A TSL Analysis of Japanese Case

Kenneth Hanson
Stony Brook University
kenneth.hanson@stonybrook.edu

July 3, 2022 (Partial Draft)

Abstract

Recent work in subregular syntax has revealed deep parallels between various syntactic phenomena, many of which fall under the computational class TSL (Graf 2018; Graf 2022). Vu et al. (2019) argue that case dependencies are yet another member of this class. But their analysis focuses mainly on English, which is famously case-poor. In this paper I present a TSL analysis of Japanese, which features a much wider range of case-marking patterns, providing additional support to the claim that case dependencies are TSL.

1 Introduction

Work on the computational complexity of strings has identified a rich hierarchy of subregular classes of patterns, showing that phonological patterns are among the simplest possible (Heinz 2018). A tantalizing possibility is that the tree-based equivalents of these classes might reveal the same result when applied to syntax. Graf (2018) adapts the class TSL (tier strictly-local), which can handle long distance dependencies, to syntax, and Graf (2022) provides preliminary evidence that many disparate syntactic operations fall under this class. But confirming this hypothesis requires much additional work, because the inherent invisibility of syntactic representations makes it difficult to claim with certainty what structures are possible.

This paper aims to shed light on the situation by focusing on syntactic case, which I define for the purposes of investigation to be those heads or features that are realized by case morphology. In other words, we are not interested in abstract Case in the sense of Chomsky (1981), but only in syntactic elements with visible realization. Vu et al. (2019) provides a proof of concept for a TSL analysis of case, using Minimalist Grammars (MGs) as the formalism of choice, and focusing primarily on English. This work extends theirs by providing an analysis of Japanese, which features a much richer range of case patterns. While far from exhaustive, the TSL analysis outlined here appears to be viable not only for the basic patterns, but for many complex constructions as well.

For reasons of space, I do not review the underlying mathematics in any detail. The presentation is intentionally non-technical so that readers unfamiliar with MGs and tree tiers should be able to follow regardless; interested readers should consult Vu et al. (2019) and the references therein. I wish to point one key difference, which is that that while Vu et al. (2019) use MG derivation trees, I use dependency trees, which are equivalent but more compact and allow the relevant generalizations to be stated more directly.

The remainder of this paper is structured as follows. Section 2 provides an overview of the basic patterns to be analyzed. In Section 3 I define several tree tiers which encode these generalizations and confirm that they correctly handle the basic data, then attempt to extend the analysis to more complex structures. Section 4 concludes.

2 Basic case patterns

Japanese has four core cases, realized by the invariant suffixes *ga* (nominative), *o* (accusative), *ni* (dative), and *no* (genitive). The prototypical functions are similar to German and other European languages: subjects receive nominative case, direct objects receive accusative, and indirect objects receive dative, while complements and possessors of nouns are genitive. Examples of simple intransitive (1a), transitive (1b), and ditransitive sentences (1c) are given below, along with examples of a nominal complement (1d) and possessor (1e). All examples are presented in topic-less sentences, since topic marking masks the underlying case.[1]

(1) a. Taroo ga hasitta.
Taroo NOM run.past
Taroo ran.

[1] Abbreviations: NOM = nominative, ACC = accusative, DAT = dative, GEN = genitive, NPST = non-past, PASS = passive, CAUS = causative.

b. Taroo ga piano o hiita.
Taro NOM piano ACC play.past
Taroo played the piano.

c. Jin ga Yumi ni hon o ageta.
Jin NOM Yumi DAT book ACC give.past
Jin gave Yumi a book.

d. Yama no e
mountain GEN picture
a picture of a mountain

e. Taroo no hon
Taroo GEN hon
Taroo's book

While these are the canonical patterns, several others are also possible. Some transitive verbs exceptionally select for a dative object (2a).[2] Stative verbs allow for two additional patterns: a dative subject with a nominative object (2b), and double nominative marking (2c). Transitive adjectives and certain complex predicates formed with a stative suffix also display the same patterns.

(2) a. Taroo ga Yumi ni atta.
Taroo NOM Yumi DAT meet.PAST
Taroo met Yumi.

b. Yumi ni tenisu ga dekiru.
Yumi DAT tennis NOM can.do
Yumi can play tennis.

c. Yumi ga tenisu ga dekiru.
Yumi NOM tennis NOM can.do
Yumi can play tennis.

