

Finnish case alternating adpositions: A corpus study*

Sander Lestrade (Radboud University Nijmegen)

Abstract

In this paper, I discuss Finnish case alternation adpositions from a theoretical and corpus perspective. First, I argue that postpositional PP constructions with genitive case denote the standard spatial meaning of which an extension is marked with partitive case. Also, I show how word order interacts with case assignment. Both findings are formalized in a bidirectional Optimality Theoretic framework. Second, I show that case alternating behavior does not occur unrestrictedly in newspaper corpora. Adpositions in principle tend to assign the same case to the same object over and over again, and only a small subgroup of highly frequent nouns is assigned both genitive and partitive case by the same adposition(s). This suggests that (adpositional case) alternation is only allowed for highly frequent adpositional objects.

1. Introduction

Finnish has six local cases and a multiple of at least 17 adpositions to denote spatial meaning (the base forms being multiplied with their possible inflected forms; cf. Sulkala and Karjalainen 1992, Huumo and Ojutkangas 2006). As it is sometimes hard to draw the line between adposition and relational noun, this number should be taken as the lower limit. One could even think of the class of adpositions as an open class (cf. Suutari 2006 for discussion). Three of these adpositions, *lähe-* ‘near’, *keske-* ‘in the middle’, and *ympäri* ‘(a)round’, can assign both genitive and partitive case, all others exclusively assign either of these cases.¹ In this paper, I will account for this adpositional case alternation. I will argue that the use of the partitive case results in an extension of the meaning expressed by the genitive PP construction. Also, I will show how word order interacts with case assignment. In Section 3, I will formalize these findings in a bidirectional Optimality Theoretic account (biOT; Blutner et al. 2006). Like standard OT, biOT makes use of violable constraints, not of absolute rules. Unlike standard OT however, biOT takes into account both the hearer’s and the speaker’s perspective, thereby optimizing over form-meaning *pairs*. Doing so, biOT is perfectly suited to deal with variation in form and meaning. I will show how a more marked adpositional construction combines with an extended meaning, the unmarked construction combining with the basic spatial meaning. After these theoretical sections, I will verify my hypotheses with corpus findings in section 4. It will be shown that the case alternating behavior of these adpositions is not attested across the board, and that there is a significant difference in frequency of occurrence between those nouns that do show the case alternation and those that appear with a single case. I will argue that this difference in frequency is due to the fact that an online case

alternation is only feasible for highly frequent adpositional objects, low frequent objects make a default choice.

2. Adpositional case alternation in Finnish

Most Finnish adpositions assign either genitive or partitive case, the first option being the most common. A very small set of adpositions may assign both cases. Not surprisingly, this yields a difference in meaning. In (1) we see this difference for *ympäri* ‘around’:²

- (1) a. Juoks-i-mme kaupungi-n ympäri
run-PAST-1PL city-GEN around
‘We ran around the city’
b. Juoks-i-mme ympäri kaupunki-a
run-PAST-1PL around city-PART
‘We were running around in the city’

As the verb remains the same in these two constructions, neither the change in meaning nor the case alternation is caused by the verb; it cannot be the verb that causes the difference in meaning. The PP construction with the genitive means ‘around’ with the city as the center of the running circle, that with partitive case ‘around in’, in which the city is the place of a criss-cross running event.

For *keske-* ‘in the middle of’ the meaning alternation is very similar. With genitive case ‘middle’ has to be taken very literally, with partitive case this is not the case:

- (2) a. Lelu-t ovat lattia-n keske-llä
toy-PL.NOM were floor-GEN in.the.middle.of-ADE
‘The toys are (exactly) in the middle of the floor’
b. Lelu-t ovat keske-llä lattia-a
toy-PL.NOM were in.the.middle.of-ADE floor-PART
‘The toys are spread around the center of the floor’

In (2a), the toys could almost be thought to be placed in a pile in the center of the room. In (2b), with a partitive case marking, they are scattered on the floor of the room. Note again that the verb remains the same in the two constructions.

The case alternation of *lähe-* ‘near’ seems to work a little bit differently. Here, both constructions can be used for the spatial meaning ‘near’. However, as we can see in (3c), an abstract object is only assigned partitive case:

- (3) a. Auto on talo-n lähe-llä
car.NOM is house-GEN near-ADE
‘The car is near the house’
b. Auto on lähe-llä talo-a

- car.NOM is near-ADE house-PART
 ‘The car is near the house’
- c. Olemme lähe-llä ratkaisu-a
 are near-ADE solution-PART
 ‘We are close to a solution’
- d. *Olemme ratkaisu-n lähe-llä
 are solution-GEN near-ADE
 ‘We are close to a solution’

Closeness to a concrete object like a car can be formulated with either the genitive or partitive case. In case of closeness to a solution, however, partitive case is required.

Thus far, the distribution of labor in alternations seems to be such that the genitive PP constructions denote the more standard spatial meanings, whereas the partitive case is used for extensions of that meaning. This is similar to the adpositional case alternation found in many Indo-European languages, where accusative case sometimes alternates with a canonical oblique case to express a more complex, directional meaning (Lestrade 2008). But, sometimes, in Finnish, word order is an additional factor of importance. Normally, postpositions assign genitive case and prepositions assign partitive case. If partitive case combines with a preposition this results in a standard, more literal meaning, but if it is combined with a postposition a more abstract meaning is obtained. This is illustrated in (4):

- (4) a. Tämä on lähe-llä minu-a
 this.NOM is near-ADE 1SG-PART
 ‘This is close to me’
- b. ?Tämä aihe on lähe-llä minu-a
 this.NOM topic.NOM is near-ADE 1SG-PART
 ‘This topic is close to me’ (i.e., sitting next to me on the couch)
- c. Tämä aihe on minu-a lähe-llä
 this.NOM topic.NOM is 1SG-PART near-ADE
 ‘This topic is close to my heart’

Spatial ‘closeness’ is expressed by a standard combination of a preposition with partitive case in (4a).³ With an abstract subject as in (4b), this combination results in an odd reading in which a topic is said to be physically present. The intended abstract meaning is expressed with the combination of partitive case with postpositional *lähellä* ‘near’, as in (4c). The same phenomenon is found for the genitive construction of *yli* ‘over’ as well:

(Vainikka, 1993)

- (5) a. aida-n yli
 fence-GEN over
 ‘over the fence’
- b. yli odotuste-n
 over expectation.PL-GEN

‘beyond expectations’

The standard position for the genitive case assigning adposition (or, in this case, for an adposition when it assigns genitive case) is after the object. In (5a) this construction is indeed used with the spatial meaning of *yli* ‘over’. In case of an abstract meaning, as in (5b), the adposition precedes the object.

Word order variation with *ympäri* ‘around’ also yields a meaning difference. This time the difference does not lie in the abstractness or metaphoricity of the meaning, but rather in its telicity. Telicity, mostly applied in the analysis of verb phrases, is about the viewpoint from which an event is described. A telic event is viewed from its endpoint, an atelic event is perceived as ongoing. Adpositional phrases can be analyzed in terms of telicity as well (Zwarts, 2005; Lestrade, 2008; Nam, 2005). Consider the following examples:

- (6) a. John walked towards the house
b. John walked into the house

In (6a), nothing is specified about the endpoint of the walking event. John could be walking towards the house for hours without having to arrive for the sentence to be true. (6b), however, necessarily includes the transition from John being outside to John being inside of the house. Therefore, the latter sentence is said to be telic, the former atelic. Such a difference in telicity with *ympäri* ‘around’ is expressed by means of word order, as can be seen in (7):

(P. Wolski, p.c.)

