it needs to be a null theory, MacSwan argues that “[...] lexical items may be drawn from the lexicon to introduce features into the lexical array, which must then be valued [...] in just the same way as monolingual features must be valued, with no special mechanisms permitted” (MacSwan 2009: 326). Within a lexicalist version of minimalism, differences between languages are attributed to differences regarding lexical and functional items (cf. Borer 1984). As MacSwan (2005: 2) puts it, “[i]n the MP, there are two central components of the syntax: C_{HL} , a computational system for human language, presumed to be invariant across languages, and a lexicon, to which the idiosyncratic differences observed across languages are attributed.” Furthermore, “[p]arameters [are] restricted to the lexicon rather than operating on syntactic rules” (MacSwan 2005: 2). This is an important part of MacSwan’s approach, because it generates predictions about which patterns that can be mixed and which that cannot. In the next section, we will consider an example of this which will be used to illustrate MacSwan’s approach.

4. A minimalist lexicalist analysis of DPs in Spanish-English mixing

In recent years, there has been a lot of research into gender marking in bilingual grammars, as seen, among others, in Liceras et al. (2008), Cantone & Müller (2008), Parafita Couto et al. (2015); Valdés Kroff (2016), Valdés Kroff et al. (2016), and Johnson Fowler (2017). We will scrutinize one particular paper because it explicitly adopts a minimalist lexicalist approach, namely Moro (2014). This paper investigates Spanish-English mixing in DPs in a linguistic

variety spoken in Gibraltar from the point of view of MacSwan's approach. The article contains very little information about the status of Spanish and English in the linguistic community that the data are drawn from, or even about the immediate linguistic context of the nominal strings that she considers (a point of some importance, as we will note below), so we have at the outset to take her data at face value.

Moro (2014) considers two possible types of mixing between D and N: one where D is Spanish and N is English, as in (3), and another where D is English and N is Spanish, as in (4). According to her, only the first pattern is well-formed in the mixing variety that she considers.

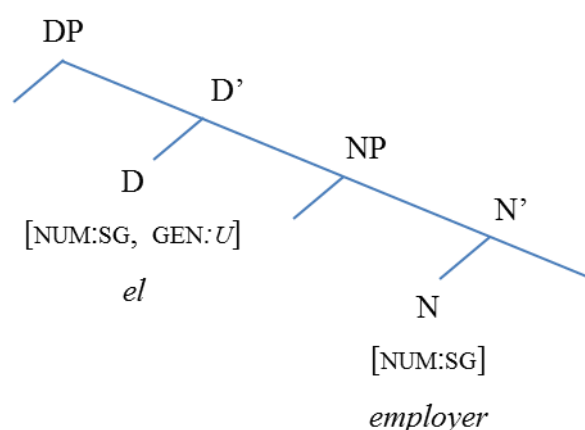
- (3) a. *el employer*
 'the employer'
- b. *la washing machine*
 'the washing machine'
- (4) a. **the casa*
 'the house'
- b. **the vecina*
 'the neighbor'

Spanish is a grammatical gender language whereas English is not, and Moro implicitly assumes that the *-er* in (3a) does not indicate natural gender. The gender difference will play a crucial role in the analysis, as we will see momentarily.

Now, consider first Moro's analysis of data of the type in (3), using (3a) as our example. D exhibits gender and number, while the English noun lacks the gender feature.

Moro (2014: 223) analyzes this in the following minimalist lexicalist manner, exploiting valued and unvalued features: “I assume [...] that the number and gender features of the determiner enter the derivation unvalued and have to be valued via Agree with the corresponding N.” Although Moro does not provide structures, we assume that the structure of (3a) is as shown in (5), where only relevant features are included.

(5)



As can be seen, the NUM category on D is valued as SG by the corresponding category on N. The GEN category on D, however, lacks a corresponding valued category elsewhere in the structure and thus appears to remain unvalued. Note that despite this, the representation in (5) is the representation of a well-formed mixing pattern, according to Moro. She suggests that “[a]ccordingly, the unvalued features number and gender in the Spanish determiner can be valued via Agree with the English noun because the former bears the full set of these features (number and gender)” (Moro 2014: 223).

This is just stipulated, and Moro does not explain exactly how the unvalued GEN feature in D can be valued by a non-existent GEN feature on the English noun, given that the determiner “bears the full set of these features (number and gender).” In our view, there would in fact not be any problems for the analysis if D failed to bear the so-called full set of

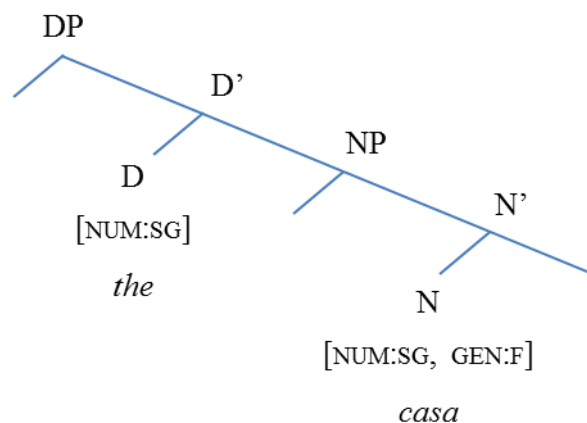
features, as Moro suggests. Rather, a problem arises precisely when the English noun does not bear the inherently valued feature required for valuation of the corresponding feature on D, as in this case. Therefore, it seems to us that Moro's solution invokes an 'impossible' theoretical rule or principle, namely a principle that amounts to a claim that an unvalued feature may be valued by a non-existent feature, i.e., that [GEN:U] in "*el* [NUM:U, GEN:U]" can be valued by [NUM:SG] in the noun *employer*.

Now, consider the mixing illustrated in (4), where the determiner is English and the noun is Spanish. We use (4a) as our example, repeated here for convenience.

- (4) a. **the casa*
'the house'

Notice, crucially, that the combination of the English determiner *the* and a Spanish noun is not well-formed according to Moro. Moro's analysis will presumably be something like the one shown in (6).

(6)



Here, N bears NUM and GEN features (since the noun is Spanish), whereas the English D does not bear a GEN feature. Given standard assumptions about feature checking, this example should be well-formed; there are no unvalued features left that could cause a crash at the interfaces.

However, this is an unwanted result as far as Moro is concerned, since this precise mixing pattern is assumed to be ungrammatical in her data. Her solution is seen in the following quotation: “On the contrary, the derivation crashes in the case of the English determiner and the Spanish noun because the feature set of the English determiner is incomplete (it lacks the gender feature)” (Moro 2014: 223). This, to us, appears to be another stipulation without any empirical or theoretical justification. First of all, we cannot see any justification for assuming something like a “complete set of features for D” that holds cross-linguistically. Moreover, since English *the* does not contain/express gender in the first place, it should not be problematic that it does not contain GEN. In fact, a more natural assumption would be that it is problematic for *the* to contain a gender feature. *Prima facie*, one should think that it would be possible for, e.g., [NUM:U] in *the* to be valued by [NUM:SG] in *casa* irrespective of the other valued or unvalued features involved. In general, an analysis that makes valuation of a particular feature dependent on the presence of features of a completely different type is in need of strong independent motivation. Unless such motivation is forthcoming, such an analysis should be discarded.

Moreover, as Licerias et al. (2008) make clear, Moro’s (2014) (cited as 2001 in Licerias et al. 2008) empirical claim is factually wrong. Contrary to Moro’s claim, instances like (4) are in fact attested in spontaneous production (see the reviews in Licerias et al. 2005, Licerias et al. 2008 and Pierantozzi 2012). For reasons of space, we will not delve into that debate here, but simply assume that (4) contains well-formed mixing data that have to be accounted for, contrary to Moro’s claim. Licerias et al. (2005, 2008) furthermore argue that the Spanish

determiner is *preferred*. They propose to account for this preference by suggesting a Grammatical Features Spell-out Hypothesis (GFSH), which claims that functional categories containing highly ‘grammaticized’ features will be chosen. Since Spanish determiners contain more features than English determiners, the speaker will choose the former. Note that the GFSH is a hypothesis about production preferences guided by a grammatical mechanism on the PF side.

In summary, we have shown that Moro’s (2014) minimalist lexicalist analysis of DP-internal language mixing is empirically inadequate. We conclude that a lexicalist feature-based analysis is the wrong tool for analyzing language mixing in the nominal domain, the reason being that such analyses require more feature matching than is actually found in mixing. In the next section, we will look at data from American Norwegian which will lend further support to this conclusion.