To summarize the patterns described so far, we see that objects of non-stative transitive verbs are accusative, nominals under the scope of a nominal are genitive, and subjects and objects of certain verbs and adjectives are dative. When no other condition applies, nominative obtains. This suggests an analysis along the following lines:

- Nominative, accusative, and genitive are structural cases.
- Some instances of dative are structural, while others are lexical.
- Nominative case is the default case, occurring when no other case can be licensed.

While this analysis is strongly reminiscent of dependent case (Marantz 2000; Baker and Vinokurova 2010), I do not believe that any of the data discussed in this paper necessarily rules out an Agree-based alternative. I remain agnostic as to the correct theoretical interpretation, focusing instead on the raw distributional patterns. We will see that these generalizations can be formalized straightforwardly using TSL constraints over MG dependency trees, and that they require only verb limited modification in order to handle more complex constructions.

[2]Certain verbs are compatible with either an accusative or dative object, with a difference in temporal properties. See Fukuda (2007) for details.

3 Analysis

Now, I will begin to formalize the generalizations made in the previous section. Ultimately, constraints over three different tree tiers will be used. The first corresponds to cases licensed in the scope of verbal predicates, i.e., accusative and dative. The second handles genitive case, licensed in the scope of nominal predicates. The last handles instances of lexical dative case.

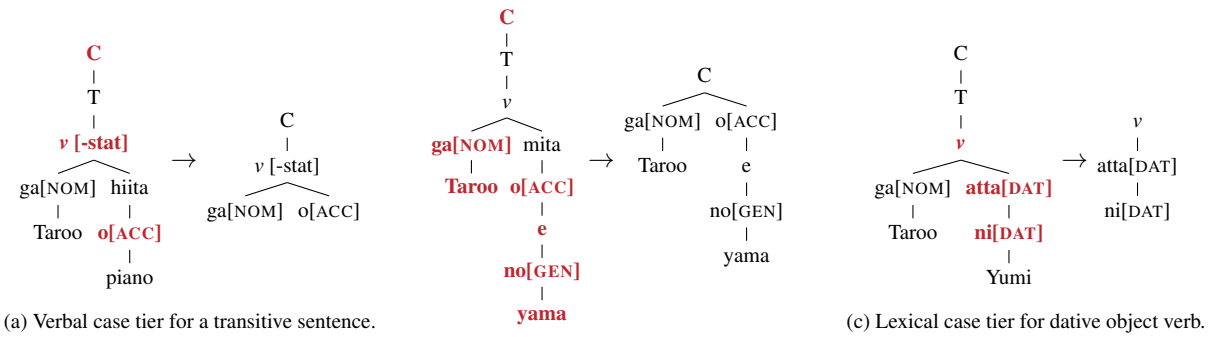
Before continuing, I wish to lay out a few syntactic assumptions. Clauses are assumed to have a basic functional hierarchy C $>$ T $>$ v $>$ V. We will also consider some constructions involving additional functional elements, including an applicative, passive, and causative head. For ease of exposition, all nominals are treated as NPs, with case markers occupying a higher KP. Finally, PPs are analyzed as taking an NP complement directly, with the K layer absent.[3] I will ignore PPs and other modifiers for the most part since they rarely interact with case marking, though we will encounter a few interesting situations related to passives and the well-known ban on multiple accusatives.

The remainder of this section is structured as follows. First, I introduce the three tree tiers and their constraints, corresponding to structural cases licensed in the verbal domain, the nominal domain, and lexical case. From there I will consider more complex constructions, including embedded clauses, passives, and causatives. The initial analysis will be refined slightly in order to account for this larger set of data.

3.1 Accusative and dative case

First, we define a tree tier to structural cases licensed in the verbal domain. On this tier we project non-stative v and all heads of category K bearing nominative, accusative, or dative case. In addition, we project C in order to limit the case licensing domain to a single clause (embedded clauses are discussed in Section 3.4). There are two constraints on this tier, licensing accusative and dative case, respectively. The rightmost K head of v must bear accusative if it has at least one KP sibling on the left. The middle of three or more K heads must bear dative if the rightmost K head bears accusative. All other K heads must bear nominative, which is the default.

