

Tal Linzen

Parallels between cross-linguistic and language-internal variation in Hebrew possessive constructions

Abstract: Grammatical constraints in one language often surface as statistical tendencies in another, suggesting that cross-linguistic comparative studies can play a central role in the study of language-internal “free” variation. This paper applies this approach to the case of the variation between two Hebrew constructions: possessive dative (PD) and ordinary possession. While both constructions convey a possessive meaning, PD additionally highlights the fact that the possessor was *affected* by an event involving his or her possessed object. To elucidate the concept of affectedness, we turn to European languages that have encoded in their grammar various concrete reflexes of this notion, such as the animacy of the possessor (animate possessors are more often perceived as affected). We show that these concrete reflexes, while not grammatically encoded in Hebrew, have a statistical effect in that language as well. This makes it possible to predict the choice of construction in any given context using these objective proxies of affectedness. Furthermore, we argue that certain categorical restrictions on PD, previously attributed to formal syntactic factors, are best captured as consequences of the semantic affectedness condition. Our results illustrate the continuum between categorical constraints and statistical tendencies, both across languages and within a single language.

Keywords: variation, possessive dative, Hebrew, external possession, affectedness

Tal Linzen: Department of Linguistics, New York University, 10 Washington Place, New York, NY 10003, USA. E-mail: linzen@nyu.edu

1 Introduction¹

Grammatical restrictions in one language often mirror statistical tendencies in another. One well-known example is the role of the person hierarchy (1st,

¹ This paper uses the following abbreviations: ACC: accusative marker; CS: construct state; PD: possessive dative; OP: ordinary possession. Transcription key: š = [ʃ], x = [x], c = [ts].

2nd > 3rd) in passivization. English does not impose any restrictions on the person of the subject in a passive sentence: *the man is known by me* is a perfectly grammatical sentence of English. A number of languages, however, require the person of the subject to be the highest in the clause. In these languages passive sentences cannot have a third person subject and a first person nonsubject argument. One example is the Salish language Lummi (Jelinek and Demers 1983): the Lummi equivalent of *the man is known by me* would be ungrammatical. Corpus studies of English have shown that these configurations, though attested, are statistically disfavored in English as well (Bresnan et al. 2001; Aissen 1999). Such parallels between grammatical constraints in one language and statistical tendencies in another follow naturally from the assumption that patterns that are frequent in discourse are more likely to become conventionalized as grammatical principles (Bybee and Hopper 2001; Hopper and Traugott 2003; Ariel 2008).

These parallels between categorical and statistical constraints across languages suggest a methodology for the study of “free” variation between two forms within a single language. In the first stage, the researcher identifies other languages that exhibit alternations similar to the one in question. In some of those languages, the alternation may be subject to categorical constraints mandating the use of only one of the forms in certain contexts. These categorical constraints are much more likely than statistical tendencies to be documented in grammars. The typologically informed constraints obtained in this way can then be applied in the language of interest to explain speakers’ choice between the two possible forms. This methodological intuition underlies much work on phonological variation (Boersma and Hayes 2001), but is less often applied in the syntactic domain.

This paper applies this methodology to the case of the variation between two Hebrew possessive constructions: the possessive dative (PD) and ordinary possession (OP). In the Hebrew PD, a dative argument of the verb is interpreted as the possessor of another entity in the sentence:

(1) POSSESSIVE DATIVE (PD)

hu šavar [li] [et ha-yad].

He broke to.me ACC the-arm

‘He broke my arm.’

A typical breaking event has two core participants, the breaker and the broken object. By contrast, in PD sentences such as (1), the verb has three syntactic arguments: the two core participants, and in addition the possessor of the broken object. Alongside the possessive dative construction, Hebrew has an ordinary

possessive construction, in which the possessor and the possessed object (the possessum) form a single noun phrase:

- (2) ORDINARY POSSESSION (OP)
hu šavar [et ha-yad šeli].
 He broke ACC the-arm my.
 ‘He broke my arm.’

The verb in (2) has two syntactic arguments, which correspond to the expected core semantic roles: the breaker and the broken object.

A PD sentence and the corresponding OP sentence typically have similar truth conditions, as illustrated by the identical English glosses of (1) and (2). In addition to their truth conditional meaning, however, PD sentences typically indicate that the possessor was *affected* by the event (Berman 1982; Landau 1999; Linzen 2009). In other words, PD is used when the event is perceived as having “happened to” the possessor, even though strictly speaking it only “happened to” the possessum (Wierzbicka 1988). Existing accounts of the Hebrew PD do not offer criteria to help determine whether the possessor in a given event is likely to be perceived as affected by the event. One goal of this paper is to establish such criteria, based on a cross-linguistic comparative study of the construction.