5. A possible minimalist lexicalist analysis of mixing in American Norwegian DPs

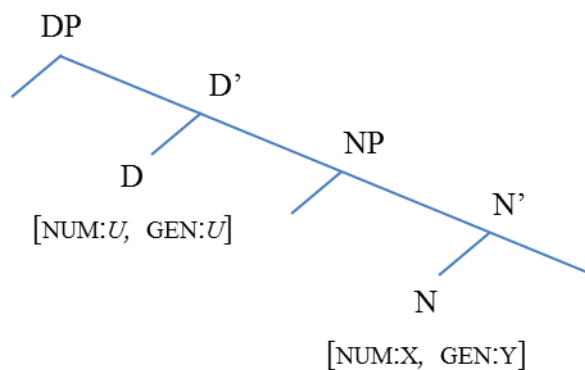
In this section, we will consider language mixing in American Norwegian DPs as a way of solidifying the conclusions reached in section 4. We will first briefly give a description of American Norwegian before we attempt to envision how a minimalist lexicalist model of the sort adopted by MacSwan and Moro could possibly handle mixing of English forms into American Norwegian DPs.

American Norwegian is a heritage variety of Norwegian spoken in North America (mainly in the US) by immigrants who came from Norway roughly from the 1850s until the 1920s, as well as their Norwegian-speaking descendants. In other words, American Norwegian is a minority language existing in the midst of a language community heavily

dominated by English. A common factor for the speakers in question is that American Norwegian is their L1 and in many cases their only language up until school age. However, through schooling and regular interaction in the community, English has come to be their dominant language. As a consequence of their language situation, these speakers often produce linguistic outcomes showing a mixture of the two languages. In the following, we will turn to analyses of such mixed linguistic outcomes in the nominal domain. The corpus we exploit for data is the *Corpus of American Norwegian Speech (CANS)* (Johannessen 2015), developed by the Text Laboratory at the University of Oslo. At the moment, this corpus comprises recordings from 50 individual speakers.

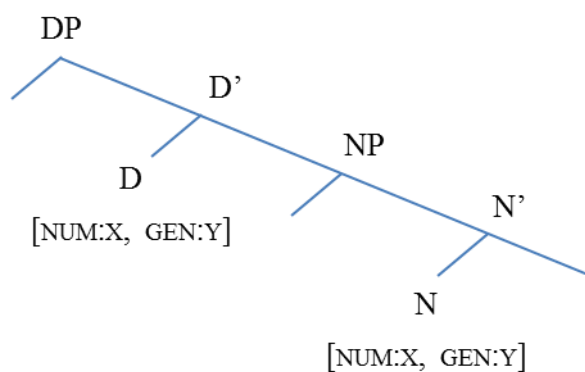
Given that American Norwegian is a heritage variety of Norwegian, let us first take a look at how the mechanism of feature valuation proceeds in a Norwegian DP without mixing. Like Spanish, Norwegian is a grammatical gender language, and we use a simplified DP structure which only contains a D-projection and an N-projection. As we are, for the moment, concerned with a minimalist lexicalist type analysis, we make the lexicalist assumption that the agreement features NUM and GEN are inherent in N, and that DP internal agreement comes about as a result of the agreement features in N valuing the corresponding unvalued features in D. This is illustrated in (7), where $[\text{NUM}:X, \text{GEN}:Y]$ in N denotes the particular inherently fixed agreement features in question, and where $[\text{NUM}:U, \text{GEN}:U]$ denotes the corresponding agreement features in D.

(7)



After valuation of the unvalued features in D by a probe – goal relation, the resulting agreement structure will be as shown in (8), where N and D have identical feature values.

(8)



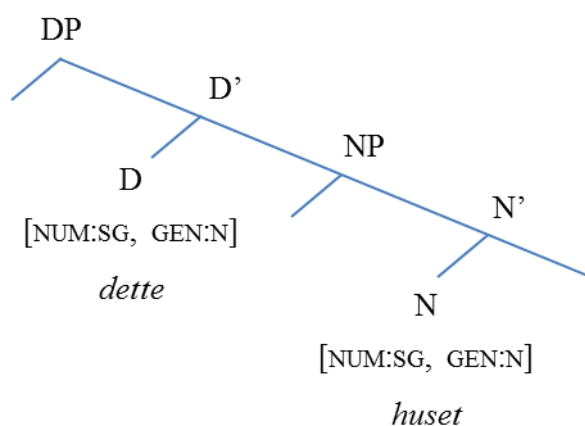
Consider now a concrete example:

- (9) a. *dette hus-et*
 this.SG.N house-SG.DEF.N
 ‘this house’
- b. **denne hus-et*
 this.SG.M/F house-SG.DEF.N
- c. **dette hus-a*

this.SG.N house-PL.DEF.N

(9a) shows the correct agreement pattern inside the DP, with both N and D marked as SG and N. (9b) is ungrammatical due to an agreement mismatch, the N being marked as SG.N, whereas D is marked as SG.M/F. In other words, there is a gender mismatch between N and D that cannot be generated given the Agree mechanism. (9b) therefore fails to be generated, as desired. The same goes for (9c), except in this case there is a number mismatch, the N being marked as PL while D is marked as SG. (9a) is shown in (10), which is the structure after valuation of the unvalued features in D.⁶

(10)

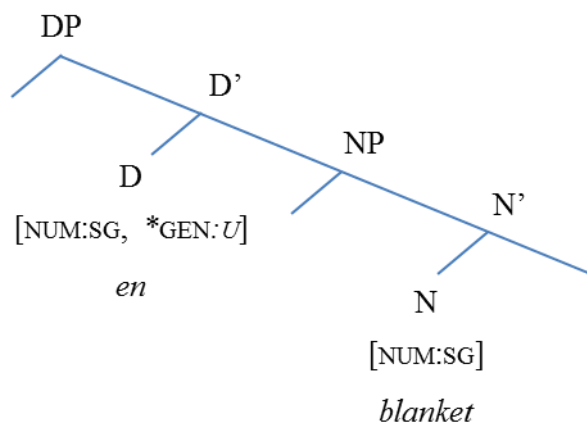


Let us next consider feature valuation in American Norwegian DPs in which English elements are mixed into the otherwise Norwegian string. In American Norwegian, it is common for an English noun to be embedded under a Norwegian determiner, e.g., an indefinite article. This is exemplified in (11). The information in parenthesis behind each American Norwegian example is a reference to the speaker in the CANS corpus who uttered that specific phrase, and the mixed English noun is boldfaced.⁷ Note that gender is not fixed across speakers and is not in general identical to their Norwegian translational equivalents.⁸

- (11) a. *en blanket* (rushford_MN_01gm)
 a.M blanket
 ‘a blanket’
- b. *ei nurse* (coon_valley_WI_02gm)
 a.F nurse
 ‘a nurse’
- c. *et crew* (westby_WI_03gk)
 a.N crew
 ‘a crew’

Adopting a standard minimalist lexicalist analysis, D has unvalued GEN and NUM features that must be valued by the corresponding fixed values on N, as explained above for standard Norwegian. However, there is a problem with this analysis given mixing cases like those in (11). Since English nouns do not have a gender feature, the GEN feature of D remains unvalued and the derivation will crash, contrary to the fact that such structures are common and therefore should converge.⁹ These examples are parallel to the Spanish cases in (3), where Moro (2014) stipulated a solution where the “full set” of features in D would enable valuation from the English noun. As discussed in section 4, we do not see how an unvalued GEN feature in D can be valued by a non-existing GEN feature in N, and therefore reject such an analysis. The ill-formed structural representation of (11a) is given in (12), the ill-formedness being indicated by an “*” on the offending feature category.