- (7) a. *ympäri talo-a*
round house-PART
‘around in the house’
b. [%]*talo-a ympäri*
house-PART round
‘(several times OR not completely) around the house’
c. **ympäri talo-n*
round house-GEN
d. *talo-n ympäri*
house-GEN round
‘(once) around the house’

In (7b) the atypical combination of partitive with postpositional *ympäri* ‘around’ yields an atelic meaning variant of the telic genitive PP construction. Although this uncommon construction is not judged completely grammatical by every speaker, the interpretation of it indeed is ‘atelic around’. With the object of postpositional *ympäri* marked with genitive case one would go round around the house precisely once, starting for example from the front door and returning at that point; postpositional *ympäri* with partitive case, however, would mean that one either did not complete an entire circle or went several times around the house. Interestingly, this is reminiscent

of the verbal domain, where objects of *unbounded* VPs (to be explained below) are assigned partitive rather than accusative case, as illustrated in (8).

(Kiparsky, 1998)

- (8) a. Ammu-i-n karhu-j-a
 shot-PAST-1SG bear-PL-PART
 ‘I shot at the bears’ OR ‘I shot bears’ OR ‘I shot at bears’
 b. Ammu-i-n karhu-t
 shot-PAST-1SG bear-PL.ACC
 ‘I shot the bears’

Kiparsky (1998) uses the notion *unboundedness* to account for the use of partitive case in the verbal domain. Partitive case both has an aspectual function in the marking of gradable events and an DP-related function in the marking of quantitative indeterminacy (including indefinite bare plurals and mass nouns), both of which are analyzed as unbounded. The diagnostic that Kiparsky proposes is the modifiability by degree adverbs such as (*some*) *more*, *a lot*, *very much*, *a bit*, *considerably*, *slightly*. The notion of unboundedness is formalized by with the properties of DIVISIVENESS, CUMULATIVITY and DIVERSITY (Kiparsky 1998:284):

- (9) a. P is DIVISIVE iff $\forall x[P(x) \wedge \neg atom(x) \rightarrow \exists y[y \subset x \wedge P(y)]]$
 b. P is CUMULATIVE iff $\forall x[P(x) \wedge \neg sup(x,P) \rightarrow \exists y[x \subset y \wedge P(y)]]$
 b. P is DIVERSE iff $\forall x\forall y[P(x) \wedge P(y) \wedge x \neq y \rightarrow \neg x \subset y \wedge \neg y \subset x]$

(10) A predicate is UNBOUNDED iff it is divisive and cumulative and not diverse

A predicate is divisive iff x is not the minimal element (9a), and cumulative iff x is not the maximal element (9b). Predicates of which the elements are atomic only and predicates whose elements are not related by the subpart relation are bounded by the diversity restriction, although they may satisfy (9a) and (9b). The unboundedness of (8a), as is clear from the translation options, could be caused by the unboundedness of the verb (the first reading), by the unboundedness of the DP (the second reading), or by a combination of both (the third reading). Using the accusative case is only possible under a reading in which both the verb and the DP are bounded, cf. (8b). Indeed, Vainikka and Maling (1996) argue that partitive case is the default case for the verbal object, the accusative having a very specific meaning. In a context of three bears and a hunter, the hunter cannot have shot only two bears or have shot more or less at the bears without changing the truth values of (8b). For (8a) to be true, however, the hunter could have shot more or less in the direction of approximately three bears. Now, the atelic meaning that is assigned to the combination of the partitive case with postpositional *ympäri* ‘around’ does not appear randomly, but can be motivated by the use of the partitive case in the verbal domain. Partitive case expresses an unbounded meaning. The bounded ‘once around the house’ is expressed

with genitive case; the unbounded ‘around in the house’ and ‘several times or not completely around the house’ are expressed by partitive case.

A final concomitant meaning difference to the case alternation concerns definiteness. Consider the following examples:

- (11) a. sopimuskaude-n keske-llä
contract.season-GEN middle.of-ADE
‘in the middle of the contract season’
b. keske-llä sopimuskaut-ta
middle.of-ADE contract.season-PART
‘in the middle of a contract season’

The genitive marked object in (11a) has a more definite reading than the partitive marked object in (11b).⁴ In Finnish, definiteness is not overtly marked in the form of articles. In the verbal domain, as we saw above, the difference in definiteness can be expressed by case. Marking the direct object with accusative instead of partitive implies that the object is definite, and therefore bounded. At the PP level too, partitive case can be used for a distinction in definiteness of objects. Apparently then, the similarity in unboundedness between the VP and PP does not only hold at the head level, but it holds at the argument level too. At both levels, genitive/accusative case is used for the bounded meanings (genitive and accusative case in Finnish have the same form *-n*), whereas partitive case is used for the unbounded ones. Of course, this only holds with respect to case alternating adpositions. If there is no optionality in case, there is none in semantics either.

In the next section the observations described above will be formalized and accounted for in a bidirectional Optimality theoretic account.

3. A bidirectional Optimality Theoretic account

In OT, we make use of violable constraints rather than inviolable rules. What are the constraints we are dealing with here? First, we saw that in the verbal domain, unbounded VPs are assigned partitive case. We can formulate this as a form constraint, that accounts for the general observation of Kiparsky that unbounded meaning is expressed with partitive case in (8):

- (12) UNBND->PART: Unbounded meaning is expressed with partitive case

Second, Vainikka (1993) convincingly argues that genitive case is the structural default case for the specifier position, whereas partitive case is the structural default case for the complement position. Structural, or configurational, case is tied to specific positions in the tree. Some examples of these default cases in different constructions are given in (13)-(16).

- (13) a. ensi vuoden suunnitelmat
 next year.GEN plan.PL
 ‘next year's plans’
 b. Liisan lähtö
 Liisa.GEN departure
 ‘Liisa's departure’

In (13), the specifier of the NP bears genitive case. The same holds for the specifier of the AP in (14). The complement of the comparative adjective *lyhyempi* ‘shorter’ bears partitive case.

- (14) a. valtavan kylmä
 enormous.GEN cold
 ‘enormously cold’
 b. Riitta on Liisaa lyhyempi
 Riitta.NOM is Liisa.PART shorter
 ‘Riitta is shorter than Liisa’

We find the same pattern in PPs (15). Specifiers of PPs bear genitive, prepositional complements bear partitive case.

- (15) a. talon takana
 house.GEN behind
 ‘behind the house’
 b. ilman sateenvarjoa
 without umbrella.PART
 ‘without an umbrella’

Finally, the specifiers of nonfinite VPs bear genitive case, the default case for the verbal object is partitive.

- (16) a. Jukka käski Pekan juosta
 Jukka.NOM asked Pekka.GEN run.INF
 ‘Jukka asked Pekka to run’
 b. Riitta luki kirjaa
 Riitta.NOM read book.PART
 ‘Riitta was reading a/the book’

In all examples, the default case for the specifier position is genitive case, the default case for the complement position is partitive case. By definition, this default case is sometimes overruled by more specific constraints. Examples are the marking of boundedness/telicity by the accusative case on the complement or the assignment of nominative case to the subject of a finite verb (Kiparsky 1998). The default case assignment in Finnish is illustrated in Figure 1.