Numerous languages have constructions in which the possessor appears as an argument of the verb rather than as a modifier of the possessed object, collectively known as external possession constructions (see Payne and Barshi 1999 for a review). However, the use of the dative case to mark the external possessor, as in Hebrew, is largely restricted to the European linguistic area. European languages typically impose an affectedness restriction of some sort on external possession, akin to the Hebrew implication of affectedness (König and Haspelmath 1998). Indeed, affectedness plays a pivotal role in dative constructions in general, both cross-linguistically and in Hebrew (Ariel et al. submitted). Finally, there is evidence that the Hebrew PD is a borrowing from European languages (Berman 1982). It is therefore natural to turn to European languages as a source of criteria for affectedness.

The restrictions that European languages place on PD often reflect concrete correlates of affectedness. In some languages, for instance, PD can only be used when the possessum is a body part. This is plausibly a grammatically encoded reflex of affectedness: the possessor is more clearly affected by an event that occurred to a possessum if that possessum is attached to the possessor’s body (e.g., a hand) than if it is not (a window). As opposed to the degree of possessor affectedness in a given situation, which can only be assessed using the analyst’s subjective intuition, these concrete correlates of affectedness can be objectively

identified: it is easy to decide whether or not the possessum in a given sentence is a body part. This makes it possible to automatically tag sentences in a corpus for the existence of these factors, and assess their respective contributions to the likelihood that a speaker would use PD in a given context.

We examine three families of typologically informed factors, inspired by Shibatani (1994) and König and Haspelmath (1998):

1. Salient possessor: PD is favored if the possessor is salient in the discourse. When the possessor is salient, the speaker is more likely to assume the possessor's point of view and describe the event as having *happened to* the possessor.
2. Personal-sphere possessum: PD is favored if the possessum is in the personal sphere of the possessor, ideally constituting a part of the possessor. Anything that happens to a part of the possessor happens by extension to the possessor as a whole.
3. Transitive event: PD is favored if the event is transitive, in the sense of Hopper and Thompson (1980): someone *did something* to the possessum, or minimally something *happened to* it.

We do not make the claim that affectedness can be reduced to these quantifiable factors. Indeed, it is unlikely that objective correlates of affectedness can ever fully explain the alternation between PD and OP: in any given context, regardless of how highly affected the possessor is, objectively speaking, a speaker can choose to deemphasize the affectedness of the possessor by using OP instead of the expected PD. The best outcome we can hope for is a statistical model that predicts with reasonable accuracy which of the constructions would be used in a given context.

In summary, the first goal of this paper is to show that these three factors affect the statistical variation between PD and OP, in cases when both constructions are grammatical. Alongside the variation between the constructions, however, there are cases in which only one of the constructions can be used, even in Hebrew. The second goal of this paper is to propose an explanatory role for affectedness in cases in which PD is ungrammatical, as in (3):

- (3) a. *ani ohev et ha-se'ar šel dana.* (OP)
 I like ACC the-hair of Dana
- b. **ani ohev le-dana et ha-se'ar.* (PD)
 I like to.Data ACC the-hair
 'I like Dana's hair.'

Previous accounts of PD have attributed the restricted distribution of the construction to fundamental differences in the syntactic structures associated with

different verb classes. In the most thorough transformational account of the Hebrew PD, Landau (1999) argues that the possessor is generated as a modifier of the possessum, and then moves to a position adjacent to the verb. This syntactic position is available when the verb is agentive (e.g. *break*), but not when it is an experiencer verb (*like*), deriving the contrast in (3). Other restrictions on the distribution of PD, such as the difference between unaccusative and unergative verbs (see Section 5), are likewise explained in terms of the syntactic reflexes of the verb's argument structure (Borer and Grodzinsky 1986).

While the transformational accounts acknowledge that PD conveys an “implication” of affectedness (Borer and Grodzinsky 1986; Landau 1999), this fact does not play an explanatory role in accounting for contrasts such as (3). In fact, this “implication” is not represented in the structure at all, making it unclear how it arises. In what follows, we will show that many of the distributional restrictions on PD, attributed by these authors to syntactic constraints, can be reduced to the affectedness condition, which we take to be part of the core meaning of the construction (see Pylkkänen 2002 for a related proposal). In (3), for instance, Dana is presumably more likely to be perceived as affected if someone breaks her arm than if someone likes her hair. Furthermore, in many cases the acceptability of PD is determined by the verb's complements rather than by the verb itself, rendering unsatisfactory any explanation that only makes reference to verb classes. Since affectedness is independently needed to derive the statistical pattern of variation between PD and OP, an account that explains the distribution of PD using affectedness only is more parsimonious than one that invokes both affectedness and restrictions on syntactic movement.