(12)



Now, consider other comparable examples where D is a demonstrative or determiner, illustrated in (13) and (14), respectively.¹⁰

- (13) a. *denne cheese* (blair_WI_04gk)
 this.M/F cheese
 ‘this cheese’
- b. *denne heritage tour-en* (flom_MN_01gm)
 this.M heritage tour-SG.DEF.M
 ‘this heritage tour’
- c. *dette computer business* (harmony_MN_01gk)
 this.N computer business
 ‘this computer business’
- (14) a. *alt det gamle stuff* (chicago_IL_01gk)
 all the.N old stuff
 ‘all the old stuff’

- b. *den digre chopper-en* (blair_WI_01gm)

the.M huge chopper-SG.DEF.M

‘the huge chopper’

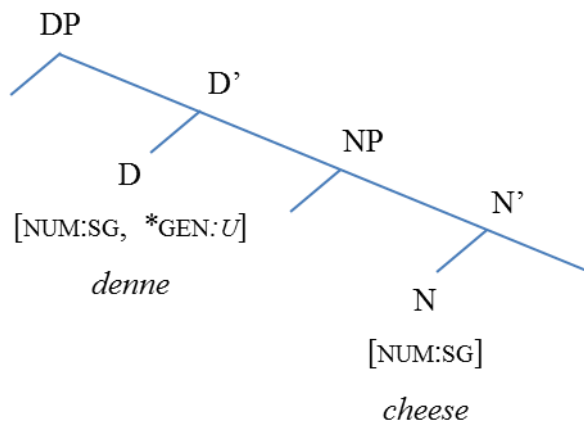
- c. *den samme lodg-en* (vancouver_WA_03uk)

the.M same lodge-SG.DEF.M

‘the same lodge’

Again, the problem is that the GEN feature of D cannot be valued because the English noun does not bear a gender feature, cf. the ill-formed structural representation of (13a) in (15).

(15)



Now, consider American Norwegian DPs where the Norwegian-English mixing pattern is switched, so to speak. We will focus on patterns where there is a mix consisting of an English definite article and a Norwegian noun, as seen in the examples in (16).

- (16) a. *the by* (chicago_IL_01gk)

the city.SG

‘the city’

- b. *the gård* (vancouver_WA_01gm)

the farm.SG

‘the farm’

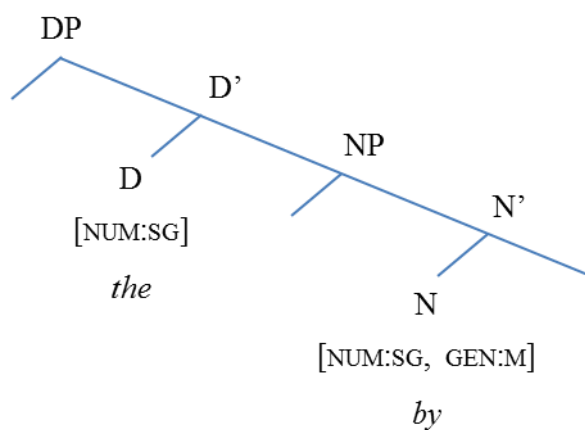
- c. *the penger* (albert_lea_MN_01gk)

the money.PL

‘the money’

In these examples, there is a Norwegian noun with gender and number features, but the English definite article *the* probably bears only a number feature and in any case lacks a gender feature. The structural representation of (16a) is shown in (17).

(17)



As can be seen in (17), there are no features that remain unvalued in this representation. As such, this representation should be deemed well-formed, at least as far as feature valuation is concerned. Note that these strings are parallel to the Spanish-English cases in (4), which Moro (2014) judged to be ungrammatical. The rationale for this conclusion was that the “incomplete” set of features in D caused an inability of feature valuation from N. Again, as discussed in section 4, we do not see how the lack of a GEN feature in D should prevent NUM in N from valuing NUM in D. (17) illustrates that after such a valuation, there are no unvalued features left in the structure that could make the derivation crash. As discussed above, Licerias et al. (2008) argue against Moro on empirical grounds, which also aligns with the evidence found in American Norwegian.

As seen in this subsection, attempting to analyze the American Norwegian data under a lexicalist approach is problematic. On the one hand, if we were to accept Moro’s analysis, mixed phrases containing Norwegian D and English N would be acceptable, whereas phrases with English D and Norwegian N would be unacceptable. Rejecting Moro’s analysis based on the discussion in section 4 and employing standard minimalist mechanisms of valuation, on the other hand, would apparently reverse the picture; the latter pattern would be acceptable and the former unacceptable. The crucial fact is nevertheless that both mixing patterns are attested in our corpus. In the next section, we will propose an alternative analysis.

6. An exoskeletal analysis of language mixing in DPs

We argue that an exoskeletal approach to grammar provides a more adequate analysis of language mixing. The following outlines such an analysis and demonstrates how the previously discussed data may be analyzed.

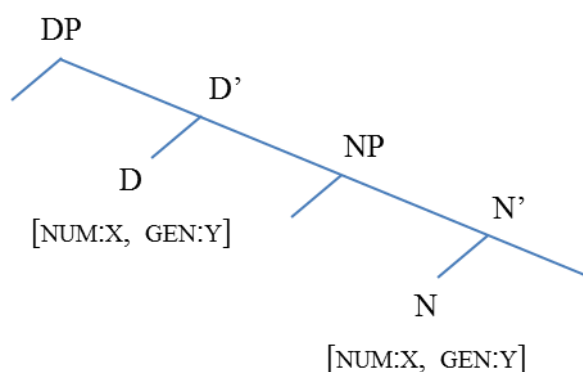
Exoskeletal approaches to grammar, also known as generative, neo-constructivist approaches, have been developed by several scholars, e.g., van Hout (1996), Marantz (1997, 2013), Borer (2005a, b, 2013), Áfarli (2007), Ramchand (2008), Lohndal (2012, 2014), and Alexiadou et al. (2015a). The details of the proposed models and analyses vary, but crucially, they share a common core, namely that the structures are generated independently from the lexical items that will come to realize them. Hence, lexical items do not have inherent, functional features that determine the derivation of the structures, but functional features are instead defined by the syntactic structure. Although exoskeletal models primarily have been applied to monolingual data, the approach has proven to be quite successful in analyzing language mixing, see, e.g., González-Vilbazo & López (2011, 2012), Pierantozzi (2012), Bandi-Rao & den Dikken (2014), Grimstad, Lohndal, & Áfarli (2014), Áfarli (2015a), Alexiadou et al. (2015b), Merchant (2015), Alexiadou (2017), and Riksem (in press).

Furthermore, the model employed in this article builds on certain aspects of Distributed Morphology (DM) (see, e.g., Harley & Noyer 1999, 2000; Marantz 1997; Embick 2000; Alexiadou 2001; Embick & Noyer 2007). The core and common ingredient is the assumption that syntactic structures consist of features that define the morphosyntactic properties. In addition, what is typically known as the lexicon is distributed in DM across three separate lists: 1) syntactic terminals, 2) vocabulary items, and 3) encyclopedic information. These are accessed at different points throughout the derivation. The structure then distinguishes between two types of terminals: functional features or feature bundles, and designated slots for root/stem insertion.¹¹ At Spell-Out, morpho-phonological exponents, known in DM as vocabulary items, are inserted and give the structure its phonological realization. This process is radically different for the two types of syntactic terminals. Following the Subset Principle (Halle 1997), functional exponents are required to match all or a subset of the functional features specified in the structure to be inserted. If two exponents

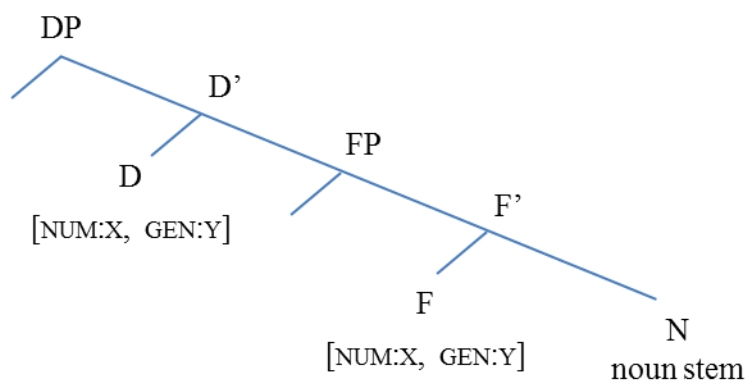
are eligible for insertion, the one matching the most features will win. This ensures that a given feature or feature bundle in the structure is spelled out by the most appropriate exponent available. Substantial exponents, i.e., roots or stems, are instead inserted into designated slots in the structure without such feature matching requirements, except that they need to match the relevant category feature.

With this model, we can now provide an analysis of the mixed American Norwegian DPs. Instead of (8), repeated below, where the inflectional properties are a property of the noun itself, we assume (18), where the inflectional properties are generated under a functional projection called F, and the noun stem (N) is generated in the complement domain of that F.

(8)



(18)



Let us briefly review some evidence in favor of the FP projection. A core piece of evidence comes from the existence of double definiteness in Norwegian and American Norwegian.