Figure 1. *Structural default case in Finnish*

This too can be formulated as two constraints, accounting for the fact that postpositions generally go with genitive case, whereas prepositions go with partitive case:

- (17) POSTP->GEN: postpositions assign genitive case
PREP->PART: prepositions assign partitive case

These two constraints are the language specific instantiations of a universal constraint that says that structural default case should be assigned. In all tableaux below it is possible to replace the two constraints by this single universal constraint. I spelled them out like this to increase the intelligibility of the tableaux.

Then, as will be further corroborated with corpus data in Section 4, postpositions are the default PP option in Finnish. Since most adpositions are derived from nouns (Huumo and Ojutkangas 2006) and genitive case is most prominently used to mark possessive relations (Blake 2001: 149ff; Sulkala and Karjalainen 1992: 88ff), it is easy to understand how genitive case assigning postpositions have become the default markers for the adpositional object. This goes against the universal tendency for SVO languages to have prepositions (Finnish being an SVO language), and should probably be seen the result of a diachronic change in basic word order (Grünthal 2003). Indeed, most prepositions seem to have only recently developed and in Finnic in general, most adpositions are postpositions (around 75% according to Grünthal 2003). In Sulkala and Karjalainen (1992) there are even no prepositions proper mentioned at all, all adpositions being either postpositions or alternating ones. Given the idea of attraction by high type frequency constructions (cf. Barðdal *to appear*), the overall preference for postpositions will push adpositions that in principle can be used in both constructions towards the postpositional option. I will dub this general tendency the MAAIVELD principle, after a Dutch proverb *Je moet je hoofd niet boven het maaiveld uitsteken* ‘One should not stick one’s head above the ground level (i.e. one should behave like all others)’, of which *PREP is the relevant and short PP instance:

- (18) MAAIVELD (*PREP): behave like most others (i.e., be a postposition)

Finally, it is generally accepted that spatial meaning is primary to abstract meaning (cf. a.o. Anderson 1971, Jackendoff 1983, O’Keefe 1996). We need some understanding of the space around us to survive, but not necessarily of abstract concepts. The latter are often understood by means of spatial metaphor, for example, time being conceptualized as one-dimensional and events being mapped to intervals like objects to locations (Bierwisch 1996). Possession too is expressed as a spatial being-at relation in many languages, as illustrated in the following Finnish example:

- (19) Sepo-lla on kirja
Seppo-ADE is book
'Seppo has a book'

Since spatial meaning is primary, we can say that abstract meaning is always a more complex and therefore less preferred interpretation for a spatial construction like a PP. Clearly, this constraint only matters if a spatial interpretation is possible, otherwise it equally violates all abstract meaning candidates.

- (20) *ABSTR: avoid an abstract interpretation.

With these five independently motivated constraints, we can account for the variation we observed in Section 2 using a bidirectional perspective. In bidirectional Optimality Theory (BiOT; Blutner et al., 2006), optimization takes place in two directions, both from meaning to form and from form to meaning, yielding two super-optimal form-meaning pairs. A form-meaning pair is found super-optimal if there is no better form (that is not already involved in another super-optimal pair) for its meaning, nor a better interpretation (that is not already involved in another super-optimal pair) for its form. Take, for example, the prototypical super-optimal pairs <*kill*, 'cause to die (directly)'\> and <*cause to die*, 'cause to die (indirectly)'\> (Blutner et al., 2006). The periphrastic causative *cause to die* is a more marked form than the lexical causative *kill*, as the former consists of more morphemes (violating a general ECONOMY constraint, see for example Kiparsky (2004)). The direct situation is said to be unmarked, as it is stereotypical, unmediated causation through physical action; the indirect, mediated situation is a more marked meaning.

Now, how do we map form to meaning? This is illustrated in Figure 2. The horizontal arrows indicate the optimization of form (i.e., the speaker's perspective), promoting an unmarked form (*f*, *kill* in the example above) over a marked one (*f'*, *cause to die* in the example above). If this would be the only optimization process at work, for both the marked (*m*, 'cause to die indirectly') and unmarked meaning (*m'*, 'cause to die directly'), the unmarked form *f* would be chosen. The vertical arrows indicate the optimization process of meaning (i.e., the hearer's perspective), promoting an unmarked interpretation over a marked one. If this would be the only optimization process at work, for both the marked (*f*) and unmarked form (*f'*), the unmarked interpretation *m* would be chosen. If a speaker would use the unmarked form for a marked meaning (which is a good thing to do from his perspective), the hearer would not get the intended interpretation, interpreting the unmarked form with an unmarked meaning. The idea of bidirectional optimization is that the perspective of the other speech participant is taken into account as well. Only when both the hearer and the speaker take into account the other perspective as well, communication of a marked meaning will be successful. When the hearer notices that a speaker uses a marked form instead of the unmarked form that he could have used, he knows that it is not the unmarked meaning that the speaker wanted to express, as there would have been a less marked form to do that. Thus, two super-optimal pairs arise: the first

combining the unmarked form and meaning; the second combining a marked form and a marked meaning.

Figure 2. *Weak bidirectional OT*

The procedure in BiOT is as follows: In a first round of optimization, the least marked meaning (that is the interpretation that violates the least meaning constraints) is combined with the least marked form (that is the form that violates the least form constraints). This form-meaning pair is called the first super-optimal pair. In a second round of optimization, the second best meaning is combined with the second best form. Obviously, when only two meaning and form candidates are present this means the marked meaning pairs up with the marked form (cf. Horn 1984). In this second round, a meaning like ‘kill indirectly’ is combined with a form like *cause to die*. Although the form *kill* would be less marked (shorter in this case), it cannot be used for the expression of ‘kill indirectly’ as it already takes part in the first super-optimal form, with the meaning ‘kill directly’. The periphrastic causative *cause to die* could in principle mean ‘cause to die directly’, but this interpretation is blocked as it is already expressed by the less marked form *kill*.⁵ Similarly, ‘cause to die indirectly’ could be expressed by the lexical causative *kill*, but this form is already in use expressing the less marked ‘cause to die directly’. Compare the following two sentences:

- (21) a. Black Bart killed the sheriff.
- b. Black Bart caused the sheriff to die.

In (21b), Black Bart could have stuffed the gun of the sheriff with cotton, causing the gun to backfire to the death of the sheriff; contrastively, (21a) would be the expression for a situation of Black Bart shooting the sheriff. The marked form is used for a marked situation; the unmarked form for an unmarked one.

Just like in standard OT, optimization processes are represented in so-called tableaux. Differently from standard OT, however, in a BiOT tableau the violation pattern of form-meaning pairs is given, rather than the violation patterns of output candidates for a given input. Thus, candidate pairs are judged simultaneously from an interpretation and a production perspective.