The paper is structured as follows. The first three sections take up the three proxies of affectedness in turn: Section 2 discusses the role of the discourse salience of the possessor, Section 3 looks into the personal sphere factor and Section 4 discusses the transitivity of the event. Section 5 evaluates the relationship between PD and unaccusative verbs (Borer and Grodzinsky 1986) in light of the results presented in this paper, and suggests that this relationship may be an epiphenomenon of the affectedness condition. Section 6 concludes with a general discussion. The paper includes an Appendix, which outlines the methodological approach of this paper and presents the corpus used in the quantitative analyses.

2 The possessor: Discourse salience

The first correlate of possessor affectedness that we examine in this paper is the discourse salience of the possessor. König and Haspelmath (1998) note that

the possessive dative is cross-linguistically more likely to be grammatical when the possessor is high on the following scale (slightly adapted):

(4) DISCOURSE SALIENCE SCALE

First person pronoun < second and third person pronouns < animate < inanimate

This scale is termed the animacy or empathy scale (Silverstein 1976; Seiler 1983; Kuno 1987). The higher an entity is on the scale, the more the speaker is likely to “empathize” with it, that is, to take its point of view when describing a state of affairs. The relevance of this scale for the possessive dative is clear: a highly salient discourse entity is more likely to be portrayed as affected by an event, even when in reality the event happens to an entity separate from it. The possessor affectedness theory therefore predicts that possessors that are high on this scale, such as first person pronouns, will favor PD to a much larger degree than those that are low on this scale. This section will survey in turn the three distinctions predicted by the scale outlined in (4): person, pronominality and animacy.

2.1 Person of pronoun

We first compare first person pronouns to other pronouns. The speaker is typically the most salient entity in the discourse, and thus the one whose point of view is most naturally adopted. According to the empathy scale, then, speakers are expected to be especially likely to portray themselves as affected by something that has happened to their own possessum. This indeed turns out to be the case: first person pronouns make up a large part of PD possessors in our corpus. Around 80% of all pronominal possessors used with PD were first person pronouns, compared to only 54% for OP sentences. Another way to interpret the counts is that the possessive dative is used roughly once in every three possessive sentences when the possessor is a first person pronoun, and only once in every eight sentences when the possessor is a second or third person pronoun. Interestingly, a similar result was reported for Latin as early as Havers (1911).

2.2 Pronouns versus full noun phrases

Pronouns are used to refer to entities that are more salient in the discourse than lexical noun phrases (Ariel 1990). The empathy scale therefore predicts that the possessive dative will be more common with pronominal possessors than with full noun phrase possessors. In some languages, this tendency takes the form of a grammatical restriction, for example in Portuguese:

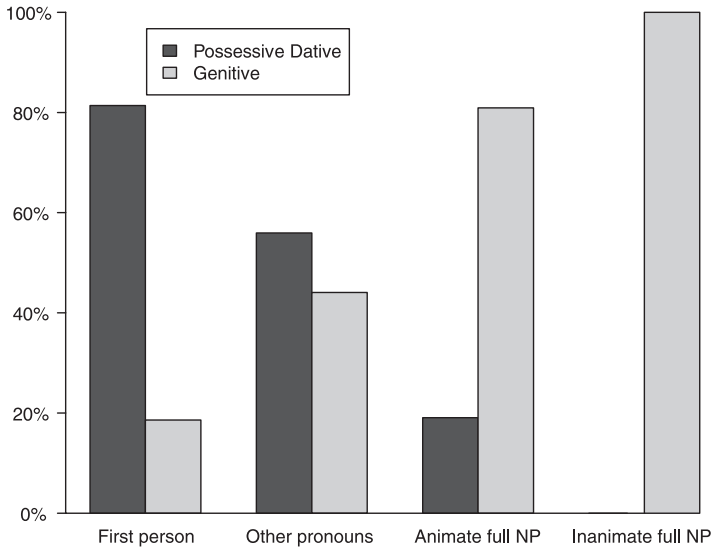


Fig. 1: Effect of possessor discourse salience on the choice of possessive construction: the more salient the possessor, the more likely is the possessive dative to be used. First person pronouns are naturally salient: the speaker most often takes his or her own perspective. Pronouns in general are used to refer to entities that have been mentioned in the discourse. Finally, animate nouns are more salient than inanimate ones; speakers are more likely to assume the perspective of another person than that of an inanimate object.