[3]PPs may alternatively they may be analyzed as KPs bearing semantic case.



(a) Verbal case tier for a transitive sentence.

(b) Genitive tier for N–N complementation.

(c) Lexical case tier for dative object verb.

Figure 1: Examples for verbal, genitive, and lexical case tiers.

Consider the example in Figure 1a. On the left is a dependency tree for the simple transitive clause from example (1b). On structural case tier, shown to the right, the only children of *v* are the K heads *ga* and *o*. The tier is well-formed because the accusative K is the second daughter of non-stative *v* and the nominative K meets the elsewhere criterion.

Simple intransitives and ditransitives are handled with no further complication. If there is only a single KP argument, it will be the only child of *v*P on the tier, whether it originated as an internal or external argument, and so it must be nominative. If there are three KP arguments, then the rightmost must be accusative, and the middle must be dative. In there could somehow be four arguments, we would predict the middle two to be dative; we will see in Section 3.6 that this is correct, though there are some problems with this analysis. Similarly, if a ditransitive were to be passivized, we might predict the return to a nominative-accusative pattern; this is also correct, though there is a small complication, as discussed in Section 3.5.

Now, we consider stative predicates, which allow either a dative-nominative or double nominative pattern. Since we do not project stative *v* on the tier, accusative case is not licensed, and the object will be nominative by default. Furthermore, the chain licensing constraint for dative case does not apply, since there are only two KPs, so the subject must be either dative or nominative. Now, it not the case that these two patterns are completely interchangeable, since the individual lexical item controls whether the subject must be dative. This fact will be handled by the lexical dative tier.

As a final note, accusative, unlike dative, never occurs twice in the same clause, and never in any position but the lowest. Because of this, the accusative constraint, but not the dative constraint, should be strengthened to a biconditional. With that said, the current tier projection rule and constraints are as stated below.

(3) **Verbal case tier (initial version)**

Project: K[NOM/ACC/DAT], *v*[-stat], C

Accusative constraint: if a head of category K is the rightmost of two or more children of *v*, then it must carry ACC. No other K may bear ACC.

Dative constraint: if a head of category K is neither the first nor last child of *v*, and its right sister bears either ACC or DAT, then it must carry DAT.

3.2 Genitives

Next, we turn to genitives, which have the simplest distribution: as a first approximation, all KPs in the domain of a nominal are genitive, and no others.

We construct the genitive tier in manner very similar to the verbal case tier. We project all KPs and all N heads, as well C in order to shield embedded clauses. There is one constraint, which is that all KP children of N must be genitive. An example shown in Figure 1b. In this sentence, there is only a single K head which is a child of a noun, and this K bears GEN as required, so the tier is well formed. There are no restrictions on the remaining KPs, though they could of course be ruled out on other tiers. Needless to say, in a complex possessive or N-N complementation structure, every embedded nominal would be genitive.

There is a possible problem regarding the particle *no*, which is that it also appears between PP complements and modifiers and their head nouns. An example is given below in (4a) below. It is likely that in at least some structures *no* is not actually a case marker, but a generalized marker of adnominal modification, as traditionally distinguished in the literature. A clear example of this use is shown in (4b).

(4) a. otera e no michi
temple to NO road
'the road to the temple'

b. aka no hana
red NO flower

'a red flower'

Fortunately, we can abstract away from this issue for present purposes. If these instances of *no* are case particles, then they still adhere to the constraint as stated; if not, then the constraint simply does not apply. To summarize, then, the genitive tier is defined as follows.

(5) **Genitive case tier**
Project: K, N, C
Genitive constraint: if a head of category K is a child of N, then it must carry GEN. No other K may bear GEN.

3.3 Lexical datives

Finally, we consider lexically controlled datives, which include both subjects and objects. While it may be possible to integrate these patterns into the verbal case tier, this would complicate the constraints considerably, so we instead form a new tier dedicate solely to controlling the distribution of lexically dative-marked nominals. If we determined that certain postpositions should be treated as lexically governed cases, these could be added to the same tier.

