SEPARATING FEATURE INTERPRETABILITY FROM FEATURE VALUES: EVIDENCE FROM NEGATIVE CONDORD IN JAPANESE[1]

Masakazu Kuno (kuno@waseda.jp)

1. INTRODUCTION

Formal features have two aspects, interpretability (interpretable vs. uninterpretable) and values (valued or unvalued). Feature interpretability is a matter of interface levels. If a given feature is interpretable, it will receive an interpretation at an interface level, and since feature interpretability is a matter of interface levels, syntax cannot inspect feature interpretability. Instead, syntax can inspect feature values: syntax can tell if a given feature is valued or not. When syntax finds an unvalued feature, it will execute an Agree operation by which an unvalued feature is made to search for a matching goal with a valued feature so that the unvalued probe will get valued. Thus, feature interpretability and feature values are two independent aspects of formal features. However, it is widely held that there is a correlation between feature interpretability and feature values in such a way that a feature F is uninterpretable iff F is unvalued. This is largely due to Chomsky's (2000, 2001) treatment of feature interpretability and feature values.

This view is challenged by Pesetsky and Torrego (2007) (P&T henceforth), who argue that feature interpretability should be separated from feature values. (1) illustrates four possible feature types.

(1) a. *i*F val *interpretable, valued*
b. *u*F [ ] *uninterpretable, unvalued*
c. ***i*F [ ] *interpretable, unvalued***
d. ***u*F val *uninterpretable, unvalued***

Chomsky's (2000, 2001) biconditional relation would allow only (1a) and (1b) while P&T's system admits all the four feature types. The feature types that are allowed by P&T's system but not by Chomsky's are (1c) and (1d).

Now let us consider what would be possible derivations if the four feature types are all allowed. One might think that if the four feature types are allowed, they will yield 16 possible derivations. In fact, it is not the case. Even if we allow the four feature types, we will have only two possible derivations. Before distinguishing possible derivations from impossible ones, let us make clear the assumptions that we adopt in the following discussion.

(2) a. Syntax inspects feature values in search of an unvalued feature.
b. An unvalued feature acts as a probe and searches for a valued counterpart feature

[1] I thank the audience at SWIGG 2011 and an anonymous reviewer for comments and questions. Needless to say, all errors are mine.

(goal). When a probe finds a goal, an Agree operation is executed, which results in the valuation of an unvalued feature.

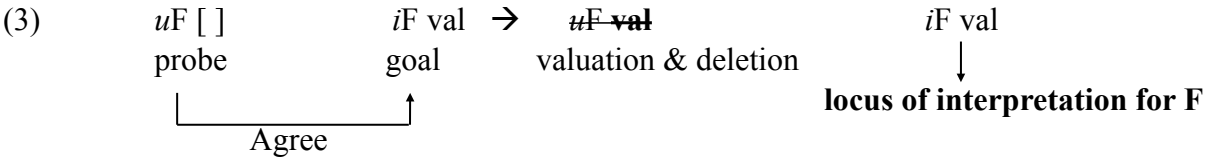
c. An uninterpretable feature must be deleted before it is sent to the interface systems.

d. A feature must be interpreted either on the probe or on the goal, not on both.

With these assumptions in mind, first consider (3). Both Chomsky's system and P&T's allow this derivation, where the probe is an uninterpretable and unvalued while the goal is interpretable and valued.

Possible Derivation 1

(3)

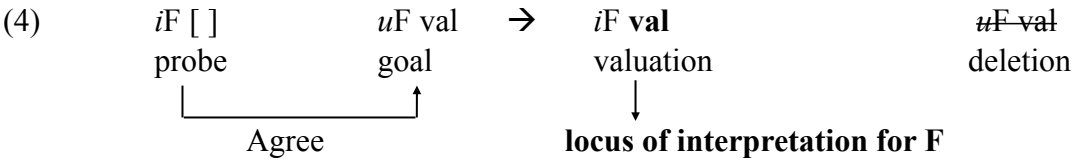


In this derivation, after Agree takes place, the probe gets valued and deleted because it is uninterpretable (by (2c)) whereas the goal will constitute "the locus of interpretable for the feature F" because it is valued and interpretable. In other words, the feature F will receive an interpretation on the goal, rather than on the probe.

The other possible derivation is shown in (4).