(5) PORTUGUESE

- a. **A mãe está lavando os cabelos ao menino.*
 the mother is washing the hairs to.the child
 ‘The mother is washing the child’s hair.’
- b. *A mãe está lhe/te/me lavando os cabelos.*
 the mother is to.him/to.you/to.me washing the hairs.
 ‘The mother is washing his/your/my hair.’
 (König and Haspelmath 1998: 570)

Even though Hebrew does not encode this distinction in the grammar of the construction, we still expect to see a higher proportion of pronouns than full noun phrases as PD possessors. This prediction is borne out strikingly: 95% of the PD possessors in the sub-corpus examined were pronouns (including reflexive pronouns, such as *acmi* ‘myself’). By comparison, only 70% of OP possessors were pronominal. Looked at another way, the probability of choosing PD over OP, which is around 25% when the possessor is a pronoun, drops to only 4% when it is a full noun phrase.

2.3 Animacy

Most European languages place the grammatical cut-off point on the empathy scale immediately after the animate nouns category. In other words, the possessive dative construction in these languages allows animate possessors, either pronominal or full noun phrases, but not inanimate possessors:

(6) GERMAN

- a. *Der Stein fällt dem Mann auf den Kopf.*
 the stone falls the:DAT man on the head
 ‘The stone falls on the man’s head.’
- b. **Der Stein fällt dem Auto aufs Dach.*
 the stone falls the:DAT car on.the roof
 ‘The stone falls on the roof of the car.’
 (Neumann 1996)

This is not the case for Hebrew, as illustrated by the following grammatical examples:

- (7) a. *im lo holex, az ešar le-haxlif la-mexonit*
 if not goes, then possible to-replace to.the-car
et ha-manoa. (PD) (Web)
 ACC the-engine
 ‘If it doesn’t work out, you can replace the car’s engine.’
- b. *nira li dafakti la-maxšev et*
 seems to.me I.screwed.up to.the-computer ACC
ha-sapak. (PD) (Web)
 the-power.supply
 ‘Looks like I screwed up the computer’s power supply.’
- c. *xaval la-haros la-ir et ha-tadmit rak biglal*
 pity to-destroy to.the-city ACC the-image only because
moaca metupešet. (PD) (Web)
 council stupid
 ‘It’s a shame to destroy the city’s image just because of a stupid city council.’

While inanimate possessors are in principle grammatical, they are infrequent: out of 25 PD sentences in the present sample that had a full noun phrase possessor, none had inanimate possessors. Inanimate possessors were the minority in OP sentences as well, but a sizable one at 25%, a proportion that is significantly

higher than the 0% proportion for PD sentences ($p = 0.002$, Fisher's exact test). In sum, we see that the Hebrew PD disprefers inanimate possessors, as predicted by the possessor salience hierarchy.

Further supporting the role of animacy, many of the inanimate PD possessors found in Google searches referred to machines, such as *car* and *computer* in (7) above. This is in line with Dąbrowska's (1998) observation that machines are often viewed as more animate than inert physical objects.

The role of animacy gains further support from the fact that the Hebrew PD disprefers dead possessors. The following sentence is marginal with PD if the mother is no longer alive at the time of the event, but perfect with OP:

- (8) a. *cavati et ha-xeder šel im-i ha-menoxa.* (OP)
 I.painted ACC the-room of mother-my the-late
 'I painted my late mother's room.'
- b. *?cavati le-im-i ha-menoxa et ha-xeder.* (PD)
 I.painted to-mother-my the-late ACC the-room
 'I painted my late mother's room.'

Similar facts are presented for German in König and Haspelmath (1998: 531).

3 The possessum: body parts and the personal sphere

A possessor is most obviously affected by an event that happened to its possessum if the possessum is a part of the possessor: if something happened to a part, it has also happened to the whole. This applies most naturally to body parts. In French, for example, body parts are acceptable PD possessums, but material possessions removed from the body are not:

- (9) FRENCH
- a. *Je lui ai cassé le bras.*
 I to.him have broken the arm
 'I broke his arm.'
- b. **Je lui ai cassé la fenêtre.*
 I to.him have broken the window
 'I broke his window.'
- (König and Haspelmath 1998: 572)

Czech is an example of a language that allows both body-part and non-body-part possessums in PD, but *disallows* the use of OP with body parts:

(10) CZECH

- a. Šlapal jí na nohy.
 step:PP:SG:MASC 3SG:FEM:DAT on foot:ACC:PL:FEM
 ‘He stepped on her feet.’
- b. #Šlapal na její nohy.
 step:PP:SG:MASC on her:ACC foot:ACC:PL:FEM
 ‘He stepped on some feet of hers.’
 (Fried 1999: 482)

Fried reports that (10b) “invokes the image of a heap of foot-like objects being stepped on”. Hebrew does not grammatically enforce the preference for body part possessums in PD in either the French or the Czech way. Hebrew PD is grammatical with non-body-part possessums, as illustrated in (11):

- (11) *hu šavar li et ha-xalon.* (PD)
 he broke to.me ACC the-window
 ‘He broke my window.’

