

The Functional Projections of Subject Splits

Mary S. Linn
The University of Kansas

Sara Thomas Rosen
The University of Kansas

1. Case and Agreement Checking

Linguistic theory (Minimalism and its precursors) by and large assumes that all subjects in all languages are syntactically equivalent. The theory designates one position in the syntax for grammatical subjects, and mandates that all morphosyntactic processes related to that subject (e.g., agreement and case) take place in that one position.

The standard theory also assumes that agreement consists of a single bundle of features. The verb agrees with the subject in person (first, second, or third), number (singular, dual, plural) and gender (e.g., masculine, feminine, neuter). The formal features associated with the functional heads must “check” against the person, number and gender features inherent to the subject, ensuring that the subject and verb properly agree. In particular, it is assumed that case and agreement are checked in Spec, AgrP (or in a second Spec position of T depending upon one’s approach to agreement checking). Under the assumption that an Agr node mediates agreement checking, it is generally assumed that accusative case and object agreement are checked in Spec, Agr-oP, and nominative case and subject agreement are checked in Spec, Agr-sP. The Minimalist Program (MP) makes a couple of assumptions critical to our focus in this paper. First, MP assumes that ϕ features (person, number, and gender) are an inseparable bundle, all selected by a single functional head. Second, MP assumes that all subjects must check ϕ features at some point in the derivation.

However, it appears that the facts in some languages are not so simple. The data presented in this paper indicate that this standard view of agreement is wrong. In particular, we argue that subjects check their ϕ features in different positions depending upon person. We will also show that some languages separate out person features and gender features as independent features, and that some subjects check only person features, while others check only gender features.

2. The Existence of Subject Splits

Preliminary research into American Indian languages shows considerable divergence in their treatment of subjects. Languages from differing language families treat subjects in multiple ways, both in the verbal morphology and in the syntax. Languages may treat first and second person subjects differently from third person subjects, and animate subjects differently from inanimate subjects. We will call any differential treatment of subjects a subject split.

Subject splits are revealed in several ways. There can be a lack of agreement marker, as in Lakhotia, a Siouan language (Rood & Taylor 1996). This is traditionally called a zero-morph. Alternatively, the agreement markers may differ in Case agreement, as in Wichita, a Caddoan language. In Wichita, for example, 1st and 2nd person agreement is marked with agent/patient relations, but the 3rd person is marked with ergative/absolute relations (Rood 1996). Importantly, both of these languages mark 1st and 2nd person differently from 3rd person.

Subject splits can also be seen in the position of the agreement morphology. In Algonquian languages, the position of agreement markers switches according to the relative animacy of the arguments. When two arguments occur, the argument higher in animacy appears as a prefix on the verb; the argument lower in animacy appears as a suffix (Wolfart 1996 for Cree, Berardo 1999 for Shawnee). Athapaskan languages provide another good example of positional differences. Athapaskan has a very complex verb morphology, including up to ten conjunct prefixes. We will not attempt to furnish all the prefixes possible across the Athapaskan languages, but rather will give the general idea of the prefixes available and the relative positions that they appear in. A detailed description of the Athapaskan verb is provided in Rice (2000), based primarily on Slave. Generalizing and simplifying a bit, the verb prefix slots are given in (1). In Athapaskan, the 3rd person agreement marker appears in a different position on the verb from that of the 1st and 2nd person marker (Rice 2000 for Slave). The 1st and 2nd person appears next to verb stem, but the 3rd person appears outside tense/aspect next to object marker, as shown in the template in (1).

- (1) A template of Athapaskan verb morphology
 Obj + (3rd**Subj**) + lexical qualifiers + tense/aspect + (1st/2nd**Subj**) + classifier + V + aspect

As can be seen from the template in (1), lexically specified material appears intermixed with agreement and tense/aspect markers. The classifier is part of the lexical entry for the verb, and a particular classifier must always occur with a particular verb. The lexical qualifiers (called “themes” in the Athapaskan literature) are also lexically determined. Incorporated nouns or incorporated postpositions may occur before the object marker.