To create the tier, we project all dative KPs and possible licensing categories, which may include *v*, V, and A. There is a single constraint tier: every dative licensor must have a KP child bearing dative case. The converse is not required because some dative KPs may be structurally licensed. An example for a transitive verb with a dative object is given in Figure 1c. In this case, we assume that it is V that licenses dative case. The only KP daughter of dative-licensing V is dative, so the tier is well-formed. If *v* was the dative licensor, then the subject would need to be dative instead.

One question that this analysis raises is what should happen if there are two KP children of the same dative case licensor. As far as I am aware, this situation never arises in Japanese. Dative experiencers are presumably external arguments, and so can always be distinguished from themes. Goals of ditransitive verbs may appear in two positions: low goals are PP daughters of VP, while high goals are KPs daughters of an applicative head, so they can be distinguished as well (cf. Miyagawa and Tsujioka 2004). If it should turn out that there are situations where a single dative-licensing head has two KP arguments, we could project all KPs and use their linear order to pick out the one that must be dative. This issue aside, our lexical dative tier is defined as follows.

(6) **Lexical case tier**
Project: K[DAT], *v*[DAT], V[DAT], A[DAT]
Dative constraint: every head of category *v*/V/A that bears DAT must have a child of category K bearing DAT.

Now, having defined three tiers modeling the canonical uses of the four core cases, we will consider several more complex structures, and see that the system outlined so far extends to handle them with little or no modification.

3.4 Embedded clauses

There are several types of finite embedded clauses in Japanese. By default, these show the same case pattern as matrix clauses, as would typically be expected. However, there are additional patterns in which the embedded subject may be exceptionally marked accusative or genitive. We will first confirm that the TSL constraints developed up to this point work correctly for the basic pattern before considering how they may be adapted to handle the exceptional patterns.

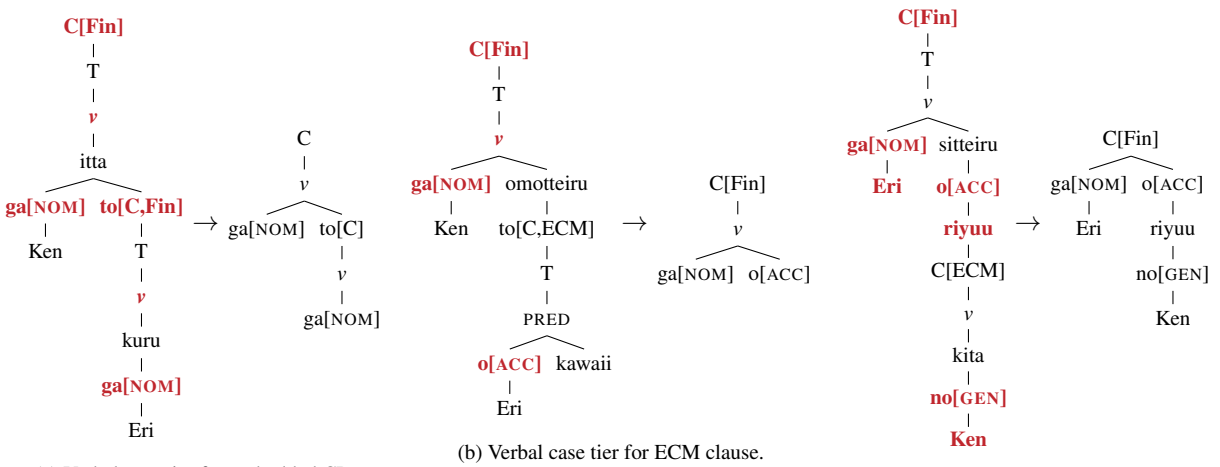
Examples of two types of finite embedded clauses are given in (7) below. Here, *to* is analyzed as a complementizer, while *koto* is an abstract noun taking a CP complement. These constructions require no additional assumptions to handle, since we already project C on the tier, forming a new case domain for each finite clause. We can confirm that this is the case by examining the verbal case tier projections of these sentences. The tier projection for sentence (7a) is shown in Figure 2a.