At the same time, OP is fine with body-part possessums, as shown in (12):

- (12) *hu darax al ha-regel šela.* (OP)
 he stepped on the-foot her
 ‘He stepped on her foot.’

Even though the distinction between body part possessums and other possessums does not play a role in the categorical grammar of the Hebrew PD, there is ample statistical evidence for the preference for body parts in the Hebrew PD. As a general illustration, a Google search for the two variants of “he broke my arms” turned up 48 matches for the PD form, and only 3 for the OP form, despite the fact that PD is in general less frequent than OP.

A more systematic survey of the possessive sentences corpus (described in Section A.1) revealed that 16% of PD sentences had a body part possessum, compared to only 5% of OP sentences ($p < 0.001$, Fisher’s exact test). Since in total there are around four times as many OP sentences as PD sentences, this means that when faced with the choice of possessive construction for a body part possessum, speakers are almost as likely to choose PD (45% of the cases) as OP (55% of the cases). In contrast, when the possessum is not one of the common body parts,

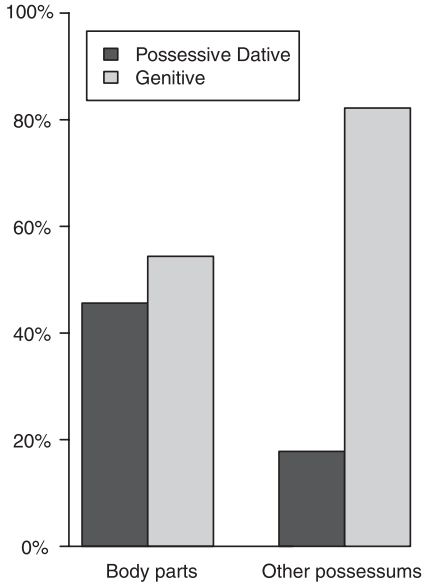


Fig. 2: Effect of type of possessum on the choice of possessive construction: PD is much more likely to be used with body part possessums than with other possessums. A possessor is more often perceived as affected by an action performed on his or her own body than by an action performed on any other object.

the proportion of PD sentences drops to 18% (see Figure 2). We see again that a grammatical restriction in other languages shows up as a dominant discourse pattern in Hebrew.

3.1 Extending the part-whole relation: the personal sphere

The notion of a part-whole relationship between possessum and possessor is often expanded beyond the possessor's body. In French, for example, clothing items are acceptable as PD possessums, as long as the possessor is wearing them at the time of the event:

(13) FRENCH

Je lui ai cassé les lunettes.

I to.him have broken the glasses

'I broke his glasses.'

(König and Haspelmath 1998: 572)

In fact, anything that is conceptualized as part of the possessor's personal sphere can serve as a PD possessum:

(14) FRENCH

On lui a tiré dans les pneus.
 one to.him has shot in the tires
 'People shot at his tires.'
 (Diffloth 1974)

Diffloth notes that (14) is only acceptable if "he was in the car, in fact in the driver's seat, at the time of shooting, with the whole vehicle, tires included, considered to be in his personal sphere" (p. 132). Moreover, an object can be part of the possessor's personal sphere in a transient way. In (13), for example, the glasses need not be permanently owned by the possessor; in fact, they can belong to someone else, as illustrated in (15):

(15) HEBREW

gil šavar le-rina et ha-miškafayim šel sigal.
 Gil broke to-Rina ACC the-glasses of Sigal.
 'Gil broke Sigal's glasses while Rina was wearing them.'
 (based on Landau 1999: 7)

The crucial distinction is whether they were within the possessor's personal sphere at the time of the event. If they were, anything that happened to them is perceived as having happened to the possessor.

The personal sphere condition also extends to mental qualities, which are relatively common as PD possessums:

- (16) a. *ze poge'a li ba-bitaxon ha-acmi.* (PD) (Web)
 it hurts to.me in.the-confidence
 'It hurts my confidence.'
 b. *šipartem li et macav ha-ruax hayom.* (PD) (Web)
 you.improved to.me ACC the mood today
 'You improved my mood today.'