First, let us look at the optimization procedure of *ympäri* ‘around’. Every prepositional construction leads to a violation of the constraint that says that a PP should be postpositional. Every *postpositional* construction with a partitive case leads to a violation of the constraint that says that postpositions govern genitive case. Every *prepositional* construction with a genitive case leads to a violation of the constraint that says that prepositions govern partitive case. The violation pattern of the unboundedness constraint is a little bit more difficult to fill in, as it requires a domain-specific interpretation. Every gradable meaning with a genitive case leads to a violation of the constraint that says that gradable meaning should be expressed with partitive case. ‘Around in’ is a gradable notion, as it is both true for a five minute walk through the city and for a walk of an hour. ‘Around’ is less so, as ‘around’ stops

being ‘around’ at some point (cf. Zwarts 2004). Therefore, the combination of ‘around in’ with a genitive form is a more serious violation of the constraint UNBND→PART than the combination of ‘around’ with genitive case. All this leads to the violation pattern in Tableau 1.

But before we can check whether the violation patterns of the candidates correctly predict the super-optimal pairs (that is, before we can check whether the proposed constraints indeed can account for the observations in section 2), we need to determine the constraint ranking. In OT, violations of constraints are only allowed in order to satisfy higher ranked constraints. But what are the higher ranked constraints? The ranking I propose has a usage-based motivation. First, following Zwarts et al. (*to appear*), I rank lexical constraints (UNBND→PART) highest. The two constraints POSTP→GEN and PREP→PART follow on a par (indicated with the dashed line in the tableaux). Since these constraints are never in direct conflict, one cannot determine their mutual ranking. These two constraints are ranked above *PREP. In every clause, there is evidence for the default case for specifiers and complements (remember that POSTP→GEN and PREP→PART are only the PP instantiations of Vainikka’s 1993 general observation), but not so for the choice between pre- and postpositions. In other words, the default case constraints are relevant in many more situations than *PREP. This difference in amount of evidence and relevancy between the constraints is translated in their mutual ranking. The general and unspecific *ABSTR, finally, can be nothing but a low ranked constraint. It specifies a general interpretation preference that only becomes apparent in situations in which an abstract and concrete meaning are equally likely. Indeed, as we will see shortly, this ranking can account for the case alternations described above.

In a first round of optimization, the form meaning pair < ‘(once) around’, GEN + *ympäri*> is found optimal, as it is the only combination that does not violate a constraint whatsoever. That means that there is no better way of expressing this meaning, nor a better way of interpreting this form. This is indicated with the symbol “☞”. For the second round, all combinations with the meaning ‘(once) around’ and the form GEN + *postposition*, are out of the competition, as there is a first super-optimal pair combining this meaning and form already (cf. the bidirectional optimization principles outlined above). This is illustrated with the strikethrough of all relevant combinations. For the meaning ‘around in’ the combination with a partitive prepositional construction is found optimal. Of the remaining form-meaning candidate pairs, this pair has the best violation pattern. *If* one now would want to make a further distinction between ‘once around’ and ‘unbounded around’, only two form-meaning pairs remain for consideration. Of these two, the combination with a partitive postpositional construction has the best violation pattern. If one wants to make a distinction between these two meanings using simple adpositional constructions only, the correct prediction is that ‘unbounded around’ combines with a partitive case taking postposition, whereas ‘once around’ combines with the postpositional construction with a genitive case.

	UNBND→PART	POSTP→GEN	PREP→PART	*PREP	*ABSTR
☞ < ‘(once) around’, <i>ympäri</i> + <i>GEN</i> > < ‘(once) around’, <i>ympäri</i> + <i>PART</i> > < ‘(once) around’, <i>GEN</i> + <i>ympäri</i> > < ‘(once) around’, <i>PART</i> + <i>ympäri</i> >		*	*	*	
☞ < ‘around in’, <i>ympäri</i> + <i>GEN</i> > < ‘around in’, <i>ympäri</i> + <i>PART</i> > < ‘around in’, <i>GEN</i> + <i>ympäri</i> > < ‘around in’, <i>PART</i> + <i>ympäri</i> >	*	*	*	*	
☞ < ‘unbounded around’, <i>ympäri</i> + <i>GEN</i> > < ‘unbounded around’, <i>ympäri</i> + <i>PART</i> > < ‘unbounded around’, <i>GEN</i> + <i>ympäri</i> > < ‘unbounded around’, <i>PART</i> + <i>ympäri</i> >	*	*	*	*	

Tableau 1: bidirectional optimization of *ympäri* ‘around’

Now let us consider *keske-* ‘in the middle’. Again, every prepositional construction leads to a violation of the constraint that says that a PP should be postpositional. Every postpositional construction with a partitive case leads to a violation of the constraint that says that postpositions govern genitive case. Every prepositional construction with a genitive case leads to a violation of the constraint that says that prepositions govern partitive case. Every gradable meaning with a genitive case leads a violation of the constraint that says that gradable meaning should be expressed with partitive case. In the first round of optimization, the pair < ‘in the precise middle’, *GEN* + *keske-* > is found super-optimal, as it does not violate any constraint. After deleting this option for the second round, the combination of ‘all over’ with the partitive prepositional construction is found the second super-optimal form-meaning pair.

Finally, consider the optimization of *lähe-* ‘near’. The violation of three constraints is similarly motivated as before. This time however, differently than before, it is not a violation of UNBND→PART that determines the first super-optimal candidate, but it’s the violation of *ABSTR by the abstract meaning of the PP. Correctly, we find < ‘near concrete Ground’, *GEN* + *lähe-* > and < ‘near abstract Ground’, *lähe-* + *PART* > to be the first and second super-optimal pairs.

	UNBND→PART	POSTP→GEN	PREP→PART	*PREP	*ABSTR
< ‘in the precise middle’, <i>keske-</i> + <i>GEN</i> > < ‘in the precise middle’, <i>keske-</i> + <i>PART</i> > < ‘in the precise middle’, <i>GEN</i> + <i>keske-</i> > < ‘in the precise middle’, <i>PART</i> + <i>keske-</i> >		*	*	*	
< ‘all over’, <i>keske-</i> + <i>GEN</i> > < ‘all over’, <i>keske-</i> + <i>PART</i> > < ‘all over’, <i>GEN</i> + <i>keske-</i> > < ‘all over’, <i>PART</i> + <i>keske-</i> >	*	*	*	*	

Tableau 2: bidirectional optimization of *keske-* ‘in the middle’

	UNBND→PART	POSTP→GEN	PREP→PART	*PREP	*ABSTR
< ‘near concrete Ground’, <i>lähe-</i> + <i>GEN</i> > < ‘near concrete Ground’, <i>lähe-</i> + <i>PART</i> > < ‘near concrete Ground’, <i>GEN</i> + <i>lähe-</i> > < ‘near concrete Ground’, <i>PART</i> + <i>lähe-</i> >		*	*	*	
< ‘near abstract Ground’, <i>lähe-</i> + <i>GEN</i> > < ‘near abstract Ground’, <i>lähe-</i> + <i>PART</i> > < ‘near abstract Ground’, <i>GEN</i> + <i>lähe-</i> > < ‘near abstract Ground’, <i>PART</i> + <i>lähe-</i> >		*	*	*	*

Tableau 3: bidirectional optimization of *lähe-* ‘near’

Importantly, the abstract meaning is not a proper feature of partitive case, whereas unboundedness is. The latter is lexically specified, the former is a result of bidirectional optimization, where the combination of a concrete meaning with the postpositional genitive construction is found super-optimal, and the partitive prepositional construction combines with the abstract meaning in a second round of optimization. The idea that abstract meaning is expressed with a more marked construction and not with a specific case per se is further founded when looking at the construction alternation of *yli* ‘over’ in (5). There we saw that spatial meaning is combined with the standard construction of a postposition with genitive case, but that abstract meaning is combined with a more marked construction, namely a *preposition*

with genitive case. This construction is more marked not for economy reasons, but because it violates the constraints *PREPP and PREP→PART.⁶

In the next section, we will evaluate this theoretical story against a corpus, with real, unelicited data. Does the adpositional case alternation indeed occur, and if it does, can it be accounted for by the theory as described above?