What is apparent from the template is that the nature of agreement is complex. First, when the subject is 1st or 2nd person, the verb agrees with that subject in person and number (singular, dual, plural). There is no gender distinction in 1st or 2nd person verbal agreement (as is the case for the majority of languages). When the subject is 3rd person, the verb agrees with it in gender and number, and there is no person feature. Second, agreement with a 1st or 2nd person subject is obligatory. The 3rd person agreement marker is optional, and tends to appear only in the dual or plural. Finally, 1st and 2nd person agreement appears relatively close to the verb stem, with only the classifier appearing between the two. 3rd person agreement appears quite far from the verb stem, outside all tense and aspect markers, just inside the object agreement marker.

Rice (2000) shows the positional difference in Slave between 1st/2nd person subject agreement and 3rd person subject agreement with the examples below. We have shown the agreement marker in bold in (2). These examples include a lexical qualifier *-n(V)-*, here indicated in italics. Notice that the first and second person marker invariably appears to the right of the lexical qualifier. The third person singular in this example has no overt marker, and the both the third person dual/plural and the impersonal agreement markers appear to the left of the qualifier. It appears from the data that there are two subject positions—one for 1st and 2nd person, and a different one for 3rd person.

- (2) Agreement in Athapaskan
 “hide” (optative)
 ní-né-o-**h**-’í 1st singular *h*
 ní-né-**q**-’í 2nd singular nasalization
 ní-n-**úú**-’í 1st dual/plural *úú*
 ní-ná-**ah**-’í 2nd dual/plural *ah*
 ní-né-o-’í 3rd unmarked
 ní-**ke-n-é**o-’í 3rd dual/plural *k*
 ní-**ts**’e-né-o-’í unspecified *ts*
 (Rice 2000: 182-3)

Finally, subject splits can appear with differential case patterns, not just differential agreement patterns. Case splits appear in languages as diverse as the Mayan and Australian languages in which the subject of only some sentences is nominative case; ergative otherwise (Garcia & Guajan 1999 for Kaqchikel Mayan, Dixon 1994 for Dyirbal). In Dyirbal, 1st and 2nd person subjects receive nominative case, whereas 3rd person subjects receive ergative case.

- (3) a. **ngaja** paninyu
I-nom come-fut
 “I’m coming.”
- b. **ngaja** nginuna palkan
I-nom you-acc hit-nfut
 “I’m hitting you.”
- c. **ngaja** payiyarapalkan
I-nom there-absman-abs hit-fut
 “I am hitting man.”
- (4) a. ngaykuna pangkul **yarangku** palkan
 I-acc there-erg **man-erg** hit-nfut
 “Man is hitting me.”
- b. payiparrkan pangkul **yarangku** jurrkanyu
 there-abswallaby-abs there-erg **man-erg** spear-nfut
 “Man is spearing wallaby.”
 (Dixon 1972: ex (28-33) 64)

In looking at Walpiri and Inuit, Bittner and Hale (1996) have argued that nominative case is checked in Spec, IP (AgrsP for us), whereas ergative case is checked inside VP. Bittner and Hale are essentially arguing that there is a positional distinction between nominative and ergative case. If this is correct, then 3rd person subjects in Dyirbal are checked lower in the structure than 1st and 2nd person subjects.

From these data, we conclude that subject marking is not necessarily uniform within a language; “subject splits” occur across languages and prevailing theories neither predict nor explain them in any natural fashion. Not only are subjects treated differently across languages, they are not necessarily uniform within a language. While the existence of subject splits in general is widely known and recognized (e.g., Rice 2000, Rice & Saxon 1994, Dixon 1994, Mithun 1991,

Bittner and Hale 1996, Rhodes 1990), there has been no cross-linguistic theoretical integration of the facts. What is not known is whether subject splits follow general patterns. Linguistic theory must seek out the patterns, and analyze them if we are to understand the agreement mechanisms and the representation of subjects in the linguistic system.

3. Position of the Subject and Object in Euchee

In order to delve into the mechanics of subject splits, we will look at agreement in Euchee. Euchee is a language isolate, originally spoken in the North American Southeast, but is now spoken in Oklahoma. Like other languages of the Americas, it is largely polysynthetic with a rich verbal agreement. It is also an active/stative language (or split-intransitive language), meaning that the core arguments of intransitive verbs are marked either with the agent marker, if the verb is an event, or the patient marker, if the verb is a state.