The personal sphere condition is usually referred to as the inalienability condition (Shibatani 1994, Haspelmath 1999). For Hebrew, at least, this name would be inaccurate. An inalienable possession is normally defined as "in-born, inherent, not conferred by purchase" (Nichols 1988: 568). This applies

equally to kinship terms and to parts of a whole, including body parts. Yet in Hebrew kinship terms make relatively poor PD possessums, as discussed below.²

When both a personal-sphere reading and a neutral possessive reading are available, it is usually the personal-sphere reading that is favored. For example, the salient reading of (17) is one in which the book the speaker was holding fell out of her hand:

- (17) *nafal li ha-sefer al ha-ricpa.* (PD) (Web)
 fell to.me the-book on the-floor
 ‘My book fell on the floor.’

While the reading in which the speaker’s book was lying on a table and fell off is not impossible, it requires more context. Similarly, example (18) generally means that the speaker’s friend took the book away from the speaker, possibly when she was holding it in her hands:

- (18) *ha-xaver šeli lakax li et ha-sefer.* (PD) (Web)
 the-friend my took to.me ACC the-book
 ‘My friend took my book.’

The alternative reading, where the friend simply took the speaker’s book (and, say, placed it on a shelf), is grammatical but less accessible.

Hebrew is relatively liberal with respect to the personal sphere restriction. For example, sentence (20), the Hebrew equivalent of the French (9b) – repeated here as (19) – is grammatical:

- (19) FRENCH
 **Je lui ai cassé la fenêtre.*
 I to.him have broken the window
 ‘I broke his window.’

² Fried (1999) reports that in Czech kinship relations are in fact more common as possessums than proximate alienable objects. She proposes the following inalienability scale:

(i) Body parts < kinship relations < close alienable entities < distant alienable entities

The axis of extension seems to vary between languages. This indicates that there is an arbitrary and learned dimension to the distinctions that different languages make, which cannot be reduced to cognitive universals.

(20) HEBREW

hu ca'ak alai ki šavarti lo et ha-xalon. (PD) (Web)
 he yelled at.me because I.broke to.him ACC the-window.
 'He yelled at me because I broke his window.'

We see that material belongings can be seen as an extension of the personal sphere (though, again, they are less common PD possessums than are body parts). On the other hand, the construction does not extend as easily to human possessums (see also Gafter 2010):

(21) a. *hu hika et ha-more šel dani.* (OP)

he hit ACC the-teacher of Dani
 'He hit Dani's teacher.'

b. *?hu hika le-dani et ha-more.* (PD)

he hit to-Dani ACC the-teacher
 'He hit Dani's teacher.'

Kinship terms and other relational nouns make slightly better possessums, as the following attested examples show:

(22) a. *nakamti ki hargu li et ha-yeled.* (PD) (Web)

I.revenged because they.killed to.me ACC the-son
 'I revenged because they killed my son.'

b. *ha-saxkan misken, metu lo ha-horim.* (PD) (Web)

the-player unfortunate, died to.him the-parents
 'This (basketball) player is unfortunate, his parents died.'

c. *yagidu še-ani ba'a kedey litfos la*

they.will.say that-I come in.order.to catch:INF to.her
 ACC the-boyfriend

'They'll say I'm going there to steal her boyfriend.'

Though not ungrammatical, human possessums are uncommon in PD sentences: none of the 1124 PD possessums in the manually annotated sample was human.

In sum, we have seen that body parts are the prototypical and most frequent type of PD possessum. At the same time, the grammar of the Hebrew PD allows considerable leeway for possessums that are not strictly speaking in the posses-

sor's personal sphere, though those are not as likely to be expressed with PD as personal sphere possessums. Based on the discussion above, the following tentative scale emerges:³

(23) THE PERSONAL SPHERE SCALE

Body parts < parts of a whole < clothing items and other objects
 in the possessor's personal sphere < mental properties < material possessions
 < relational human possessums < other human possessums

Despite the broad range of possessums that are in principle grammatical, the further down the possessum is on the scale, the less likely it is to be used in PD.

3.2 PD and non-possessive uses of OP

The personal sphere scale proposed in (23) only covers a subset of the contexts in which OP can be used. Consider, for example, the distinction between creators and owners, illustrated in (24):

- (24) a. *be-zman še-moti haya ba-šerutim axalti*
 at-time that-Moti was at.the-bathroom I.ate
lo et ha-stek. (PD)
 to.him ACC the-steak
 'When Moti was at the bathroom I ate his steak.'
- b. **šamati še-moti šef me'ule, aval od lo*
 I.have.heard that-moti chef great but still not
axalti lo et ha-stekim. (PD)
 I.ate to.him ACC the-steaks
 'I've heard that Moti is a great chef, but I haven't eaten his steaks yet.'
- c. *šamati še-moti šef me'ule, aval od lo*
 I.have.heard that-moti chef great but still not
axalti et ha-stekim šelo. (OP)
 I.ate ACC the-steaks his
 'I've heard that Moti is a great chef, but I haven't eaten his steaks yet.'