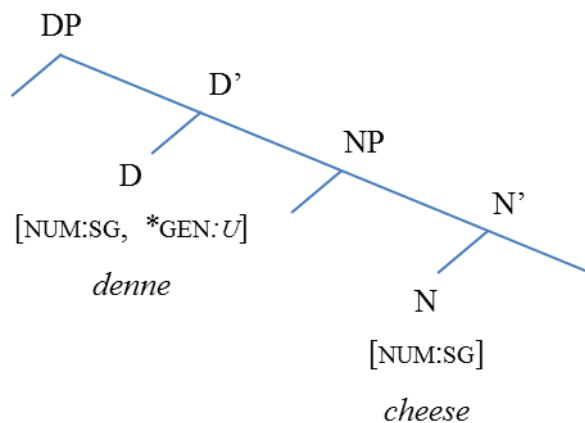
Consider (19):

- (19) *den gamle mann-en*
 the-DEF.SG.M old man-DEF.SG.M
 ‘the old man’

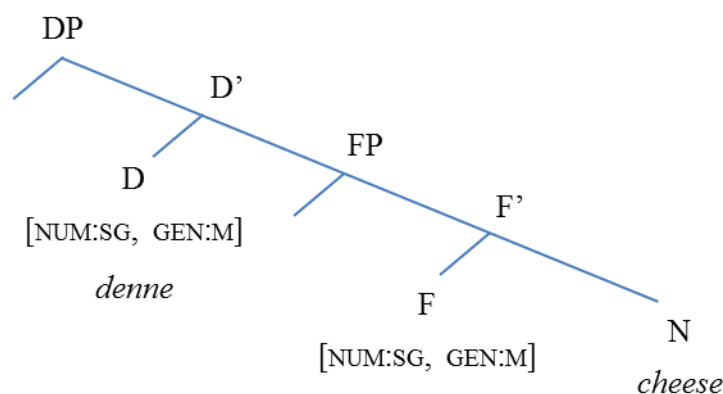
In (19), definiteness is encoded both pre- and postnominally. Julien (2005), building on an extensive review of previous research into the nominal phrase in Norwegian, argues that there is a functional projection for each of the two definiteness features. The prenominal one is situated in D, whereas she suggests that the postnominal one serves as the head of a definiteness projection in the lower domain of the DP. The adjective, situated as the specifier of its own projection (α P in Julien 2005), then agrees with the features of the definiteness and D heads. Julien also assumes that there is a separate NumP. In (18), FP is a different label for the lower definiteness projection, which also encompasses number. We have collapsed both definiteness and number onto one head, both because we have not been able to find evidence in American Norwegian for distinguishing them into two projections, and because the choice between one or two projections is not crucial for present purposes. Furthermore, we assume that FP is always present, but its feature content differs across varieties. In English, FP only has a NUM feature, whereas the Norwegian and American Norwegian FP both have NUM, GEN and DEF features. Just like subject-verb agreement at the sentential level differs between English and Norwegian, with Norwegian not exhibiting features for subject-verb agreement, so does the presence of the functional features in the nominal domain.

Given (18), noun stems from any language will acquire the inflectional properties of the language that specifies the syntactic frame.¹² Thus the prediction is that mixing is fairly free, as is actually the case in American Norwegian. Let us consider *denne cheese*, “this cheese”. As the structure in (15) shows, repeated below, the lexicalist approach fails to predict this pattern, as the feature matching process will leave the GEN feature of D unvalued. The exoskeletal representation in (20) fares better. The functional features are here determined by the syntactic structure, which in this case is Norwegian. Given the Subset Principle, inserted functional exponents must provide the best match to the functional features in the structure. In this case, the Norwegian exponent will provide the best match, as the structure holds a GEN feature. The complement position of F, on the other hand, is available for insertion of an English noun stem.¹³ Mixing patterns like these can thus be characterized as English stems being inserted into Norwegian structure, and are in fact the most frequent in the corpus.

(15)



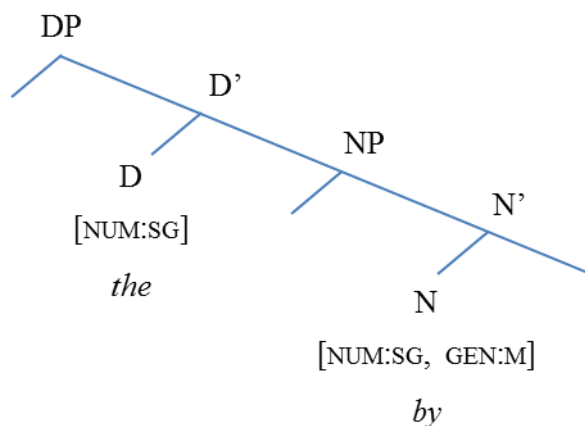
(20)



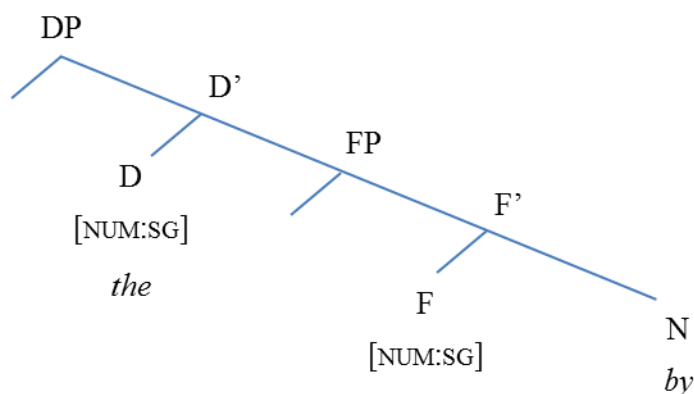
As mentioned in footnote 10, earlier American Norwegian (and also European Norwegian) would typically have double definiteness in examples corresponding to (15). As shown in (13b) and (14b, c), that is also very often the case in contemporary American Norwegian; although as noted, there is vacillation among contemporary speakers (see Riksem 2017 for an analysis of instances where the postnominal definite suffix are missing, arguing in favor of features being either rearranged or erased from the structure). However, it is important to point out that the presence of double definiteness in American Norwegian mixing examples like (13b) and (14b, c) provides further empirical support for an exoskeletal analysis. The reason is that the definiteness suffix bears number and gender features which must be Norwegian, thus providing evidence for the existence of the F head as part of the exoskeletal frame. See Riksem (in press) for further data and analysis.

Considering the second pattern of determiner-noun mixing in American Norwegian, i.e., where the determiner is English and the noun Norwegian, this, too, can be analyzed with an exoskeletal model. The lexicalist version of (16a) *the by* is repeated in (17), followed by an exoskeletal representation in (21).

(17)



(21)



As the structures show, cases of English D plus Norwegian N are successfully analyzed in both frameworks, despite Moro's claim that such combinations are ungrammatical. The crucial difference is, however, that whereas the exoskeletal model is capable of accounting for both mixing patterns attested in the corpus, the lexicalist analysis will fail in cases like (15).

The process of generating the structure in (21) is not considered to be quite the same as the one for the structure in (20). Whereas the main language in (20) is Norwegian, hence the GEN feature in D and F, the main language for the DP in (21) is assumed to be English, meaning neither D nor F have a GEN feature. Evidence in favor of this is the absence of double definiteness, i.e., the speaker says *by* 'city' and not *byen* 'the city', which together with the English determiner suggests that the determiner causes a change into English for the rest of the noun phrase. In other words, what we assume we have here is a Norwegian stem

inserted into an English structure, the opposite of what we have seen so far.¹⁴ This English DP can further be inserted into the DP slot of a larger Norwegian utterance, which is the case for this particular example:

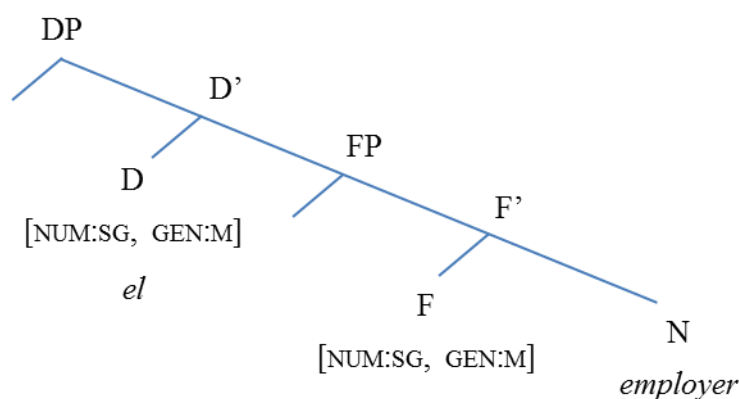
(22) *Jeg husker ikke **the** by der vi stoppet.* (chicago_IL_01gk)

I remember not the city there we stopped

“I don’t remember the city where we stopped.”