On the surface, the agreement morphology appears to have an Object (patient)-Subject (actor)-Verb order, given as a preliminary template in (5).

- (5) A template of Euchee verb morphology—**FIRST PASS**
Obj + Subj + V + aspect + tense

This order is clearly seen when participants are 1st and 2nd person, as in (6).

- (6) a. $h\ddot{o}-\mathbf{di}-'ne \rightarrow$ 1st person subject = O-S-V
him(obj)-I(subj)-see
'I see him.'
- b. $nedze-\mathbf{di}-'ne$
you(obj)-I(subj)-see
'I see you.'
- c. $h\ddot{o}-\mathbf{ne}-'ne \rightarrow$ 2nd person subject = O-S-V
him(obj)-you(subj)-see
'You see him.'
- d. $dze-\mathbf{ne}-'ne$
me(obj)-you(subj)-see
'You see me.'
(Linn 2001)

It becomes more complicated, however, when the subject (actor) is in the 3rd person. The relative positions of the subject and object agreement markers are flipped when the subject (actor) is in the 3rd, as the data in (7) show.

- (7) a. $h\ddot{o}-dze-'ne \rightarrow$ 3rd person subject = S-O-V
he(subj)-I(obj)-see
'He sees me.'
- b. $h\ddot{o}-nedze-'ne$
he(subj)-you(obj)-see
'He sees you.'
(Linn 2001)

In order to account for this, we might revise the template in (5), and posit a template as in (8), where the 3rd person subject is separate from 1st and 2nd person subjects. This would make the Euchee agreement system somewhat parallel to the Athapaskan data.

- (8) A template of Euchee verb morphology—**SECOND PASS**
3Subj + 1/2/3Obj + 1/2Subj + V

Unfortunately, this order does not work in Euchee. The data in (9) illustrate the agreement markers when both arguments are 3rd person. Euchee does not allow two 3rd person markers. If both arguments are 3rd, the verb agrees with the subject only; the object must appear as an independent phrase, and no verbal agreement marker appears for the object argument.

- (9) a. *sedi hõ'ne*
 her he-see
 'He sees her.'
- b. **se-hõ-'ne*
 her(obj)-he(subj)-see
 'He sees her.'
- c. **hõ-se-'ne*
 he(subj)-her(obj)-'ne
 'He sees her.'
 (Linn 2001)

This fact cannot be accounted for if there are two slots in the template, one for the 3rd person subject and one for the (3rd person) object.

In addition, recall that Euchee has active/stative marking. Active/stative marking appears only on 1st and 2nd person arguments. The agreement for the event roles actor and patient are only found on the 1st and 2nd persons, as seen in (10). In (10a) 1st and 2nd get event role-marking. The generalization seems to be that 1st and 2nd person arguments get event role-marking; 3rd person gets gender marking.

- (10) a. Active marking in 1st and 2nd
- | | |
|---------------------------------|----------------------|
| di the | 'I ran' |
| 'õ the | 'we (inclusive) ran' |
| nõ the | 'we (exclusive) ran' |
| ne the | 'you ran' |
| 'ane the / 'ã the | 'you (plural) ran' |
- b. Stative marking in the 1st and 2nd (bi-morphemic)
- | | |
|--------------------|----------------------------|
| dze s'i'e | 'I am short' |
| 'õdze s'i'e | 'we (inclusive) are short' |
| nõdze s'i'e | 'we (exclusive) are short' |
| nedze s'i'e | 'you are short' |
| 'adze s'i'e | 'you (plural) short' |

- c. No event role; gender in 3rd person

hõ the	‘he ran’
hõs ’i’eõ	‘he is short’
* hõdzes ’i’eõ	‘he is short’
se the	‘she ran’
ses ’i’e	‘she is short’

- d. No inanimate subject (agent) of an event verb.

*hithe	‘it ran’
hish’o	‘it is withered’

(Linn 2001)

Interestingly, Euchee does not allow inanimate subjects of eventive verbs, as the example in (10d) shows. Our assumption is that an event must have initiation, and the subject plays the role of initiating events. An inanimate argument is by definition a non-initiator (cannot be an actor, for example), and therefore there can be no inanimate actors in Euchee.