4. A corpus study

Actual language use may differ from the individual speaker's or linguist's intuitions. What is claimed to be strictly ungrammatical may turn out to be a frequently used, rule-governed option or just an idiosyncratic judgment. The context that lacks in example sentences on which grammaticality judgments are based may turn the most ill-sounding examples in perfectly grammatical constructions. This context is available in language corpora, in which the written or spoken language of different language users is recorded. Because of the recent technological and statistical developments and the growth of these computer-readable texts and recordings, large amounts of actual language use can now be studied. These data sometimes force linguists to rethink grammatical theory, as is for example shown for the dative alternation by Bresnan et al. (2007). In this section, I show how corpus findings indeed force us to rethink the BiOT analysis proposed above as well. It is only the highly frequent constructions that exhibit the alternation that the BiOT grammar formalizes. Frequency seems to be a necessary condition for variation.

For my corpus study, I randomly selected five newspapers from the CSC, a Finnish on line corpus:⁷ Demari 2000, a newspaper of the Social Democrats; Karjalainen 1998 and 1999 Aamulehti 1999, two pretty high standard, big newspapers, the latter especially for the region of Tampere; Hyvinkään Sanomat 1997 and Hämeen Sanomat 2000, two lower quality, smaller newspapers. This totaled to a corpus of approximately 27 million words. I extracted all instances of the three case alternating adpositions *ympäri* 'around', *keske-* 'in the middle' and *lähe-* 'near' from the corpus. Only the combination of prepositional *kesken* with genitive case was removed from the data set, as this form of *keske-* has an unrelated meaning ('during . . . , suddenly . . . ') that cannot be expressed with a partitive case equivalent.

In total, I have considered 3671 adpositional objects, after removing those objects that appeared less than three times. As illustrated in Table 1, indeed both the pre- and postpositional construction are used, and both genitive and partitive case are assigned. At least in terms of raw frequency, the postpositional construction (and genitive case) are the more common option for *case alternating* adpositions too.⁸ Therefore, these numbers support the postulation of the *PREP constraint

Table 1. *Finnish case alternating PPs: construction and case*

n postpositions	1930	N genitive case	1987
n prepositions	1741	N partitive case	1684
total n	3671	total n	3671

Before we have a closer look at the case alternation, let us consider the effect of word order. As said above, case assignment and word order are related: Postpositions generally assign genitive case, prepositions assign partitive. In Figure 3, the relation between the proportions of postpositional use and genitive case assignment is given. The black diagonal in the plot represents a hypothetical, perfect, positive linear relation between the proportion of postpositional use and that of genitive case assignment. As can be seen, almost all words are on this line, which means that there is almost a linear relation between the two proportions. That is, when a case alternating adposition is more often used as a preposition (it is more likely that) it assigns partitive case, and when it is more often used as a postposition (it is more likely that) it assigns genitive case.

Figure 3. *Relation proportions postpositional use and genitive case assignment*

Most words always combine with either a pre- or with a postposition. This explains the big clusters in the lower left and upper right corner the graph. Those words that are assigned different cases can be found in the middle of the y-axis, with partitive case proportions between 0 and 1. An interesting group is the group of nouns that tend to take some case disproportionately. These are the nouns that pattern away from the diagonal, towards the upper left (genitive) and lower right (partitive) corners of the plots. It turns out that such atypical case assignment to these objects is marginal and due to particular adposition-object combinations. The *lähe-* ‘near’ objects *sydän* ‘heart’ and *puolue* ‘(political) party’ prefer partitive case irrespective of their position. As argued for in Section 2, indeed both objects that appear in partitive case with postpositional *lähe-* ‘near’ could be used as the thing one is metaphorically rather than literally close to: I could have affinity with some party (*puolee*), and something could be close to my hear (*sydän*). Only the genitive preference of the two *ympäri* ‘around’ objects *vuorokausi* ‘day and night’ and *vuosi* ‘year’ is unexpected, as a prepositional construction with genitive case is predicted to be ungrammatical (a prediction which is indeed borne out for other objects). Since this construction only seems to occur with these two highly frequent time expressions, they could probably be thought of as fixed archaic expressions. This is not an uncommon thing cross-linguistically: highly frequent constructions or items are known to resist standardization (cf. Haspelmath 2008). Given the almost perfect match between word order and case, I will neglect word order in what follows.

In Table 2, the different case groups of Table 1 above are further inspected. Four groups are discerned: *Gen* and *Part* are noun types occurring with respectively genitive and partitive case only; *Diff* is the group of nouns that occur with different cases, but always in the same case-adposition combinations, for example always with genitive case for *keske-* and with partitive when combined with *lähe-*; the group *Alt* consists of those nouns that are assigned different cases by the same adposition(s). The different groups are illustrated in Examples (22)-(25). An example of a *Gen* noun is *saari* ‘island’. This noun always occurs in genitive case with both *keske-* ‘in the middle of’ and *ympäri* ‘round’, and never combines with *lähe-* ‘near’:

Gen

- (22) a. saare-n ympäri
island-GEN around
'around island'
b. saare-n keske-llä
island-GEN in.the.middle.of-ADE
'in the middle of the island'

An example of a *Part* noun is *lattia* 'ground'. This noun always occurs in partitive case with both *keske* 'in the middle of' and *lähe*- 'near', and never combines with *ympäri* 'around':

Part

- (23) a. Keske-llä lattia-a
in.the.middle.of-ADE ground-PART
'In the middle of the ground'
b. lähellä lattia-a
near-ADE ground-PART
'close to the floor'

An example of a *Diff* noun is *risteys* 'intersection'. This noun always occurs in genitive case in combination with *lähe*- 'near' and in partitive case in combination with *keske*- 'in the middle'; it does not occur in combination with *ympäri* 'around':

Diff

- (24) a. risteykse-n lähe-llä
intersection-GEN near-ADE
'close to the crossroad'
b. keske-lle risteyksi-ä
in.the.middle.of-ALL intersection.PL-PART
'in the middle of crossroads'

An example of an *Alt* noun, finally, is *talo* 'house'. Crucially, this noun occurs in both partitive and genitive case with *ympäri* 'around'. In combination with *lähe*- 'near' it occurs in genitive case only and in combination with *keske*- 'in the middle' it always appears in partitive case:

Alt

- (25) a. ympäri talo-a
around house-PART
'around (in) the house'
b. talo-n ympäri
house-GEN around
'around the house'

- c. talo-n lähe-llä
house-GEN near-ADE
'close to the house'
- d. keske-llä talo-a
in.the.middle.of-ADE house-PART
'in the middle of the house'

Strikingly, the group of *Alt* nouns is very small and the case alternation occurs only rarely. Although the three adpositions in principle should be able to assign either case to most adpositional objects, they tend to assign the same case to the same object over and over again (239 out of 272 nouns occur in only one case, i.e. *Gen*, *Part*, *Diff*). Only 25 nouns (listed in Table 3 below) are assigned both genitive and partitive case by the same adposition(s) (*Alt*; cf. Table 2).