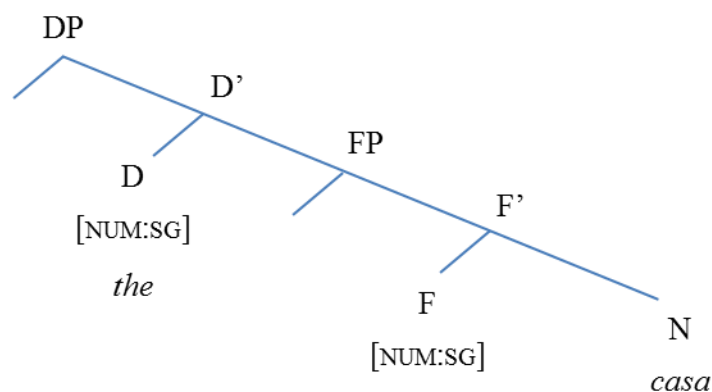
As for the Spanish – English mixing data, the mixing pattern that is well-formed according to Moro (2014) corresponds to the American Norwegian structure (20), see (23):

(23)



In addition, the pattern that Moro (2014) claims is ill-formed is also, like the American Norwegian (21), predicted to be grammatical given an exoskeletal analysis, cf. (24).

(24)



As mentioned in section 4, Moro (2014) provides little information concerning the status of Spanish and English in the linguistic community she draws her data from. There is also virtually nothing about the informants, the collection of data or even the immediate linguistic context of the nominal strings in question. This lack of information is problematic in several ways. From an exoskeletal perspective, in a linguistic community where Spanish is the main language spoken, it is expected that NPs like *el employer* will be more common than those like *the casa*. This is simply because if you are speaking Spanish, the structures will be Spanish as well, whereas you would need an English DP for the Subset Principle to licence an English determiner like *the*.

Moreover, social factors such as prestige are likely to impact the notion of something being acceptable versus unacceptable, which is a crucial point of discussion when employing acceptability judgments in language mixing. The alleged unacceptability of phrases like *the casa* may thus actually stem from sociolinguistic norms in the language community. In order to give a thorough analysis of language mixing patterns, one must consider both which language constitutes the main one as well as other factors that may influence the judgements of mixed phrases. See Liceras et al. (2008) for further discussion of preferences in the linguistic production of speakers who mix English and Spanish.

7. An extension: Lexicalism does not fare better in the verbal domain

So far we have seen that mixing patterns in the nominal domain provide evidence against a minimalist lexicalist analysis and that an exoskeletal analysis is descriptively more adequate. In this section, we will briefly show that mixed verb forms in American Norwegian provide additional evidence against a minimalist lexicalist analysis, and, subsequently, that the way the mixed verbs pattern is predicted by an exoskeletal analysis.

An example of mixed verb forms is illustrated in (25).

- (25) *vi bare satt der og watch-a da* (sunburg_MN_03gm)
 we just sat there and watch-PAST then
 “We just sat there watching then.”

Here the verb stem, *watch*, is clearly English, while the tense inflection is clearly Norwegian, *-a* being a past tense suffix belonging to the main class of Norwegian weak verbs. (26) provides additional data.

- (26) a. *teach-er* (rushford_MN_01gm)
 teach-PRES
 b. *play-de* (coon_valley_WI_03gm)
 play-PAST

The main pattern we find in CANS when lone English verbs are mixed into otherwise Norwegian utterances is that the mixed English verb stems appear with the appropriate Norwegian inflection, as shown in (25) and (26). This is in accordance with what we already

saw for nouns in, e.g., (13b) and (14b, c), where the English nouns for the most part occur in exactly the position their Norwegian counterparts would in the noun phrases, with appropriate inflections.

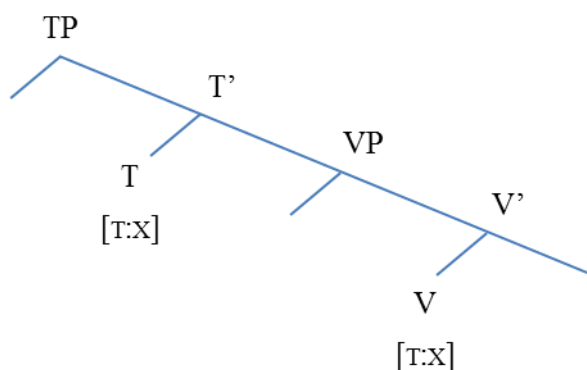
Since the verbal stem and the tense affix belong to different languages, it is not self-evident that standard analyses of the T – V relation can explain these data. Chomsky (1995: 195) provides a classic formulation of the theoretical possibilities when analyzing the T – V relation, which again amounts to an exoskeletal vs. lexicalist analysis:

The main verb typically "picks up" the features T and Agr [...], adjoining to an inflectional element to form [V I]. There are two ways to interpret the process, for a lexical element *a*. One is to take *a* to be a bare, uninflected form; PF rules are then designed to interpret the abstract complex [*a* I] as a single inflected phonological word. The other approach is to take *a* to have inflectional features in the lexicon as an intrinsic property (in the spirit of lexicalist phonology); these features are then checked against the inflectional element I in the complex [*a* I].

Chomsky embraces the second solution, assuming that the inflected form of the verb already is created in the lexicon and subsequently inserted into the syntax with an inherent feature bundle, i.e., fully tensed. This amounts to a fullblooded lexicalist analysis of the T – V relation, and is also the one MacSwan makes use of.¹⁵

Just as we saw for the DPs, the mechanism of feature checking or valuation plays a crucial role within the lexicalist analysis of the T – V relation. In order to prohibit arbitrary insertion of tensed verbs, a given tensed form that is inserted into syntax must be checked against a corresponding feature in T to ensure that it occurs in a structurally correct position.¹⁶ (27) shows a relevant structure ("x" denotes a particular tense feature value).

(27)



However, there are at least three problems with this lexicalist analysis of the T – V relation.

The first problem is how to account for the overwhelming occurrence of Norwegian tense suffixes on English verbal stems if the verb has “inflectional features in the lexicon as an intrinsic property”. As argued in Åfarli (2015b: 168-169), this assumption would make us expect that an English verb stem should have English tense inflection, and it remains a mystery that the inflection instead is Norwegian.

The second problem concerns the feature checking between T and V. An English verb will, given lexicalist assumptions, bear not only a tense feature, but also agreement features. The most standard assumption is that those features of the verb must be checked against corresponding features in T. However, if an English verb is employed in a Norwegian structure, T will only contain a tense feature. Thus, the putative agreement features of the verb cannot be checked. This is parallel, *mutatis mutandis*, to the failure of feature checking in the nominal domain that we discussed earlier.

The third problem has to do with the richness of mixing varieties. As was the case for the DPs, the main pattern for mixed verbs seen in (25) and (26) is not the only type found in the corpus. We have found 292 lone English verbs occurring in otherwise Norwegian utterances in the corpus, and out of those, 210 had Norwegian inflectional suffixes whereas 82 did not, instead seemingly appearing with the appropriate English morphology.¹⁷ Examples are given in (28), where # marks a pause.

- (28) a. *så e # I kunne ikke # e sing #* (north_battleford_SK_02gk)
 so e # I could not # e sing #
 ‘so I couldn’t sing’
- b. *og han sends # han er #* (vancouver_WA_01gm)
 and he sends # he is #
 ‘and he sends’
- c. *e v- vi # translated “Synnøve Solbakken”* (gary_MN_01gm)
e w- we # translated “Synnøve Solbakken”
 ‘we translated (the novel) *Synnøve Solbakken*’

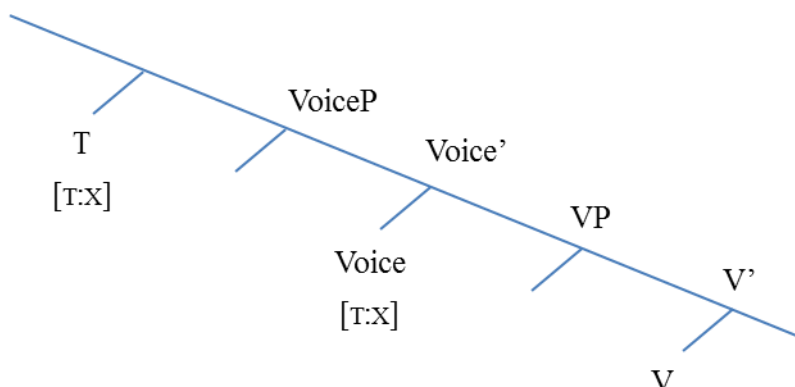
Since these appear to be English verbs with English inflection, (28b) even displaying the appropriate 3.P.SG. suffix *-s*, we can assume that the entire TP structure is English.¹⁸

Importantly, since both the earlier patterns and the patterns in (28) are attested in the corpus, we would like an analysis that is flexible enough to handle both.