To sum up, since Euchee does not allow 3rd person agreement markers, and since 3rd person objects do not show agreement, we posit that there is only one 3rd person slot in the Euchee verbal template. A final revised template appears in (11).

- (11) A template of Euchee verb morphology—**FINAL**
3 + 1/2Obj + 1/2Subj + V

The following facts emerge from our discussion of Euchee:

- i. First and second person agreement markers appear in a different position in the verb from third person agreement markers.
- ii. First and second person agreement is inflected for person and number but not gender.
- iii. Third person agreement is inflected for gender and number but not person.
- iv. When both arguments are 3rd person, the verb agrees with the subject only; the object must appear as an independent DP.
- v. Active/stative marking appears on 1st and 2nd person arguments.

4. Hypotheses

In order to account for the facts listed above, we propose, following Ritter and Rosen (2000, 2001) and others, that there are languages that organize their arguments along an animacy hierarchy. In such languages, the positions SUBJECT and OBJECT have less significance in the organization of the clause and the agreement markers than animacy. We provide a standard animacy hierarchy in (12) for the sake of reference. Notice that the hierarchy is based on animacy, person, and type of NP. It seems that languages that are sensitive to animacy make a binary split somewhere along the continuum in (12) between animate and inanimate arguments. The split may be almost anywhere on a hierarchy like that in (12).

- (12) Croft’s (1990: 112) Animacy Hierarchy (based upon Silverstein)
1st/2nd person pronoun > 3rd person pronoun > proper name > human common noun > non-human animate common noun > inanimate common noun

The person hierarchy, as pointed out by Croft (1990: 113) plays an important role in the expression of subjects and objects in many languages, and it represents the most radical difference between Subject/Object marking language and other languages in the world.

From these facts given above, we make the following hypotheses:

- i. Subjects check person features in one position and gender features in a separate position; number may appear with either person or gender.
- ii. First and second person have person features but no gender features.¹
- iii. Third person has gender features but no person features (Benveniste 1956, Noyer 1997).

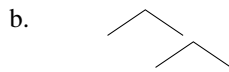
The hypotheses have the following consequences for the structure of a language like Euchee, and is schematized in (16) below:

- i. Subjects that are animate have person features. Subjects that are inanimate have gender features, and these features are checked in a Gender Phrase.
- ii. There are two functional projections where subjects may be checked—Agr for subjects with person features, and Gen for subjects with gender features.²
- iii. TP may appear between AgrP and GenP (in Athapaskan tense and aspect comes between the two subject positions), or may dominate AgrP.

We assume, along with Baker (1996), that the pronominal markers on the verb are agreement, and not, for example, the arguments themselves. Along these lines, the actual verbal arguments appear as *pro* in the relevant specifier positions. In line with Baker's claims, all independent noun phrases are adjuncts.

In order to derive the Euchee morpheme and word order, we also assume that the morphology is represented as a mirror image of the syntax. This is known as Mirror Image projection and checking (Baker 1995, Pollock 1997). Pollock (1997: 257) proposes that in a checking theory, functional features are visible only on the outer layers of the morphology. After a feature is checked and deletes, the next layer is visible to be checked. In his approach, checking requires functional projections to be a mirror image of the morphology on the verb, as schematized in (13).

(13) a. [[]]



The example in (14) shows Pollock's proposal for the relation between the morphology and the syntax for agreement and tense in French. In this example, Pollock argues that because the agreement marker appears outside the tense marker, it must be lower in the syntactic structure.

¹Semitic languages have gender in the 2nd person. Our proposal is that gender is used to check animacy requirements of 3rd persons. Since 2nd person is inherently animate, there is no need for gender to be a separate projection for 2nd person.

²Rice and Saxon (1994) also propose two functional projections for the subject in Athapaskan. Their analysis is compatible with ours. One difference is that Rice and Saxon propose a Number Phrase rather than a Gender Phrase. Given that 1st, 2nd and 3rd all include number distinctions, but not all include gender distinctions, we propose that gender be separated from person rather than number.

When the verb raises to Agr-s, it will check the agreement, it's formal feature will delete, and the tense marker will be visible for checking. If the visibility approach to the formal features is correct, then the Agr node must be lower than T in French.