³ The corpus used in this paper did not allow us to present quantitative evidence for all of the rankings – for example, we did not have enough tokens to substantiate the claim that mental properties make better PD possessums than material possessions. We leave this question to future research.

PD can be used if the steaks belong to Moti (24a) (for example, if he is about to eat them), but not when they are the steaks cooked by him (24b). OP, on the other hand, can be used in either context, as demonstrated in (24c).

It should be noted that Landau (1999: 5) argues that creators are in fact acceptable PD possessors, based on the following example (his 5b):

- (25) *gil higdil le-rina et ha-tmuna.* (PD)
 Gil enlarged to-Rina ACC the-picture.
 ‘Gil enlarged Rina’s picture.’ (Rina = creator/possessor)

Landau’s claim is that (25) can be used whenever Rina was the person who took the picture, even if it didn’t belong to her at the time of the picture enlarging event. In my judgment, this claim is incorrect; sentence (25) is only acceptable when the creator is also the owner of her creation. This contrast is brought out more clearly in (24) above.

This contrast between creators and owners is only surprising if we assume that every kind of relation that can be expressed using the OP construction can also be expressed using PD (Landau 1999), as long as it is embedded in an appropriate syntactic structure (in Landau’s framework, a Specifier position). In reality, only the particular set of relations listed in (23), formed by the gradual relaxation of the personal sphere condition, is compatible with PD. An object created by a person is not necessarily in that person’s personal sphere; a sentence such as *I read Wittgenstein’s book* is not going to be expressible using PD.

4 The event: Transitivity

We now turn to the last correlate of affectedness proposed in this paper: the transitivity of the event. This is a generalization of König and Haspelmath’s (1998) “situation hierarchy” and “syntactic relation hierarchy”, which they use to characterize the distribution of PD in the languages of Europe. The possessive dative is most commonly used to describe an event that happened to the possessum, and, by extension, to the possessor. Hopper and Thompson’s (1980) notion of transitivity offers a way to assess the degree to which the event “happened to” an entity – in their words, “the effectiveness or intensity with which the action is transferred” (p. 252). This notion can be decomposed into a set of features, three of which will be examined in this section. Section 4.1 considers the number of event participants mentioned in the sentence, and shows that PD is favored when the entity acting on the possessum is explicitly mentioned in the sentence – that is, with transitive rather than intransitive verbs (König and Haspelmath’s syntac-

tic relations hierarchy). Section 4.2 shows that PD is more commonly used to describe events, which involve a clear action, than states (König and Haspelmath's situation hierarchy proper). Finally, Section 4.3 shows that volitional actions are more compatible with PD than involuntary ones. All of these factors are gradient: intransitive verbs, states and involuntary actions are all attested with PD, but are disfavored.

4.1 Number of participants

There is a clearer “transfer of action” when the event involves two participants than when it only involves one. Hence we expect PD to be more common with transitive than with intransitive verbs. This distinction is grammatically enforced in Russian, where PD can only be used with transitive verbs:

(26) RUSSIAN

a. *Ona nastupila emu na nogu.*

she walked to.him on leg

‘She walked [stepped] on his leg.’

b. **Mne drožat ruki.*

to.me shake hands

‘My hands are shaking.’

(König and Haspelmath 1998: 532, 538)

We assessed the effect of this distinction on the choice of possessive construction in Hebrew using Google searches. The searches were performed for both possessive constructions, with both the transitive and intransitive (unaccusative) forms of the verb *haras* ‘destroy’ / *neheras* ‘was destroyed’. These two forms were combined with three common possessums: *bayit* ‘house’, *maxšev* ‘computer’ and *xaim* ‘life’.⁴ The results are striking: with the intransitive form of the verb, 45% of the matches (208 out of 455) were PD sentences; with the transitive form, this proportion increased to 91% (270 out of 296), a highly significant difference ($p < 0.001$, Fisher’s exact test).

⁴ A potential confound is the fact that intransitive verbs with PD typically precede the subject, which is not the case for most Hebrew sentences, including the OP sentences in question (see Melnik 2002 for discussion). For this reason, queries were performed both with the intransitive verb preceding the subject and with the verb following the subject.