(7) a. Ken ga [Eri ga kuru to] itta.
Ken NOM Eri NOM come C said
Ken said that Eri will come.

b. Eri ga [Ken ga tegami o okutta C]
Eri NOM Ken NOM tegami ACC sent C
koto o sitteiru.
thing ACC know
Eri knows that Ken sent the letter.

At this point, it may not be completely clear why C needs to be projected at all, since the lower *v* in these examples already shields the embedded verb's arguments. We will see several relevant examples shortly.

Turning now to the alternative patterns, we first examine the Japanese ECM construction, in which the embedded subject apparently takes accusative case (8a). In some cases the *o*-marked nominal may in fact be a matrix object, with a pro subject in the embedded clause, but Kishimoto (2018) argues that at least some ECM subjects are genuine. Another similar phenomenon is *ga-no conversion*, in which subject of a nominal clause exceptionally takes genitive case (8b) (see Maki and Uchibori (2008) for details). Both of the examples below are also grammatical with a nominative embedded subject.



(a) Verbal case tier for embedded CP.

(b) Verbal case tier for ECM clause.

(c) Genitive case tier for *ga-no* conversion.

Figure 2: Example case tiers for embedded clauses.

(8) a. Ken ga [Eri {ga/o} kawaii to]
Ken NOM Eri {NOM/ACC} be.cute C
omotteiru.
think
'Ken thinks that Eri is cute.'

b. Eri ga [Ken {ga/no} kita C] riyuu
Eri NOM Ken {NOM/GEN} came C reason
o sitteiru.
ACC know
'Eri knows the reason that Ken came.'

Much has been written about the syntax of these constructions. For the purpose of this paper, I wish to draw attention to the fact that ga-no conversion occurs within the scope of a higher nominal, while in the ECM construction only the particle *to* intervenes between the embedded clause and the higher verb. Futhermore, the ECM verb is always one that could otherwise license accusative case.

This suggests a way to capture the alternation that does not involve adding additional contexts for accusative and genitive case licensing. I posit that these lexical items may select for a special C head which is transparent to case licensing, call this C[ECM]. Under the present analysis, we state that while the ordinary C[Fin] is projected on the structure case tier, C[ECM] is not. Examples tier projections (8a) and (8b) are shown in Figures 2b and 2c, respectively. For simplicity, I treat the subject of an adjective as the specifier of a PRED head, and ignore aspectual morphology of the complex verbs *omotteiru* and *sitteiru*.

The updated verbal case tier definition is provided below. All said, we see that the facts about embedded clauses are handled quite well under the TSL perspective.

(9) **Verbal case tier (final version)**
Project: K[NOM/ACC/DAT], *v*[-stat], C[-ECM]
Accusative constraint: if a head of category K is the rightmost of two or more children of *v*, then it must carry ACC. No other K may bear ACC.
Dative constraint: if a head of category K is neither the first nor last child of *v*, and its right sister bears either ACC or DAT, then it must carry DAT.

3.5 Passives

Now that we have dealt with simple and embedded clauses, I now examine several ways of forming complex predicates within a single clause, including the passive and causative constructions.[4] These interact with case marking in a manner which is difficult to capture in traditional approaches but which is surprising straightforward under the present analysis.

The Japanese passive suffix *rare* has several functions. One of these, the *direct passive*, decreases the valency of a transitive verb by eliminating the agent, which may be optionally reintroduced in a *by*-phrase. (Another function, the *indirect passive* is not considered here.) As is typical in a nominative-accusative system, the theme is then promoted to become a subject. The passive of a ditransitive verb, however, has two possible argument frames: one in which the (higher) goal is promoted, and one in which the (lower) theme is promoted. Examples

[4]More precisely, the cases of passive and causative verbs considered here are those where the suffix appears as a distinct functional head in the syntax. These are commonly referred to as *biclausal* analyses, as this was how they were treated in earlier forms of generative grammar. This wording contrast with an analysis in which a complex predicate is formed within the lexicon and inserted as a single verb with a modified argument structure. At least some productive causatives are probably best understood in this way.

are given in (10) and (11).