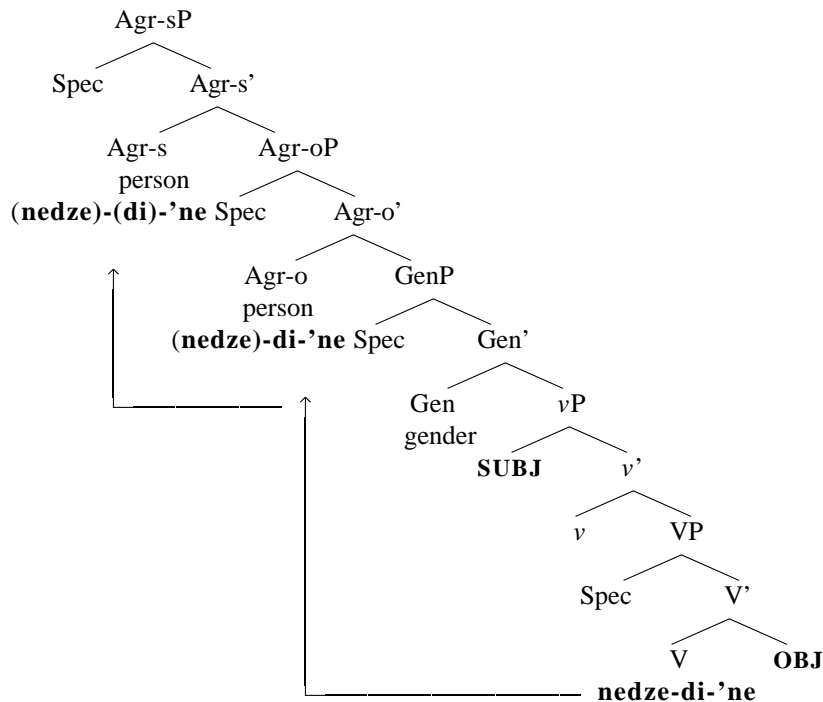
- (14) a. Nous parlerons (French)
'We will speak.'
- b. [[[_{root} parl] -er_{tense/mood}] -ons_{Agrs}] morphological structure
[_{TP} NP T]_{AgrP} Agr [_{VP} -] syntactic structure

If we apply Pollock's approach to Eucree, we get a structure like that in (15). Agr-s must dominate Agr-o, which in turn dominates Gender.

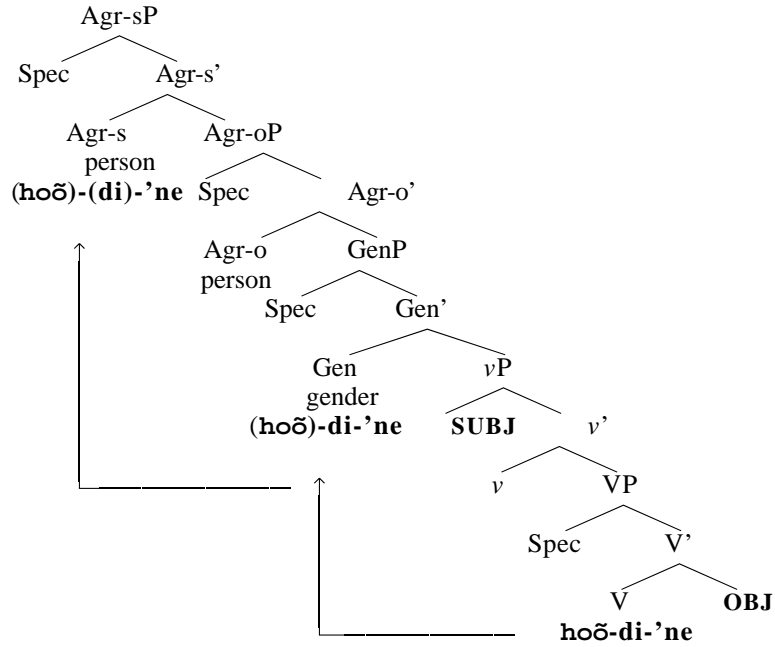
- (15) a. I see you. (Eucree)
[_{gen} [_{AgrP} nedze- [_{Agrs} di- [_{root} 'ne] morphological structure
[AgrsP [AgrP [GenP [VP]]]] syntactic structure
- b. He sees me.
[_{Gen} ho- [_{AgrP} dze- [_{Agrs} [_{root} 'ne] morphological structure
AgrsP [AgrP [GenP [VP]]]] syntactic structure

Thus, a separate gender phrase (GenP) between vP and Agr-o, and using the Mirror Principle, derives the agreement morpheme order in Eucree. This can be seen in examples (16)-(19). Each morpheme layer is checked from the inside out, and the parentheses are used to show that a morpheme has been checked and the next layer is visible for checking.