We will now sketch an exoskeletal analysis of mixed verb forms and show that it can handle the problems reviewed above and make the right predictions for American Norwegian. As in our exoskeletal analysis of nouns, we will assume an exoskeletal analysis of verbs where the inflectional properties of the verb are not an intrinsic property of the verb itself, contrary to the lexicalist analysis, but of a functional head whose complement contains the verb stem. Thus, the language of the structure and its inflectional elements may belong to a language different from the stem, as is the case for the examples in (25) and (26).

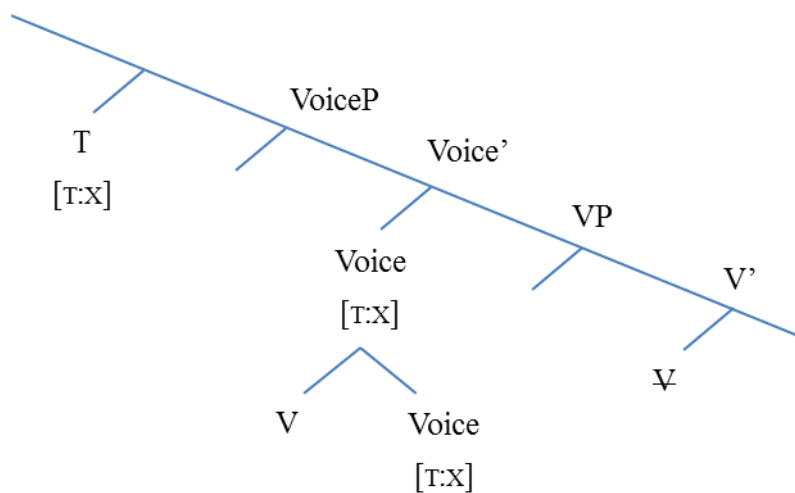
Let’s start with a structure where the verb phrase of a simple clause is c-commanded by T and the verb phrase itself consists of a Voice-projection (Kratzer 1996) with a VP in its complement domain, as in (29).¹⁹

(29)



Focusing on the T – Voice – V relation for Norwegian verb phrases, we assume, as mentioned, that the tense feature is generated under T and that Voice contains a corresponding tense feature that is checked against the feature in T. V acquires tense by obligatory movement to Voice, giving us the structure sketched in (30).

(30)



Note that V, being a stem, can be inserted into the verb phrase from any language, like N can be inserted from any language into a given noun phrase structure.

The main exoskeletal point of this analysis is that the generation of the tense inflection is divorced from the generation of the verb stem, and that these two elements are syntactically integrated during the derivation. Thus, the first and second problems for a lexicalist analysis, as noted earlier, are solved. A Norwegian tense affix is expected on the English verb, since the Voice projection, containing the lower tense feature, is part of the Norwegian structure. Furthermore, feature checking between T and the verb, now situated in Voice, is unproblematic because both T and Voice belong to the same abstract syntactic frame and thus contain corresponding features, whereas the verb stem alone may belong to another language.

The exoskeletal analysis can also deal adequately with the other attested verb pattern, illustrated in (28), where the mixed English verb has English inflection although the rest of the clause may be Norwegian. Such a pattern may occur if the main exoskeletal structure in fact is English, with Norwegian phrases inserted in argument and adjunct positions, which is a possible mixing pattern given this analysis.

Summing up this brief section on mixing in the verbal domain, our hypothesis is that just like in the nominal domain, the exoskeletal analysis can account for all the attested mixing patterns in the corpus whereas the lexicalist one is problematic in several respects, as noted. This clearly favors the exoskeletal approach.²⁰

8. Conclusion

According to MacSwan (2014: 18), generative theories prior to the Minimalist Program did not provide sufficient theoretical tools for implementing an adequate language mixing or codeswitching analysis; prior attempts were simply doomed to fail, seeing as the formal model they built on was wrong to begin with. MacSwan contrasts this approach with his own, the minimalist lexicalist one, which he finds more satisfying:

Within the [Minimalist Program], structures are built from a stock of lexical items, with lexical insertion [...] taking place at the outset. This important development permits [codeswitching] researchers to probe the structural consequences of particular lexical items from specific languages [...].

By scrutinizing one paper explicitly adopting such a minimalist lexicalist approach, Moro (2014), we contest the claim that syntactic structure is dependent on the language of the lexical items involved. In fact, we show that Moro's analysis of English – Spanish mixing between a determiner and a noun fails to predict the observed patterns: Under standard minimalist feature valuation mechanisms the pattern Moro argues is well-formed, i.e. a Spanish D and English N, would crash the derivation, whereas the alleged ill-formed pattern, an English D and Spanish N, actually should converge.

Adopting an exoskeletal approach to grammar, instead, a model that entails the independent generation of syntactic structures, we successfully analyze both patterns. This is further supported by data showing English – Norwegian mixing between a determiner and a noun in the heritage language American Norwegian, as well as mixing in the verbal domain in the same language. For instance, sections 5-6 illustrate that an English noun inserted into American Norwegian has no power to influence the overall syntactic structure of the phrase, not even on its own inflection; and likewise, section 7 shows that the same is true for English verbs inserted into otherwise Norwegian phrases in American Norwegian.

Thus, it remains to be seen what empirical shortcomings exoskeletal models might be subject to cross-linguistically, but as we have tried to argue in this article, they seem to remain our most viable options for analyzing language mixing data at present.

References

- Ackerman, F., G. Stump, & G. Webelhuth. 2011. Lexicalism, periphrasis, and implicative morphology. In R. D. Borsley & K. Börjars (eds.), *Non-transformational syntax: Formal and explicit models of grammar*, pp. 325-358. Malden, MA: Blackwell-Wiley.
- Adger, D. 2003. *Core Syntax*. Oxford: Oxford University Press.
- Áfarli, T. A. 2007. Do Verbs have Argument Structure? In E. Reuland, T. Bhattacharya & G. Spathas (eds.), *Argument Structure*, pp. 1-16. Amsterdam: John Benjamins.
- Áfarli, T. A. 2015a. A Syntactic Model for the Analysis of Language Mixing Phenomena: American Norwegian and Beyond. In B. R. Page & M. T. Putnam (eds.), *Moribund Germanic Heritage Languages in North America*, pp. 12-33. Leiden: Brill.
- Áfarli, T. A. 2015b. Hybrid verb forms in American Norwegian and the analysis of the syntactic relation between the verb and its tense. In J. B. Johannessen & J. C. Salmons (eds.), *Germanic Heritage Languages in North America: Acquisition, attrition and change*, pp. 161-177. Amsterdam: John Benjamins.
- Alexiadou, A. 2001. *Functional structure in nominals: nominalisation and ergativity*. Amsterdam: John Benjamins.
- Alexiadou, A. 2017. Building verbs in language mixing varieties. *Zeitschrift für Sprachwissenschaft*, 36, 165-192.
- Alexiadou, Artemis, E. Anagnostopoulou, & F. Schäfer. 2006. The properties of anticausatives crosslinguistically. In M. Frascarelli (ed.), *Phases of Interpretation*, pp. 187-212. Berlin: Mouton de Gruyter.
- Alexiadou, A., E. Anagnostopoulou, & F. Schäfer. 2015a. *External arguments in transitivity alternations: a layering approach*. Oxford: Oxford University Press.