Table 2. *Case distribution per group*

	<i>n types</i>	<i>n tokens</i>	<i>mean (log) frequency</i>
<i>Gen</i>	175	1241	7.1 (0.85)
<i>Part</i>	64	561	8.8 (0.94)
<i>Diff</i>	8	132	16.5 (1.2)
<i>Alt</i>	25	1737	69.5 (1.8)
	Total: 272	Total: 3671	Overall mean (log): 13.5 (1.13)

Before we address the question why the case alternation is restricted to this small group only let us see if the grammar proposed above at least makes the right predictions for this group. In Table 3 below it is shown which case is assigned to these nouns by which adposition. By definition, at least one of the adpositions assigns both cases (*alt*). The other adposition(s) that combine with this noun then could consequently combine with either case (*gen*, *part*, or *both*), or never select this particular noun (-). For every noun in Table 3 the case alternation can indeed be explained by its concomitant meaning alternation described in Section 2. Recall from above that nouns governed by *lähe-* 'near' could appear in any case, except for abstract nouns. Indeed, the abstract nouns *luonto* 'nature' and *kenttä* 'area' always appear in partitive case.

The case alternation of *keske-* was about precise vs. sloppy 'middle'. Constructions with genitive case denote the very center of the ground; those with partitive case something like 'all over'. For the nouns that exhibit this alternation, this meaning distinction makes sense: for example, shops can be found 'all over' a city center (*kaupunki*, *asutukse*, *kylä*), whereas the town hall could stand precisely in its middle, or, one could walk in the middle of a forest (*metsä*), but never see the well that is precisely in the middle of it. For words like *sota* 'war' and *vuosi* 'year' this alternation may denote a difference in definiteness, as in example (11) above, repeated here for convenience.

- (26) a. sopimuskaude-n keskellä

contract.season-GEN middle.of
 ‘in the middle of the contract season’

- b. keskellä sopimuskaut-ta
 middle.of contract.season-PART
 ‘in the middle of a contract season’

For *ympäri* the meaning alternation between ‘around’ and ‘around in’ indeed is applicable to nouns like *koulu* ‘school’, *talo* ‘house’, *kaupunki* ‘city’ and *maailma* ‘world’. *Talo* ‘house’ and *kaupunki* ‘city’ were the nouns used in the examples above. Also, one could wander all over the world (*maailma*) “partitively”, whereas a satellite circles around the world “genitively”.

Table 3. *Case alternating nouns*

Stem	Meaning	Case assigned by		
		<i>lähe-</i>	<i>keske-</i>	<i>ympäri</i>
<i>asutukse</i>	settlement	alt	alt	-
<i>elämä</i>	life	-	alt	-
<i>huonee</i>	room	-	alt	part
<i>ihminen (pl)</i>	human being, man	alt	gen	-
<i>kaupunki</i>	city	alt	alt	part
<i>kenttä</i>	field, area	part	part	alt
<i>keskusta</i>	center	alt	-	part
<i>kirkko</i>	church	alt	gen	gen
<i>koti</i>	home	alt	-	-
<i>koulu</i>	school	gen	gen	alt
<i>kylä</i>	village	gen	alt	part
<i>lentokenttä</i>	airport	alt	-	-
<i>luonto</i>	nature, wildlife	part	alt	-
<i>maa</i>	country, earth	alt	gen	part
<i>maailma</i>	world	gen	-	alt
<i>maapallo</i>	globe	-	-	alt
<i>metsä</i>	forest, woods	-	alt	-
<i>pääkaupunki</i>	capital	alt	-	-
<i>raja</i>	boundary, limit	alt	-	-
<i>ranta</i>	border, edge	alt	-	-
<i>rautatieasema</i>	railway station	alt	-	-
<i>sota</i>	war	-	alt	-
<i>talo</i>	house	gen	part	alt
<i>tori</i>	marketplace	-	alt	gen
<i>vuosi</i>	year	-	alt	gen

Some of the nouns that appear in partitive case only can be explained by their abstract meaning. The majority of this group (48 out of 80)⁹ consists of words that are

assigned partitive case by *lähe*- ‘near’. Half of these are abstract nouns like ‘average’, ‘agreement’, ‘surprise’, ‘balance’, ‘succes’, ‘retirement age’, and as we have seen in Section 2, *lähe*- ‘near’ obligatorily combines with partitive case with abstract nouns. But abstractness of meaning cannot explain why other complements of *lähe*- ‘near’ only take partitive case. Also, it cannot explain why some others take genitive case only, let alone the non-occurrence of case alternations for the complements of the other two Ps.

So, what restricts the case alternation to a small group of 25 nouns? I think the answer is frequency. If we look at the mean frequency of the different groups in Table 2, we see huge differences. To correct for outliers, the mean logged frequencies of the different object classes are given between parentheses. This reduces the effect of very highly frequent items (for a more detailed motivation of this correction cf. Baayen 2008). The logged frequency of case alternating nouns is significantly higher than that of nouns that are consequently assigned one particular case (Wilcoxon rank sum test, $W = 5772$, $p\text{-value} \sim 0$). Also, it is noteworthy that those nouns that exhibit a case alternation are very often also selected as the adpositional object of some other adposition. This is not to say that the case alternation arises from the fact that these nouns are assigned genitive case by the one and partitive case by the other adposition. As indicated above, only those nouns that are assigned both cases by the same adposition(s) are called “case alternating”. At first sight, one could think that the non-alternating nouns just might have been too infrequent to appear in both cases. But the relation between the higher mean frequency and the case alternation cannot be explained by a lack of frequency of the single case nouns: There were a number of very highly frequent nouns (like *päivä* ‘day’ (logged frequency = 1.77), *sydän* ‘heart’ (1.68), and *maakunna* ‘province’ (1.83)) that did not take both cases.

Importantly, this finding is not restricted to Finnish. Consider the Dutch adposition *in* ‘in’. As a preposition, it generally has a locative meaning (27a), but in the right context, it can get a directional interpretation (27b; cf. Nikitina 2008). When used postpositionally, it unambiguously has this directional meaning (27c).

- (27) a. Ik loop in de kamer
 I walk in the room
 ‘I walk (around) in the room’
 b. Ik schuif de doos in de kamer
 I push the box in the room
 ‘I push the box into the room’
 c. Ik loop de kamer in
 I walk the room in
 ‘I walk into the room’

I withdrew all *in* PPs, both with pre- and postpositions, from the syntactically annotated part of the Corpus of Spoken Dutch (CGN). In the following table, the mean frequency for two types of complements is given. *Pre* are complements that only occur in prepositional constructions and *Both* are complements that occur in both the prepositional and postpositional variant. In the computation of the means, I have

left out all complements with a frequency below 3 again. These complements would disproportionately lower the mean frequency of the most frequent type (i.e. Pre) to the advance of my proposal. There were no complements that only occurred in a postpositional construction that reached this threshold, therefore this logical third type candidate is lacking in the overview in Table 4. The *mean total frequency* is the mean of the total frequency of the complements in both the prepositional and postpositional construction. For complements of the Pre type, this is necessarily the same as the *mean prepositional frequency*, the mean of the frequency of complements in the prepositional construction only. This latter statistic is given to show that the difference in frequency between the two types is not due to fact that complements that occur in both constructions are more frequent because they are counted twice.