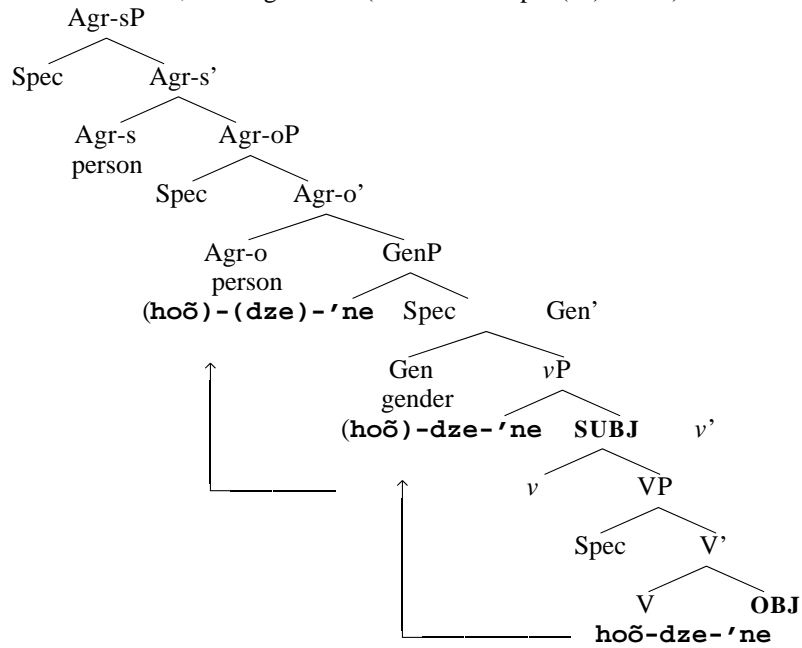
- (16) O-S-V order, 1/2 acting on 1/2 (same as example (6b) above)



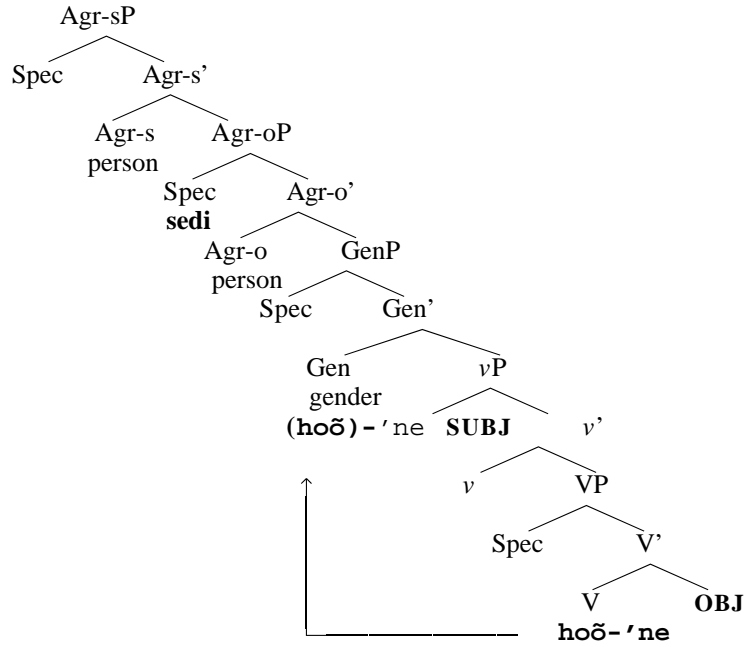
- (17) O-S-V order, 1/2 acting on 3 (same as example (6a) above)