- Alexiadou, A., T. Lohndal, T. A. Áfarli & M. B. Grimstad. 2015b. Language Mixing: A Distributed Morphology approach. In T. Bui and D. Özyildiz (eds.), *Proceedings of NELS 45*, pp. 25-38. Create Space.
- Bandi-Rao, S., & M. Den Dikken. 2014. Light Switches: On v as a pivot in codeswitching, and the nature of the ban on word-internal switches. In J. MacSwan (ed.), *Grammatical Theory and Bilingual Codeswitching*, pp. 161-184. Cambridge, MA: MIT Press.
- Borer, H. 1984. *Parametric Syntax*. Dordrecht: Foris.
- Borer, H. 2005a. *Structuring Sense I: In Name Only*. Oxford: Oxford University Press.
- Borer, H. 2005b. *Structuring Sense II: The Normal Course of Events*. Oxford: Oxford University Press.
- Borer, H. 2013. *Structuring Sense III: Taking Form*. Oxford: Oxford University Press.
- Cantone, K.F. & N. Müller. 2008. Un nase or una nase? What gender marking within switched DPs reveals about the architecture of the bilingual language faculty. *Lingua*, 118, 810-826.
- Chan, B. H-S. 2008. Code switching, word order and the lexical/functional category distinction. *Lingua*, 118, 777-809.
- Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, N. 2000. Minimalist inquiries: the framework. In R. Martin, D. Michaels and J. Uriagereka (eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, pp. 89-155. Cambridge, MA: MIT Press.
- Chomsky, N. 2001. Derivation by Phase. In M. Kenstowicz (ed.), *Kenneth Hale: A Life in Language*, pp. 1-52. Cambridge, MA: MIT Press.
- Embick, D. 2000. Features, syntax, and categories in the Latin perfect. *Linguistic Inquiry*, 31, 185-230.

- Embick, D., & R. Noyer. 2007. Distributed Morphology and the syntax morphology interface. In G. Ramchand and C. Reiss (eds.), *The Oxford Handbook of Linguistic Interfaces*, pp. 289-324. Oxford: Oxford University Press.
- Folli, R., & H. Harley. 2007. Causation, Obligation, and Argument Structure: On the Nature of Little *v*. *Linguistic Inquiry*, 3, 197-238.
- Goldrick, M., M. Putnam, & L. Schwarz. 2016. Coactivation in bilingual grammars: A Gradient Symbolic Computation account of code mixing. *Bilingualism: Language and Cognition*, 19, 857-876.
- González-Vilbazo, K., & L. López. 2011. Some properties of light verbs in code-switching. *Lingua*, 121, 832-850.
- González-Vilbazo, K., & L. López. 2012. Little *v* and parametric variation. *Natural Language and Linguistic Theory*, 30, 33-77.
- Grimstad, M. B. 2017. The CS/borrowing debate: Evidence from English-origin verbs in American Norwegian. *Lingue e Linguaggio XVI*: 3-34.
- Grimstad, M. B., T. Lohndal, & T. A. Åfarli. 2014. Language mixing and exoskeletal theory: A case study of word-internal mixing in American Norwegian. *Nordlyd*, 41, 213-237.
- Halle, M. 1997. Distributed morphology: Impoverishment and fission. In B. Bruening, Y. Kang & M. McGinnis (eds.), *MITWPL 30: Papers at the Interface*, pp. 425-449. Cambridge: MITWPL.
- Harley, H. 1995. Subjects, Events, and Licensing. Doctoral dissertation, MIT.
- Harley, H., & R. Noyer. 1999. Distributed Morphology. *Glott International*, 4, 3-9.
- Harley, H., & R. Noyer. 2000. Licensing in the non-lexicalist lexicon. In B. Peeters (ed.), *The Lexicon/Encyclopedia Interface*, pp. 349-374. Amsterdam: Elsevier Press.

- Jake, J. L., C. Myers-Scotton, & S. Gross. 2002. Making a minimalist approach to codeswitching work: Adding the Matrix Language. *Bilingualism: Language and Cognition*, 5, 69-91.
- Jake, J. L., C. Myers-Scotton, & S. Gross. 2005. A response to MacSwan (2005): keeping the matrix language. *Bilingualism: Language and Cognition*, 8, 271-276.
- Johannessen, J. B. 2015. The Corpus of American Norwegian Speech (CANS). In B. Megyesi (ed.), *Proceedings of the 20th Nordic Conference of Computational Linguistics, NODALIDA 2015, May 11-13, 2015, Vilnius, Lithuania*. NEALT Proceedings Series 23.
- Johnson-Fowler, C. 2017. Cross-language interaction during language production in L1 German-L2 Italian bilinguals. Doctoral dissertation, Penn State University.
- Julien, M. 2003. Double definiteness in Scandinavian. *Nordlyd*, 31, 230-244.
- Julien, M. 2005. *Nominal Phrases from a Scandinavian Perspective*. Amsterdam: John Benjamins.
- Koopman, H. 2006. Agreement configurations: In defense of “Spec head”. In C. Boeckx (ed.), *Agreement Systems*, pp. 159-199. Amsterdam: John Benjamins.
- Kratzer, A. 1996. Severing the External Argument from the Verb. In J. Rooryck, & L. Zaring. (eds.), *Phrase Structure and the Lexicon*, pp. 109–137. Dordrecht: Kluwer.
- Kroll, J. F., & T. Gollan. 2014. Speech planning in two languages: What bilinguals tell us about language production. In M. Goldrick, V. Ferreira, & M. Miozzo (eds.), *The Oxford Handbook of Language Production*, pp. 165-181. Oxford: Oxford University Press.
- Liceras, J. M., R. F. Fuertes, S. Perales, R. Pérez-Tattam, & K. T. Spradlin. 2008. Gender and gender agreement in bilingual native and non-native grammars: A view from child and adult functional-lexical mixings. *Lingua*, 118, 827-851.

- Liceras, J. M., K. T. Spradlin, & R. F. Fuertes. 2005. Bilingual early functional-lexical mixing and the activation of formal features. *International Journal of Bilingualism*, 9, 227-252.
- Lohndal, T. 2012. Without Specifiers: Phrase Structure and Argument Structure. Doctoral dissertation, University of Maryland.
- Lohndal, T. 2014. *Phrase structure and argument structure: A case-study of the syntax semantics interface*. Oxford: Oxford University Press.
- Lohndal, T., & M. Westergaard. 2016. Grammatical Gender in American Norwegian Heritage Language: Stability or Attrition? *Frontiers in Psychology*. DOI: <http://dx.doi.org/10.3389/fpsyg.2016.00344>
- MacSwan, J. 1999. *A minimalist approach to intrasentential code switching*. New York: Garland.
- MacSwan, J. 2000. The architecture of the bilingual faculty: evidence from intrasentential code switching. *Bilingualism*, 3, 37-54.
- MacSwan, J. 2005. Codeswitching and generative grammar: A critique of the MLF model and some remarks on “modified minimalism”. *Bilingualism: Language and Cognition*, 8, 1-22.
- MacSwan, J. 2009. Generative approaches to codeswitching. In B. E. Bullock & A. J. Toribio (eds.), *Cambridge Handbook of Linguistic Codeswitching*, pp. 309-335. Cambridge: Cambridge University Press.
- MacSwan, J. 2014. Programs and proposals in codeswitching research: Unconstraining theories of bilingual language mixing. In J. MacSwan (ed.), *Grammatical theory and bilingual codeswitching*, pp. 1-33. Cambridge, MA: MIT Press.
- Mahootian, S. 1993. A null theory of code switching. Doctoral dissertation, Northwestern University.

- Marantz, A. 1997. No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon. In A. Dimitriadis, L. Siegel, C Surek-Clark, & A. Williams (eds.), *Proceedings of the 21st Penn Linguistics Colloquium*, pp. 201-225. University of Pennsylvania, UPenn Working Papers in Linguistics.
- Marantz, A. 2013. Verbal argument structure: Events and participants. *Lingua*, 130, 152-168.
- Merchant, J. 2015. On ineffable predicates: Bilingual Greek-English code-switching under ellipsis. *Lingua*, 166, 199-213.
- Moro, M. Q. 2014. The semantic interpretation and syntactic distribution of determiner phrases in Spanish-English codeswitching. In J. MacSwan (ed.), *Grammatical theory and bilingual codeswitching*, pp. 213-226. Cambridge, MA: MIT Press.
- Muysken, P. 2000. *Bilingual Speech. A Typology of Code-Mixing*. Cambridge: Cambridge University Press.
- Parafita Couto, M., A. Munarriz, I. Epelde, M. Deuchar, & B. Oyharçabal. 2015. Gender conflict resolution in Spanish-Basque mixed DPs. *Bilingualism: Language and Cognition*, 18, 304-323.
- Pfaff, C. W. 1979. Constraints on Language Mixing: Intrasentential Code-Switching and Borrowing in Spanish/English. *Language*, 55, 291-318.
- Pierantozzi, C. 2012. Agreement within early mixed DP: What mixed agreement can tell us about the bilingual language faculty. In C. Gabriel & K. Braunmüller (eds.), *Multilingual Individuals and Multilingual Societies*, pp. 137-152. Amsterdam: John Benjamins.
- Poplack, S. 1980. "Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL": Toward a typology of code-switching. *Linguistics* 18, 581-618.