Possible Derivation 2

(4)



This derivation is allowed only under P&T's system because the probe is interpretable but unvalued while the goal is uninterpretable but valued. But syntax can establish an Agree relation between them. Since the probe is unvalued, it is made to search for a matching goal, and since the goal is valued, it can assign its value to the probe. As a result, the probe gets valued whereas the goal gets deleted because it is uninterpretable (by (2c)). In this derivation, it is the probe that constitutes the locus of interpretation for the feature F. To put it differently, the feature F will get interpreted on the probe, rather than on the goal.

Now notice that the two derivations in (3) and (4) differ in terms of the locus of interpretation of the feature F. The goal constitutes it in (3) while the probe does so in (4). This is one of the consequences of P&T's system, i.e., it gives both the probe and the goal a chance to get interpreted at the interface. By contrast, Chomsky's system allows only the goal to get interpreted. This raises an empirical question of whether there is a case in which the probe constitutes the locus of interpretation. In what follows, we will demonstrate that the answer is yes by analyzing Negative Concord phenomena in Japanese.

Before moving on, however, let us consider and preclude other derivations. We will take up four representative cases.

Impossible Derivation 1

(5) *i*F val *u*F val **Probe has be unvalued. (2b)**
probe goal
*Agree

When the probe is valued, syntax will not make it act as a probe.

Impossible Derivation 2
(6) *i*F [ ] *u*F [ ] **Goal has be valued. (2b)**
probe goal
*Agree

The goal has to be valued; otherwise, it cannot assign its value to the probe.

Impossible Derivation 3
(7) *i*F [ ] *i*F val → ******i*F val *i*F val
probe goal valuation
Agree **A feature must be interpreted either on the probe or on the goal, not on both. (2d)**

When both the probe and the goal end up in an interpretable feature, the feature F cannot constitute a single locus of interpretation.

Impossible Derivation 4
(8) *u*F [ ] *u*F val → ***~~*u*F val~~ ~~*u*F val~~**
probe goal valuation & deletion deletion
Agree **A feature must be interpreted either on the probe or on the goal, not on both. (2d)**

When both the probe and the goal are uninterpretable, the feature F will get deleted on both sites and will never get interpreted. Derivations in (5) through (8) are thus ruled out.

2. NEGATIVE CONCORD IN JAPANESE

It is well known since Watanabe (2004) that Japanese has Negative Concord Items (NCIs). Typically, NCIs in Japanese take the form of an indeterminate WH-phrase + a focus particle *mo* 'also, even' and have to occur in the presence of a sentential negation marker, as shown in (9) and (10).

Wh-mo type
(9) a. **Dare-mo** ko-**nakat**-ta
who-MO come-Neg-Past
'Nobody came.'
b. * **Dare-mo** ki-ta

who-MO come-Past

(10) a. John-wa kino **nani-mo** tabe-**nakat**-ta
John-Top yesterday what-MO eat-Neg-Past
'John did not eat anything yesterday.'

b. John-wa kino **nani-mo** tabe-ta
John-Top yesterday what-MO eat-Past

Watanabe (2004) convincingly argues that the WH-*mo* expressions are NCIs, not Negative Polarity Items (NPIs). Here we will see two arguments. First, the WH-*mo* expressions can only be licensed by a sentential negation marker. As indicated by the ungrammaticality of (11), the relevant expressions cannot occur in contexts where NPIs would normally be licensed.

(11) a. * **Dare-mo** ki-ta-ra, osiete-kudasai. (Conditional)
who-MO Come-Past-condi tell-please
'If anyone comes, please tell me.'

b. John-wa kino ***nani-mo** tabe-ta-no? (Question)
John-Top yesterday what-MO eat-Past-Q
'Did John eat anything yesterday?'

Another piece of evidence for treating the WH-*mo* expressions as NCIs rather than NPIs is that they can be used as an elliptical answer, which would best be replaced by a negative quantifier, rather than a NPI, as shown in (12) and (13).

(12) Q: **Dare-ka** ki-masi-ta-ka
who-KA come-Polite-Past-Q
'Did anybody come?'

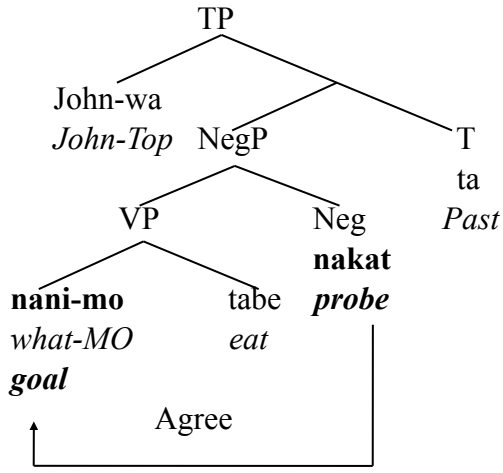
A: **Dare-mo**
who-MO
'Nobody.' (= Nobody came.)