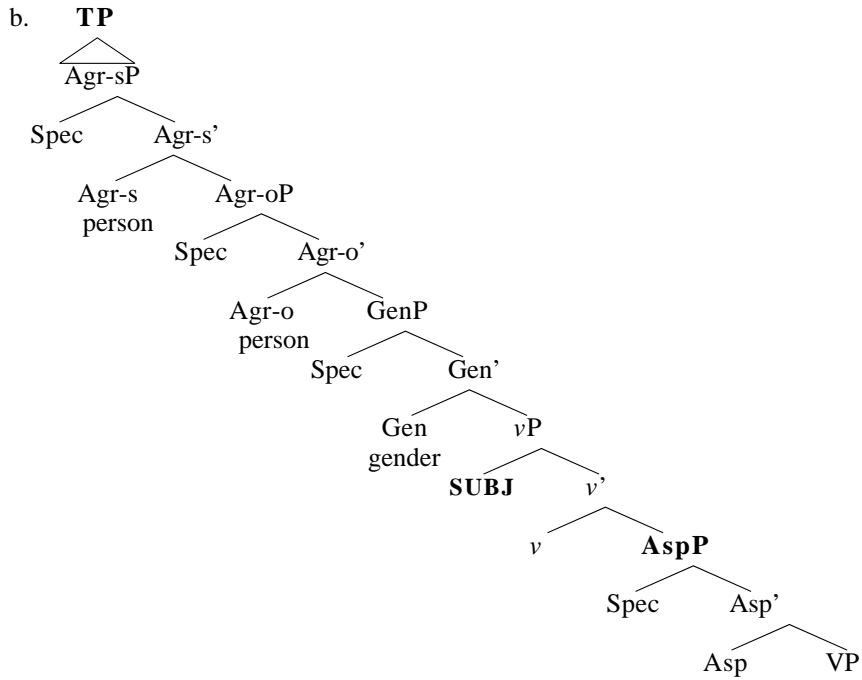
- (18) S-O-V order, 3 acting on 1/2 (same as example (7a) above)



(19) S-V order, 3 acting on 3 (same as example (9a) above)



(20) a. ho-dze-'neõ-ne-jeõ
 he-me-see-HAB-PST.IMP
 'He used to see me.'



While the agreement markers are consistent with mirror image checking, the tense and aspect markers in Euchee are not aligned in mirror image fashion and are more consistent with incorporation. Tense and aspect are suffixed, with an order of V–Aspect–Tense. The placement of the tense and aspect markers are illustrated in (20).

Tentatively, we conclude that agreement is checked, but tense/aspect is incorporated. If this is on the right track, it means that languages can have a mixed system of incorporation for some functional features and checking for others.³

5. Ongoing Inquiry and Conjectures

Without the availability of a functional projection between the *vP* and *Agr-o*, we are not able to account for the Euchee morpheme order. We have argued that this functional projection is a Gender Phrase. Subjects that are animate have person features. Subjects that are inanimate have gender features, and these features are checked in a Gender Phrase. In Euchee, only animate participants may be subjects of events. Gender is used to regulate the animacy of a 3rd person, making an animate 3rd person available for subjecthood.

Event structure in the syntax builds on the aspectual notions of events (activities, accomplishments, and achievements vs. states) as having an object of origin/actualization, the event, and the object of termination (Tenny 1994, Van Voorst 1988, Dowty 1979, Vendler 1967). The initiator is realized as the subject and the termination or delimiter is realized as the object. There has been much research indicating that a verb is read as an event when it has a terminus, or is delimited by a direct object. However, there has been little research conducted on the role, if any, that the initiator plays in eventhood. Ritter and Rosen (2000, 2001), for example, have argued that some languages organize their arguments around quantization of the object and the event, and others do not. They suggest there is a typology of languages based on events in the syntax.