(10) Active/passive transitive verb

a. Sensei ga gakusei o hometa.
teacher NOM student ACC praised
'The teacher praised the student.'

b. Gakusei ga (sensei ni) homerareta.
student NOM teacher by praise.PASS.PAST
'The student was praised (by the teacher).'

(11) Active/passive ditransitive verb

a. Mari ga kodomo ni okasi o ataeta.
Mari NOM child DAT candy ACC gave
'Mari gave the child candy.

b. Kodoma ga (Mari ni) okasi o
child NOM Mari by candy ACC
ataerareta.
give.PASS.PAST
'The child was given candy (by Mari).'

c. Okasi ga (Mari ni) kodomo ni
candy NOM Mari by child DAT
ataerareta.
give.PASS.PAST
'The candy was given to the child (by Mari).'

The fact that the theme of a passivized transitive verb raises is straightforwardly understood under the common assumption that agent is not projected in Spec-*v*P in passives, and that the optional *by*-phrase is an adjunct PP. Miyagawa (1989) argues that this is indeed the case in Japanese. In this case, the complex predicate effectively has the argument structure of an unaccusative verb, and so the theme receives nominative case.

The same can be said of the goal passive of a ditransitive verb. Once the agent is eliminated, the next highest argument is the goal, so it becomes the subject, even though it otherwise would have received dative case, not accusative case. This would be surprising if dative was always a lexical case, but is completely natural if dative case licensed by a ditransitive verb is structural. Thus, the goal passive of a ditransitive verb is behaves exactly like a transitive verb.

There is an elegant solution for the theme passive as well. Recall from Section 3.3 that the goal of a ditransitive verb may occupy one of two positions, and that the higher position is a KP while the lower position is a PP. In other words, there are two structures parallel to the English double object construction and *to*-dative, respectively. This is obscured by the fact that the case marker *ni* and the postposition *ni* are homophonous, along with the fact that the default word order is the same. Thus, it should be possible to passivize the direct object by selecting the low goal structure, in which the goal is a PP and therefore ineligible for promotion.

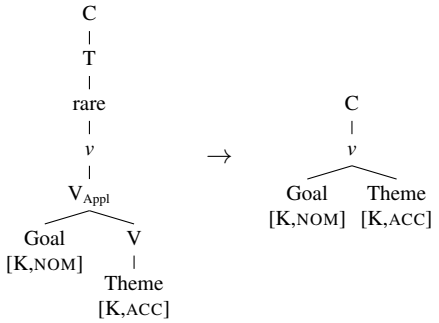


Figure 3: Verbal case tier for goal passive of ditransitive

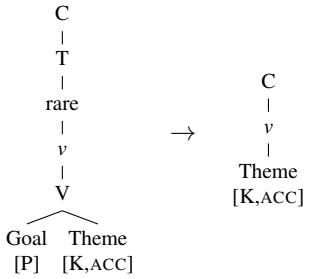


Figure 4: Verbal case tier for theme passive of ditransitive

So we see that the facts about passives fall out naturally in the present analysis. Dependency trees and verbal case tiers for the ditransitive goal passive and theme passive are shown in Figures 3 and 4, respectively.

3.6 Causatives

As our final case study, we consider the syntactic causative construction, which *increases* the valency of the complex predicate. The causative head *sase* is compatible with verbs of any valency. Causative equivalents of the examples in (1) are given in (12) below. These examples are called *ni*-causatives, as the causee, usually the agent of the main verb, takes dative case. There is also an *o*-causative, in which the causee receives accusative case, but is subject to various restrictions, and I will not analyze it here.

(12) a. Ken ga Taroo ni hasiraseta.
Ken NOM Taroo DAT run.CAUS.PAST
Ken made Taroo run.

b. Ken ga Taroo ni piano o
Ken NOM Taro DAT piano ACC
hikaseta.
play.CAUS.PAST
Ken made Taroo play the piano.

c. Ken ga Jin ni Yumi ni hon o
Ken NOM Jin DAT Yumi DAT book ACC

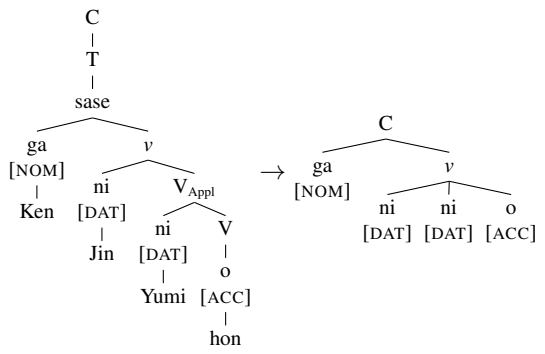


Figure 5: Verbal case tier for causative ditransitive

agesaseta.
give.CAUS.PAST
Ken made Jin give Yumi a book.