(13) Q: Ohiru-ni **nani-ka** tabe-masi-ta-ka
Lunch-for what-KA eat-Polite-Past-Q
'Did you eat something for lunch?'

A: **Nani-mo**
what-MO
'Nothing. (= I did not eat anything.)'

Given that a WH-*mo* expression needs to be licensed by a sentential negation (SN), the relation between a WH-*mo* expression and a SN can be seen as an Agree relation. The question is exactly how this Agree relation is established. As answer to this question lies in the structural relationship between a NCI and a SN. Consider (14).

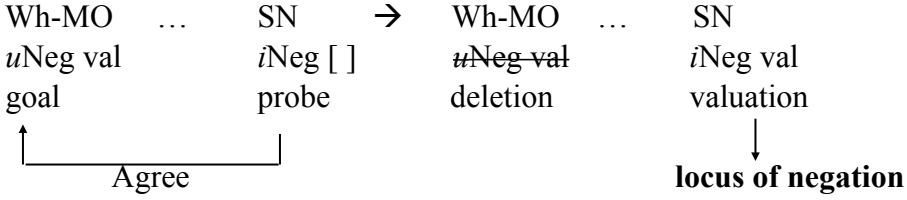
(14)



Given the tree diagram, it is clear that the SN is higher than the NCI, and therefore we can regard the SN as a probe and the NCI as a goal. Then another question arises: what feature is involved. Since the SN is negative by itself and the NCI can be used as an elliptical answer, the Neg-feature is a likely candidate.

We can analyze the probe-goal relation between the SN and the NCI as follows.

(15)



In (15), the SN acts as a probe, hence its Neg-feature has to be unvalued. As for interpretability, we assume that the Neg-feature of a SN is interpretable because the SN generally has the ability to negate sentences. Given this, it follows that the Neg-feature of a WH-*mo* has to be uninterpretable but valued, which falls under the possible derivation seen in (4). When an Agree operation takes place, the resulting structure is the one on the right hand side of the arrow, where the Neg-feature of WH-*mo* has deleted whereas that of the SN constitutes the locus of the interpretation of the Neg-feature. This amounts to saying that in a sentence that contains a WH-*mo* NCI and a licensing SN, it is the SN that receives a negative interpretation.

This conclusion is in contradiction with Watanabe's (2004), who argues that in such a sentence, negation is expressed by a NCI rather than by a SN. In order to show that the present proposal is superior to Watanabe's, let us do a test. The test takes the following form. First, construct a sentence that contains a WH-*mo* NCI and a licensing SN. Then construct another sentence where TP is elided and then check whether the elided TP is interpreted as negative or non-negative. If the locus of negation is constituted by a SN, then the elided TP will be interpreted as negative since the TP in the antecedent sentence contains a projection of a SN. If a SN does not receive a negative interpretation, the elided TP will receive a non-negative interpretation. The point of this test is that we can tell where negation is expressed in the sentence that contains a NCI and a licensing SN by looking into the meaning of an elided material that is licensed by the sentence with the NCI and SN in light of the condition on

ellipsis that dictates that the elided material be semantically identical to its antecedent. With this in mind, consider (16).

(16) a. Watashi-wa daietto-chu-dakara karori-no takai-mono-wa **nani-mo**
I-Top diet-during-because calorie-Gen high-stuff-Top what-MO
tabe-taku-nai.
eat-want-Neg
'Because I'm dieting, I do not want to eat anything with high calorie'

b. Tokuni [$_{CP}$ nani-o [$_{TP}$ Δ (= tabe-taku-nai)] ka] osiete-kudasai
Especially what-Acc eat-want-Neg Q tell-place
'Please tell me especially what (you do not want to eat).'

As shown in (16b), the elided TP must be interpreted as negative. This indicates that the TP of the antecedent sentence (16a) is also interpreted as negative. Thus, we can conclude that the SN in (16a) constitutes the locus of interpretation of the Neg-feature.