- Delimitation-based languages

- Accomplishments form a natural class with achievements
- Sensitive to specificity of the object
- Accusative case may be restricted on the basis of the object or on the basis of aspect/tense
- Ergative splits on the basis of perfective aspect/past tense
- =>D-languages identify events via the object

The question is whether some languages organize events around the initiator of the event, or the subject. Thus, we suggest that there is a class of animacy based languages (Ritter & Rosen's initiation-based languages):

- Animacy-based languages

- Accomplishments form a natural class with activities (??)
- Sensitive to animacy and agentivity properties of subjects

³We do have a complication concerning the universality of functional projection ordering. The dominance relations of the agreement phrases in Athapaskan would be *AgrsP>GenP>AgroP* if we assume mirror imaging. It is not clear at this stage of our research what to make of such apparent language-particular ordering.

- May make a grammatical distinction between topic and subject
- Ergative splits on the basis of properties of the subject
- Subject and object agreement specified for person features
- Show overt animacy hierarchies
- =>A-languages identify events via the subject

We propose that Euchee and other subject split languages organize arguments around animacy, particularly the animacy of the subject.

References

- Baker, M. C. (1985). The Mirror Principle and Morphosyntactic Explanation. *Linguistic Inquiry* 16, 373-415.
- Baker, M. C. (1996). *The Polysynthesis Parameter*. Oxford University Press, London.
- Benveniste, E. (1956). La Nature des Pronoms. In M. Halle, H.G. Lunt, H. MacLean & C.H. van Schooneveld (eds.), *For Roman Jakobson: Essays on the Occasion of his Sixtieth Birthday*, 24-37. Mouton, The Hague.
- Berardo, M. (1999). Animacy and Shawnee Verbal Inflection. Ph.D. dissertation, University of Kansas, Lawrence.
- Bittner, M. & K. Hale. (1996). Ergativity: Toward a Theory of a Heterogeneous Class. *Linguistic Inquiry* 27, 531-604.
- Chomsky, N. (1995). *The Minimalist Program*. MIT Press, Cambridge, MA.
- Croft, W. (1990). *Typology and Universals*. Cambridge University Press, Cambridge England.
- Dixon, R.M.W. (1994). *Ergativity*. Cambridge University Press, Cambridge, England.
- Garcia, P. M. L., and J. O. R. P. B. Guajan. (1999). *Rukemik ri Kaqchikel chi': Gramatica Kaqchikel*. Editorial Cholsamaj, Guatemala.
- Linn, M. S. (2001). *A Grammar of Euchee (Yuchi)*. Ph.D. dissertation, University of Kansas, Lawrence.
- Noyer, R. R. (1997). *Features, Positions and Affixes in Autonomous Morphological Structure*. New York: Garland Publishing.
- Pollock, J.-Y. (1997). Notes on Clause Structure. In L. Haegeman (ed.), *Elements of Grammar: Handbook of Generative Grammar*. Kluwer Academic Publishers, Dordrecht.
- Rice, K. (2000). *Morpheme Order and Semantic Scope: Word Formation in the Athapaskan Verb*.
- Rice, K., and L. Saxon. (1994). The Subject Position in Athapaskan Languages. In H. Harley and C. Phillips (eds.), *The Morphology-Syntax Connection*, 173-195. Cambridge, MA: *MIT Working Papers in Linguistics* 22.
- Ritter, E., and S. T. Rosen. (2000). Event Structure and Ergativity. To appear in J. Pustejovsky and C. Tenny (eds.), *Events as Grammatical Objects*. CSLI, Stanford.
- Ritter, E., and S. T. Rosen. (2001). The Interpretive Value of Object Splits. To appear in *Language Sciences*.
- Rood, D. S. (1996). Sketch of Wichita, a Caddoan Language. In I. Goddard (ed.), *Handbook of North American Indians*, Volume 17, Languages, 580-608. Smithsonian Institution, Washington.
- Rood, D. S., and A. R. Taylor. (1996). Sketch of Lakhota, A Siouan Language. In I. Goddard (ed.), *Handbook of North American Indians*, Volume 17, Languages, 440-482. Smithsonian Institution, Washington.
- Silverstein, M. (1976). Hierarchy of Features and Ergativity. In R.M.W. Dixon (ed.), *Grammatical Categories in Australian Languages*, 112-71. Australian Institute of Aboriginal Studies, Canberra.
- Wolfart, H.C. (1996). Sketch of Cree, and Algonquian Language. In I. Goddard (ed.), *Handbook of North American Indians*, Volume 17, Languages, 390-439. Smithsonian Institution, Washington.