Table 4. *Mean (log) frequencies for in complements*

	<i>n types</i>	<i>n tokens</i>	<i>mean prepositional frequency</i>	<i>mean total frequency</i>
Pre	352	4310	12.24 (1.09)	12.24 (1.09)
Both	30	803	25.30 (1.40)	26.77 (1.43)
	Total: 382	Total: 5113	Overall mean (log): 13.27 (1.22)	Overall mean (log): 13.38 (1.13)

Just like we saw in Table 2 for Finnish, the mean frequencies for complements that may occur in both constructions is significantly higher than that of complements that occur in the prepositional construction only (Wilcoxon rank sum test, $W = 7462.5$, p -value = 0.0001484).

In this section we saw that variation is restricted to highly frequent items. Apparently, the costly operation of having an online construction alternation is only feasible if the collocation of that adposition with some object, the *collostruction* in terms of Stefanowitsch and Gries (2003), is frequent enough. If not, the collostruction will shift to a default choice. Now, why would frequency be a necessary condition for variation? The answer is that only highly frequent words have the “mass” to behave idiosyncratically; less frequent members need to behave like others, otherwise they cannot be learned. Thus, the corpus findings can be seen as empirical evidence for the claim in Pullum and Scholz (2007) that it is because of memory limits that less frequent items have to behave regularly.

“Sets of grammatically idiosyncratic forms exhibiting partially overlapping subregularities are a familiar feature of the most frequently occurring items in the vocabulary, but out in the long tail of the frequency distribution, where the rare words are, there has to be a degree of regularity and predictability – some clusters of items sufficiently unfamiliar that all their syntactic behavior can be inferred on the basis of general facts about whole equivalence classes of words.” (Pullum and Scholz 2007: 400)

My corpus findings hint at an even more important role for frequency in grammar. Frequency differences not only explain iconicity (“marked meanings go with marked forms”) as argued for by Haspelmath (2008), they make variation possible in the first

place. Adpositional complements that are less frequent either go with a preposition or with a postposition. They do not enter a second round of optimization.

5. Conclusion

In this paper, I discussed Finnish case alternation adpositions. I argued that PP constructions with genitive case denote a standard spatial meaning of which an extension is marked with partitive case. Unbounded extensions are lexically motivated, abstract ones are due to the relative markedness of abstract meaning of spatial constructions. Also, I have shown how word order interacts with case assignment: an atypical word order of the adposition and its adpositional object expresses a more marked meaning. Both findings were formalized in a bidirectional Optimality Theoretic framework.

Using on line newspaper corpora, I have shown that this case alternating behavior does not hold across the board. Only a small subgroup of highly frequent nouns is assigned both genitive and partitive case by the same adposition(s). Only these highly frequent words have the “mass” to behave idiosyncratically; less frequent members need to behave like others, otherwise they cannot be learned. Frequency is a necessary condition for variation.

BiOT does not predict variation, it can only account for existing variation. The Bidirectional OT grammar I propose correctly describes the existing variation in the Finnish adpositional domain.

References

- Anderson, John M. (1971). *The grammar of case: towards a localistic theory*, Cambridge: Cambridge University Press.
- Baayen, Harald (2008). *Analyzing Linguistic Data: A practical introduction to Statistics*, Cambridge: Cambridge University Press.
- Barðdal, Jóhanna (to appear). The Development of Case in Germanic. To appear in: *The Role of Semantics and Pragmatics in the Development of Case*, Jóhanna Barðdal and Shobhana Chellia (eds), Amsterdam and Philadelphia: John Benjamins.
- Bierwisch, Manfred (1996). How much space gets into language. In *Language and Space*, Paul Bloom, Mary A. Peterson, Lynn Nadel and Merrill F. Garrett (eds), 31-76. Cambridge, Massachusetts: MIT Press.
- Blake, Barry J. (2001). *Case*, Cambridge: Cambridge University Press
- Blutner, Reinhard; de Hoop, Helen; and Hendriks, Petra (2006). *Optimal Communication*, Stanford: CSLI.
- Bresnan, Joan; Cueni, Anna; Nikitina, Tatiana; and Baayen, Harald R. (2007). Predicting the Dative Alternation. In *Cognitive Foundations of Interpretation*, Gerlof Bouma, Irene Krämer, and Joost Zwarts (eds), 69-94 Amsterdam: Royal Netherlands Academy of Science.

- Grünthall, Riho (2003). *Finnic adpositions and cases in change* [Suomalais-Ugrilaisen Seuran Toimituksia/Mémoires de la Société Finno-Ougrienne 244], Helsinki: Soci  t   Finno-Ougrienne.
- Haspelmath, Martin (2008) Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics* 19 (1), 1-33.
- Horn, L. (1984). Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In *Meaning, Form, and Use in Context: Linguistic Applications*, Deborah Schiffrin (ed.), 11-42. Washington, D.C.: Georgetown University Press.
- Huomo, Tuomas and Ojutkangas, Krista (2006). An introduction to Finnish spatial relations: Local cases and adpositions. In *Grammar from the Human Perspective. Case, space and person in Finnish* [Current Issues in Linguistic Theory 277], Marja-Liisa Helasvuo and Lyle Campbell (eds.), 101-128. Amsterdam and Philadelphia: John Benjamins.
- Jackendoff, Ray S. (1983). *Semantics and Cognition*. Massachusetts: MIT Press.
- Kiparsky, Paul (1998). Partitive case and aspect. In *The Projection of Arguments*, Miriam Butt and Wilhelm Geuder (eds.), 265-307. Stanford: CSLI.
- Kiparsky, Paul (2004). Blocking and periphrasis in inflectional paradigms. *Yearbook of Morphology*, 113-135. Dordrecht: Springer. 2004
- Lestrade, Sander (2006). Spatial adpositions and partitive case in Finnish. In *Proceedings of the Third ACL-SIGSEM Workshop on Prepositions*, Trento, Italy.
- Lestrade, Sander (2008). The correspondence between directionality and transitivity. In *Syntax and Semantics of Spatial P* [Linguistik Aktuell/Linguistics today 120], Anna Asbury, Jakub Dotla  il, Berit Gehrke and Rick Nouwen (eds), Amsterdam/Philadelphia: John Benjamins.
- Nam, S. (2005). Directional locatives in event structure: Asymmetry between goal and source. *Eoneohag (Linguistics)* 43, 85-117.
- O'Keefe, John (1996). The spatial prepositions in English, Vector Grammar, and the Cognitive Map Theory. In *Language and Space*, Paul Bloom, Mary A. Peterson, Lynn Nadel and Merrill F. Garrett (eds), 277-316. Massachusetts: MIT Press
- Pullum, Geoffrey K. and Barbara C. Scholz (2007) Systematicity and natural language syntax. *Croatian Journal of Philosophy* 7 (21), 375-402.
- Sulkala, H. and Karjalainen, M. (1992). *Finnish*, London: Routledge.
- Stefanowitsch, Anatol and Stefan Th. Gries (2003). Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8 (2): 209-43.
- Suutari, Toni (2006) Body part names and grammaticalization. In *Grammar from the Human Perspective. Case, space and person in Finnish* [Current Issues in Linguistic Theory 277], Marja-Liisa Helasvuo and Lyle Campbell (eds.), 101-128. Amsterdam and Philadelphia: John Benjamins.
- Talmy, L. (2001). *Toward a Cognitive Semantics*, Cambridge and London: MIT Press.
- Vainikka, A. (1993). Three structural cases in Finnish. In *Case and Other Topics in Finnish Syntax*, A. Holmberg and U. Nikanne (eds.), 129-159. Holland: Kluwer Academic Publishers.