Notice how in (12a) a dative argument is introduced without there being an accompanying accusative argument. This suggests that lexical dative case is being licensed. As the valency of the main verb increases, the main verb objects appear in the expected cases, suggesting that these are the usual structural cases. Unfortunately, the tiers as they are defined now cannot quite handle this.

Consider just the accusative/data tier for the ditransitive causative, shown in Figure 5. The rightmost (lowest) KP has more than one KP sister to the right, and so it is accusative. Moving leftward, the goal KP has an accusative KP to the right and another KP to the left, so it is dative. The third KP has no left sister, so while it may be dative, it is not required to be. Finally, the causer is outside the scope of *v*P, so it must be nominative.

The problem is that, as it stands now, we cannot ensure that the causee is dative rather than nominative. On the contrary, if it is the causative head *sase* that licenses dative, we would expect the causer to be dative rather than the causee. Clearly, one of the assumptions we made so far needs to be changed. The option that seems most appealing to me is to change the way that the verbal case tier is constructed such that all four KPs end up as sisters, and analyze all instances of dative case in causative verbs to be structural dative, not lexical. The downside to this approach is that we cannot explain why the causee of an intransitive verb may be dative; we would predict that it must be accusative. For now, I leave this issue open.

4 Conclusion

In this paper, I developed a TSL analysis of Japanese case, and showed that the basic generalizations are captured neatly with a system of three tiers and a small number of constraints, and that the analysis generalizes to a wide range of constructions, including stative predicates, embedded clauses, ECM, and passives. Causatives proved somewhat more difficult, with either intransitives or ditransitives remaining unexplained depending on whether they are taken to license lexical dative case. It seems likely that an examination of other complex predicates of this sort, such as indirect passives and potentials, might help sort out this issue. While the analysis has been far from exhaustive, the results so far point towards the viability of the TSL approach to Japanese case.

References

Baker, Mark and Nadya Vinokurova (2010). "Two modalities of case assignment: Case in Sakha". In: *Natural Language & Linguistic Theory* 28.3, pp. 593–642.

Chomsky, Noam (1981). *Lectures on government and binding: The Pisa lectures*. 9. Walter de Gruyter.

Fukuda, Shin (2007). "Object Case and Event Type: Accusative-Dative Object Case Alternation in Japanese". In: *Annual Meeting of the Berkeley Linguistics Society*. Vol. 33. 1, pp. 165–176.

Graf, Thomas (2018). "Why movement comes for free once you have adjunction". In: *Proceedings of CLS* 53, pp. 117–136.

– (2022). "Typological Implications of Tier-Based Strictly Local Movement". In: *Proceedings of the Society for Computation in Linguistics 2022*, pp. 184–193.

Heinz, Jeffrey (2018). "The computational nature of phonological generalizations". In: *Phonological Typology, Phonetics and Phonology*, pp. 126–195.

Maki, Hideki and Asako Uchibori (2008). "Ga/no conversion". In: *The Oxford handbook of Japanese linguistics*.

Marantz, Alec (2000). "Case and licensing". In: *Arguments and case: Explaining Burzio's generalization*, pp. 11–30.

Miyagawa, Shigeru (1989). *Structure and Case Marking in Japanese*. Academic Press.

Miyagawa, Shigeru and Takae Tsujioka (2004). "Argument structure and ditransitive verbs in Japanese". In: *Journal of East Asian Linguistics* 13.1, pp. 1–38.

Vu, Mai Ha et al. (2019). "Case assignment in TSL syntax: A case study". In: *Proceedings of the Society for Computation in Linguistics (SCiL) 2019*, pp. 267–276.