- Poplack, S., D. Sankoff, & C. Miller. 1988. The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics*, 26, 47-104.
- Pylkkänen, L. 2008. *Introducing Arguments*. Cambridge, MA: MIT Press.
- Ramchand, G. 2008. *Verb Meaning and the Lexicon: A First Phase Syntax*. Cambridge: Cambridge University Press.
- Riksem, B. R. In press. Language Mixing in American Norwegian Noun Phrases. *Journal of Language Contact*.
- Riksem, B. R. 2017. Language Mixing and Diachronic Change: American Norwegian Noun Phrases Then and Now. *Languages* 2, doi:10.3390/languages2020003
- Riksem, B. R., M. B. Grimstad, T. Lohndal, & T. A. Åfarli. To appear. Language mixing within verbs and nouns in American Norwegian. *Journal of Comparative Germanic Syntax*.
- Sankoff, D., & S. Poplack. 1981. A Formal Grammar for Code-Switching. *International Journal of Human Communication*, 14, 3-46.
- Shim, J. Y. 2013. Deriving Word Order in Code-Switching: Feature Inheritance and Light Verbs. Doctoral dissertation, The City University of New York.
- Valdés Kroff, J. R. 2016. Mized NPs in Spanish-English bilingual speech: Using a corpus-based approach to inform models of sentence processing. In R. E. Guzzardo Tamargo, C. Mazak & M. C. Parafita Couto (eds.), *Spanish-English codeswitching in the Caribbean and the U.S.*, 281-300. Amsterdam: John Benjamins.
- Valdés Kroff, J. R., P. E. Dussias, C. Gerfen, L. Perrotti, & T. Bajo. 2016. Experience with code-switching modulates the use of grammatical gender during sentence processing. *Linguistic Approaches to Bilingualism*, 7, 163-198.
- van Hout, A. 1996. Event Semantics and Verb Frame Alternations. Doctoral dissertation [TILDIL Dissertation Series], Tilburg University.

¹ We are grateful to audiences at a range of conferences around the world for feedback on various incarnations of ideas contained in this paper. We would also like to thank the two anonymous reviewers for constructive feedback that we believe made the paper better.

² We will not discuss how to capture the fact that certain patterns of mixing are more common than others. For an approach that incorporates a probabilistic distribution of attested forms, see Goldrick et al. (2016).

³ In this article, we will use the broader term “language mixing” instead of code-switching. This is mainly to set aside the issue of how to distinguish between code-switching and nonce-borrowing; see Grimstad, Lohndal & Åfarli (2014) and Grimstad (2017) for discussion.

⁴ For MacSwan, there is an important distinction between code-switching and nonce borrowing. As mentioned, we use “language mixing” as a more theoretically neutral and descriptive label.

⁵ It is important to note that minimalism is a program, hence it is hard to accurately characterize what minimalism as such adopts. Nevertheless, we believe that it is fair to say that most practitioners adopt a lexicalist feature-driven view of derivations.

⁶ Notice that (9a)/(10) exhibit double definiteness, i.e. definiteness realized both by a determiner and by a functional suffix on the noun, which is characteristic for Norwegian DPs, see Julien (2003, 2005) for discussion. Still, the definiteness feature is not shown in the representation (10) (nor in our subsequent representations of the Norwegian DP) for expository purposes, since what we concentrate on here is the logic of Moro’s (2014) analysis of DP internal mixing, where definiteness is left out of consideration. In section 6, double definiteness will play a crucial role in motivating our exoskeletal structure for the DP.

⁷ We have not indicated what the Norwegian counterparts of the English words would be, but they are very different from the English ones.

⁸ A few examples of this are: *choiren*, M (coon_valley_WI_07gk) – *koret*, N (Norwegian); *ferryen*, M (harmony_MN_04gm) – *ferga*, F (Norwegian); *et tittel*, N (stillwater_MN_01gm) – *ein tittel*, M (Norwegian).

⁹ An alternative approach could be to argue that the speaker has internalized two lexical items for each English noun: one with gender and one without gender. We assume that this analysis is implausible, in part because some speakers of American Norwegian vary in their gender assignment (Lohndal & Westergaard 2016).

¹⁰ From the perspective of (European) Norwegian, we would, in data like (13) and (14), expect (near) obligatory double definiteness, that is, we would expect the English nouns to have a post-nominal Norwegian definite suffix as well. This is what we find in earlier varieties of American Norwegian, but in contemporary varieties of

American Norwegian, we find double definiteness only in approximately half of the relevant cases . The remaining half lacks either the suffix or the determiner, which probably is an effect of attrition. Space prevents us from discussing this further, but see Riksem (2017).

¹¹ We will not go into the discussion of roots (their nature, when they are inserted, etc.) in this article. Note, however, that the NP is a simplification and most likely contains more structure – for example a nominalizing phrase, nP, with a bare root in its complement space, as linguists working within Distributed Morphology would typically assume.

¹² Please note that although we will speak of an English or a Norwegian structure in the remainder of this article, this is merely informal: We do not assume any "language features" whereby syntactic features are annotated for language. Rather, we assume that the "language mode" of the speaker will determine what kind of features are selected as the basis for the abstract syntactic structure. The idea is that the speaker is attempting to speak a certain language, e.g., American Norwegian, and that this manifests itself by virtue of the overall structure of the sentence mimicking or resembling this language. Depending on the features, different structures and thereby potentially different morphosyntactic realizations will be produced. We set aside the precise implementation of this important issue for future research.

¹³ Note that this model would work just as well for the data in this paper if the smallest lexical building blocks for the syntax were uncategorized roots merged with a categorizer. For ease of exposition, we will assume that the mixed unit is a categorized stem, but see see Riksem et al. (to appear) for a discussion of whether these mixed items below word-level in American Norwegian are roots or stems.

¹⁴ Another possible solution would be to assume that the structure below D is Norwegian, so that the feature bundle in F is unchanged, whereas the one in D contains only the NUM feature. Thus, the feature bundle in D allows insertion of the English determiner. We have a few examples of structures with English determiner, English or Norwegian noun stem, and Norwegian definiteness suffix, i.e., double definiteness, such as *the roaden*, "the road" and *the andre dagen*, "the other day". However, we are talking about only a couple of examples found in the corpus, so we will leave this issue aside for now (see Riksem 2017). The key takeaway is that both of these solutions are compatible with an exoskeletal analysis, and both possibilities may also co-exist in the contemporary American Norwegian speech community.

¹⁵ Note that MacSwan himself actually escapes the whole issue of word-internal mixing by claiming that they are not proper examples of code-switching (or language mixing) at all. Instead, they are what he calls "nonce borrowings" (after Poplack, Sankoff and Miller 1988), and they come about by being copied from one mental

lexicon to the other, thus receiving the appropriate feature bundle. This solution is not falsifiable (unless neural imaging one day were to show that that is not how the brain does it), but for an extensive discussion of this and other positions, see Grimstad, Lohndal & Åfarli (2014) and Grimstad (2017).

Furthermore, if one assumes with Distributed Morphology and other theories that syntax operates “all the way down”, i.e., word-internally, then words are constructed in the syntax in the sense that the stem (or root) is inserted in one position, the inflectional morpheme in another, and syntactic or post-syntactic operations ensure that they combine or appear adjacent to one another. From this perspective, the American Norwegian data can be easily accommodated, as we will show below.

¹⁶ We will not go into technicalities concerning checking vs. valuation here. For the purpose of this section, we just notice that somehow it must be ensured that the lexically given tense of the verb itself will be identical to the corresponding tense specified in T.

¹⁷ Note that even though the written Norwegian standards, Bokmål and Nynorsk, both have tense suffixes for all weak verb forms, many of those are omitted in several Norwegian dialects. We have not checked this further (and as these are heritage speakers, placing them dialect-wise is often not possible); but we can assume that several of these 82 verb forms without Norwegian tense suffixes in fact are correct, dialectal forms of Norwegian, reducing this group even further. See Grimstad (2017) for further discussion.

¹⁸ This hypothesis is further elaborated on and supported in Grimstad (2017).

¹⁹ Note that the VoiceP between VP and TP in fact mirrors the FP between NP and DP, and that just like we saw for NP, VP is a simplification which most likely contains more structure. See Harley (1995), Alexiadou et al. (2006, 2015a), Folli & Harley (2007), Pylkkänen (2008), Ramchand (2008), and others.

²⁰ Additional support for the exoskeletal analysis comes from argument structure patterns in the verbal domain. These patterns are also unexpected given a lexicalist analysis, see Åfarli (2015a: 18-20) for discussion.