The discussion so far supports the derivation of the type illustrated in (4). Japanese has another type of NCI, which consists of a XP + a particle *sika* 'except'. The *sika*-marked XP can also be regarded as a kind of NCIs because it has to occur in the presence of a sentential negation marker, as shown in (17) and (18). In the following discussion, we will focus on the case where *sika* is attached to a NP, but it should be noted that this particle can be attached to other categories as well.

NP-SIKA type

(17) a. John-**sika** ko-**nakat**-ta
John-SIKA come-Neg-Past
'Only John came.' = 'Nobody except John came.'

b. * John-**sika** ki-ta
John-SIKA come-Past

(18) a. John-wa yasai-**sika** tabe-**nai**
John-Top vegetables-SIKA eat-Neg
'John eats only vegetables.' = 'John eats nothing except vegetables.'

b. John-wa yasai-**sika** tabe-**nai**
John-Top vegetables-SIKA eat-Neg

Notice that the sentences that involve a *sika*-marked NP and a licensing SN are interpreted as positive propositions. This suggests that the SN involved in these sentences is in fact non-negative. We will return to this point soon.

What is the meaning of *sika*? In Japanese, this particle has no other usage and native speakers of Japanese do not have clear intuition about its meaning. They only know that it means something like 'except' when it is used together with a SN. Thus, we propose that NP-*sika* means 'nothing/nobody except NP', which is equivalent to 'only NP'. This proposal is supported by the Korean counterpart of *sika*, which is *pakkey*, because the latter can be decomposed into 'outside-Dat'. Take a look at (19) for concrete examples and (20) for the analysis of *pakkey*.

Korean

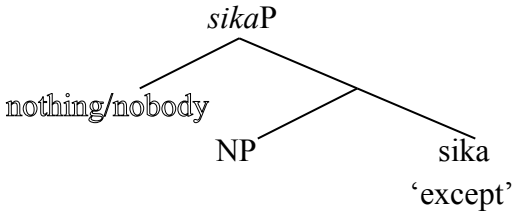
(19) a. Swuni-ka ku chayk-**pakkey** ilk-ci **anh**-ass-ta
Swuni-Nom that book-except read-Comp Neg-Past-Decl
'Swuni read only that book.' = 'Swuni did not ready anything except that book.'

b. Swuni **pakkey** Seoul-ey **anh**-ka-ss-ta
Swuni-except Seoul-to Neg-go-Past-Decl
'Swuni went to only Seoul.' = 'Swuni did not go anywhere except Seoul.'

(20) pakkey = pack-ey
outside-Dat

In view of the meaning of NP-*sika*, it is not unreasonable to suppose (21).

(21) NP-*sika* comes with a silent negative quantifier (NQ) and *sika* itself means 'except.'



The idea is that a *sika*-marked NP contains a phonologically null negative quantifier and the *sika*-marked NP expresses an exception to the domain of quantification.

We are now ready to analyze the structures that involve a *sika*-marked NP and a licensing SN. The analysis is shown in (22).

(22)

NQ NP-sika …	SN	→	NQ NP-sika …	SN
*i*Neg val	*u*Neg [ ]		*i*Neg val	~~*u*Neg val~~
goal	probe			valuation & deletion
	Agree		**locus of negation**	

Here the Neg-feature of the SN is unvalued by assumption (in order to act as a probe) as well as uninterpretable because of the lack of negativity in the interpretation (see the discussion above). On the other hand, the goal is analyzed as bearing an interpretable and valued Neg-feature because it comes with a silent negative quantifier. Thus, this derivation falls under the type of derivation schematized in (3). After Agree is executed, the Neg-feature of the probe is valued and deleted whereas the goal constitutes the locus of interpretation.

As one may have noticed, this analysis leads us to predict that when a sentence that involves a *sika*-marked NP and a licensing SN antecedes an elided TP, the elided TP must be interpreted as non-negative because the SN in the antecedent does not constitute the locus of interpretation for the Neg-feature. This prediction is borne out. Consider (23).

(23) a. Watashi-wa daietto-chu-dakara karori-no hikui-mono-**sika** tabe-taku-**nai**.
I-Top diet-during-because calorie-Gen low-stuff-Top eat-want-Neg

'Because I'm dieting, I do not want to eat anything except low calorie food.'

b. Tatoeba [CP nani-o [TP Δ (= tabe-tai)] ka] osiete-kudasai
for example what-Acc eat-want Q tell-place
'Please tell me, for example, what (you want to eat).'

Just as predicted, the elided part of (23b) must be interpreted as non-negative. The contrast between (23) and (16) has hitherto been unnoticed and thus offers significant insight to look into the difference between WH-*mo* NCIs and NP-*sika* NCIs.