- Vainikka, A. and J. Maling (1996). Is partitive case inherent or structural? In *Partitives. Studies on the Syntax and Semantics of Partitive and Related Constructions*, J. Hoeksema (ed.), 179-208. Berlin and New York: Mouton de Gruyter.
- Zwarts, Joost (2004). Competition between Word Meanings: The Polysemy of (A)Round. In *Proceedings of SuB8*, Cecile Meier and Matthias Weisgerber (eds.), 349-360. Konstanz.
- Zwarts, Joost (2005). Prepositional aspect and the algebra of paths. *Linguistics and Philosophy* 28, 739–779.
- Zwarts, Joost, Lotte Hogeweg, Sander Lestrade, and Andrej Malchukov (to appear), Semantic markedness in gender opposition, blocking and fossilization. To appear in: *STUF*.
- R Development Core Team (2006). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.

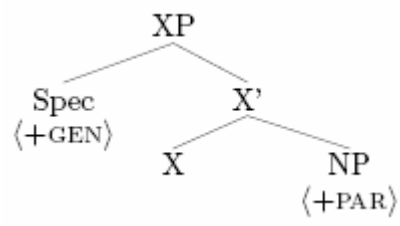


Figure 1. *Structural default case in Finnish*

$$\begin{array}{ccc} \langle f, m \rangle & \leftarrow & \langle f', m \rangle \\ \uparrow & & \uparrow \\ \langle f, m' \rangle & \leftarrow & \langle f', m' \rangle \end{array}$$

Figure 2. *Weak bidirectional OT*

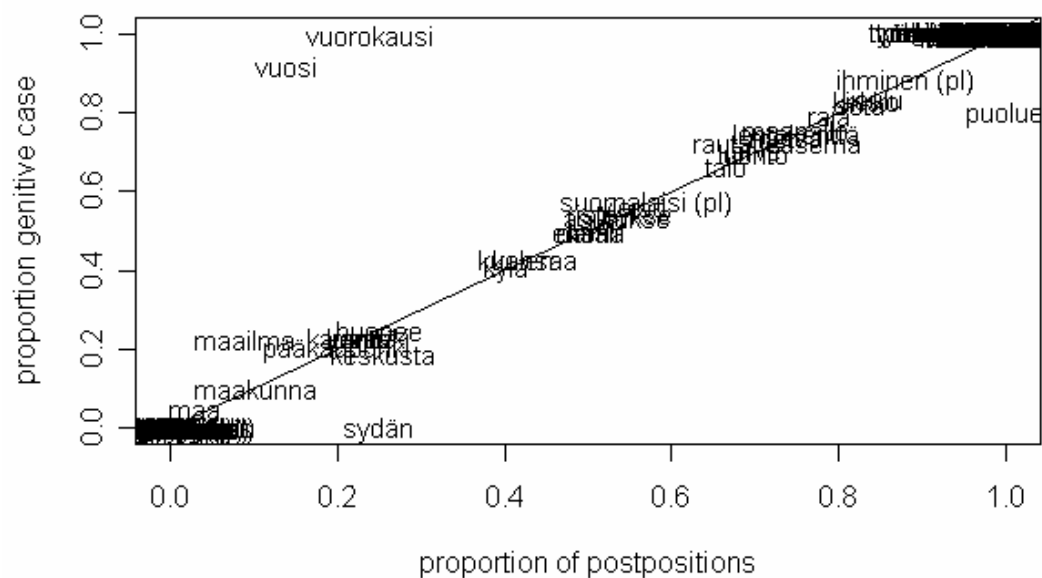


Figure 3. *Relation proportion postpositions and proportion genitive case*

* I would like to thank Arto Anttila, Seppo Kittilä and Paula Wolski for their help with the Finnish data. Many thanks to the members of the Optimal Communication research group for fruitful discussion of this paper. Also, I gratefully acknowledge the VSBfonds and the Netherlands Organisation for Scientific Research (NWO), grant 220-70-003 PIONIER project Case cross-linguistically, principal investigator Helen de Hoop, for making this research possible.

¹ In fact, a fourth adposition, *yli*- ‘over’, exhibits case alternation as well. This alternation, however, does not yield a spatial contrast, as is shown in the following example:

- (i) a. *Tämä-n auto-n hinta on yli 25000 euro-a*
 this-GEN car-GEN price.NOM is over 25000 euro-PART
 ‘The price of this car is higher than 25000 Euro’
 b. *Tämä-n auto-n hinta on yli 25000 euro-n*
 this-GEN car-GEN price.NOM is over 25000 euro-GEN
 ‘The price of this car is higher than 25000 Euro’

The case alternation of this adposition only occurs in its non-spatial use – genitive case always being used for spatial meaning – and rather lies in pragmatics (see Lestrade 2006 for a more elaborate discussion).

² Abbreviations used in this paper: 1,2,3 first, second, third person; ADE adessive; ALL allative; GEN genitive; INF infinitive; NOM nominative; PART partitive; PAST past tense; PL plural; PP adposition.

³ A less marked way of expressing the literal meaning in (4a) would be the following:

- (i) *tama on minu-n lähellä-ni*
 this.NOM is 1sg-GEN near-1SG.POSS
 ‘This is close to me’

In this case, genitive case is used in combination with a postposition to which a co-referring morpheme is added.

⁴ The change in form of the stem *sopimuskaude-* ‘contract season’ is phonologically driven and is not meaningful in any way.

⁵ Note that markedness per definition is a relative notion: Some form or meaning is never marked per se, but only in relation to its alternatives.

⁶ An anonymous reviewer noted that the proposed grammar does not account for example (5). It wrongly predicts that that if $\langle \text{yli} + \text{GEN} \rangle$ exists, $\langle \text{yli} + \text{PART} \rangle$ should exist as well. Except for the example described in Footnote 1, this alternation is not productive. However, this prediction only follows from a (unidirectional) production perspective. *Bidirectional* OT does not so much predict which forms are grammatical, but rather accounts for combinations of existing forms and meanings. Indeed, the BiOT grammar correctly predicts that given the form candidates in (5) and Footnote 1, the meanings should be distributed as they are.

⁷ <https://hotpage.csc.fi>, consulted in the summer of 2006.

⁸ If we take as our null hypothesis that there is no preference for either construction or case, we expect a probability of .5 for both constructions and cases to occur. The chance of getting the distribution of Table 1 under this assumption approaches zero, hence, we can say that the prepositional construction and the partitive case occur significantly less often than the postpositional construction and the genitive case. (For this and all other computations R is used.)

⁹ As the reader may note, in Table 2 a number of 64 was mentioned. Here, I count all nouns that only occur with partitive case with one adposition, but possibly with genitive or both cases with some other.