Let us now turn to the case where multiple NCIs occur. The first case to consider is the one in which two WH-*mo* NCIs occur in the same structure, as shown in (24).

Sentences with Multiple Wh-MO expressions

(24) **Dare-mo nani-mo** tabe-**nakat**-ta
who-MO what-MO eat-Neg-Past
'Nobody ate anything.'

The sentence is perfectly grammatical in the sense indicated in the translation. The present analysis can explain how this interpretation is derived. Take a look at (25).

(25)

Wh-MO$_1$	Wh-MO$_2$	SN	→	Wh-MO$_1$	Wh-MO$_2$	SN
*u*Neg val	*u*Neg val	*i*Neg [ ]		~~*u*Neg val~~	~~*u*Neg val~~	*i*Neg val
goal	goal	probe		feature deletion		feature valuation

Multipe Agree — **locus of negation**

Here the SN bears the unvalued but interpretable Neg-feature and acts as a probe for the two NCIs, hence multiple Agree operation takes place. As a result, the Neg-features of the NCIs are both valued and deleted while the Neg-feature of the SN remains and constitutes the locus of interpretation. This is how a single negation reading obtains in a structure with multiple WH-*mo* NCIs.

Unlike the case of multiple WH-*mo* NCIs, a *sika*-marked NP cannot occur multiply, as shown by ungrammaticality of (26).

Sentences with Multiple NP-sika expressions

(26) * John-**sika** yasai-**sika** tabe-**nai**
John-SIKA vegetable-SIKA eat-Neg
'Only John eats only vegetables.'

This can also be accounted for under the present proposal. (26) can be analyzed as shown in (27), where the two *sika*-marked NPs both bear an interpretable and valued Neg-feature.

(27)

NQ NP-sika$_1$	NQ NP-sika$_2$	SN	→	NQ NP-sika$_1$	NQ NP-sika$_2$	SN
*i*Neg val	*i*Neg val	*u*Neg [ ]		*i*Neg val	*i*Neg val	~~*u*Neg val~~
goal	goal	probe				

Multipe Agree — **Two loci of negation yield a gibberish sentence.**

As a result of multiple Agree, the two *sika*-marked NPs constitute the locus of the interpretation for the Neg-feature. But such a structure cannot be interpreted as the interface because a formal feature must be interpreted once and only once, as stipulated in (2d). The contrast between (26) and (24) can be easily captured under the present proposal.

3. CONCLUSION

This paper has argued, following P&T, that feature interpretability should be separated from feature values. Unlike Chomsky's (2000, 2001) system, P&T's approach allows the derivation in which the probe constitutes the locus of interpretation as well as the one in which the goal does so. We have presented evidence for the former type of derivation from the realm of WH-*mo* NCIs in Japanese. Finally note that an anonymous reviewer points out that there are some attempts to get rid of the four-way feature distinctions of P&T. Specifically, the reviewer mentions Zeijlstra 2010, which proposes that Agree involves a probe α and a goal β, such that:

(28) a. α and β are in a proper local domain
b. α has some uninterpretable feature [*u*F]
c. β has a matching interpretable feature [*i*F]
d. α is c-commanded by β
e. There is no matching goal carrying [*i*F] between α and β

If this system is extended to negative concord in Japanese, we would have to say that a SN will carry a [*u*Neg] while a NCI will bear an [*i*Neg] and c-command the former. At this point, it is not clear how such an analysis can account for the empirical facts discussed in section 2 (that is, a SN expresses negation when a NCI that it licenses is a WH-*mo* expression whereas it does not do so when a NCI occurs in the form of a *sika*-marked NCI). Thus, I would like to leave the reviewer's suggestion for future research.

REFERENCES

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

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

Pesetsky, D, & E. Torrego. (2007) "The Syntax of Valuation and the Interpretability of Features", in S. Karimi, V. Samiian & W. Wilkins (ed.) *Phasal and Clausal Architecture: Syntactic Derivation and Interpretation*, John Bejamins, Amsterdam, 262-294.

Watanabe, A. (2004) "The Genesis of Negative Concord: Syntax and Morphology of Negative Doubling", *Linguistic Inquiry* 35, 559-612.

Zeijlstra, H. 2010. "There is only one way to Agree", Paper presented at the 33rd Generative Linguistics in the Old World (GLOW 33), Wroclaw.