4.2 Events versus states

Another factor that influences the transitivity of a situation is its eventivity: whether it involves an action or only a state. This factor has a dramatic effect on the acceptability of PD. Consider the contrast between the action in (27a) and the mental state in (27b) (after Shibatani 1994):

- (27) a. *ha-sapar šeli icev la et ha-se'ar.* (PD) (Web)
 the-hairdresser my styled to.her ACC the-hair.
 'My hairdresser did her hair.'
- b. **ha-sapar šeli ahav la et ha-se'ar.* (PD)
 the-hairdresser my liked to.her ACC the-hair.
 'My hairdresser liked her hair.'

Eventivity also captures the fact that PD is disfavored in predicational sentences, which refer to states and not to events (Horvath and Siloni 2008). This is a tendency rather than an inviolable constraint, as illustrated by the following examples:

- (28) a. *meluxlax li ha-masax.* (PD) (Web)
 dirty to.me the-screen
 'My screen is dirty.'
- b. *adumot lexa me'od ha-einayim.* (PD) (Web)
 red to.you very the-eyes
 'Your eyes are very red.'

Even though there are isolated cases of stative predicational sentences involving the possessive dative, as shown in (29), each of these situations is normally expressed using the OP construction. Google searches for the analogous OP sentences turned up significantly more matches:

- (29) a. *ha-masax šeli meluxlax.* (OP) (14 matches)
 the-screen my dirty.
 'My screen is dirty.'
- b. *ha-einayim šelxa me'od adumot.* (OP) (69 matches)
 the-eyes your very red.
 'Your eyes are very red.'

While not frequent in PD, adjectival passives as in (28a) are consistently attested in this construction. This is not surprising from the transitivity perspective: adject-

tival passives often implicitly refer to the action that has brought about the state they describe (Meltzer-Asscher 2011), and hence are more transitive than simple adjectives such as *red*. Other examples found online are *mekulkal* ‘malfunctioning’ and *šavur* ‘broken (glass)’:

- (30) a. *kodem haya mekulkal li ha-maxšev.* (PD) (Web)
 earlier was broken to.me the-computer
 ‘Earlier my computer was broken.’
 b. *yeš li patifon aval švura li*
 exists to.me record.player but broken to.me
ha-maxat. (PD) (Web)
 the-needle
 ‘I have a record player but my needle is broken.’

On the other hand, PD is absolutely incompatible with individual-level adjectives, which refer to permanent properties (as opposed to stage level adjectives, which refer to temporary states):

- (31) a. *ha-einayim šela kxulot.* (OP)
 the-eyes her blue
 ‘Her eyes are blue.’
 b. **kxulot la ha-einayim.* (PD)
 blue to.her the-eyes
 ‘Her eyes are blue (she has blue eyes).’

The distinction between stage level and individual level predicates, or between temporary states and permanent properties, is not one of Hopper and Thompson’s features of transitivity, but constitutes a natural extension of them: the state of being dirty could be perceived as something that is happening to the screen in (28a), or at least showing the effects of an action that has happened to it in the past. The condition of being blue, on the other hand, is not something that is “happening to” the eyes, under any circumstances.

The following eventivity scale emerges:

(32) EVENTIVITY SCALE

Actions < adjectival passives < other stage-level predicates < individual-level predicates

The higher the predicate is on the scale is (32), the more likely the speaker is to use PD.

4.3 Volitionality: the case of perception events

If someone is performing an action intentionally, the “transfer of activity” associated with the action is cross-linguistically perceived to be stronger than if it happened by accident, even if the outcome in the world was identical (Hopper and Thompson 1980). Indeed, volitional events are more common with PD than non-volitional ones. Landau (1999) notes, for example, that the agentive *histakel* ‘look’ is acceptable in PD in a broader set of contexts than *ra’a* ‘see’:

- (33) a. *Gil ra’a et ha-bayit šel Rina.* (OP)
 Gil saw ACC the-house of Rina.
 ‘Gil saw Rina’s house.’
- b. ??*Gil ra’a le-Rina et ha-bayit.* (PD)
 Gil saw to-Rina ACC the-house
 ‘Gil saw Rina’s house.’
- (34) a. *Gil histakel al ha-bayit šel Rina.* (OP)
 Gil looked on the-house of Rina.
 ‘Gil looked at Rina’s house.’
- b. *Gil histakel le-Rina al ha-bayit.* (PD)
 Gil looked to-Rina on the-house
 ‘Gil looked at Rina’s house.’

While there is no question that there is a difference in acceptability between (33b) and (34b), this difference appears to be a matter of degree, rather than a clear-cut grammaticality contrast. First, as noted in the literature, *ra’a* ‘saw’ is compatible with PD in invasion-of-privacy contexts, particularly with intimate body parts and clothing items, but not exclusively:

- (35) a. *hi hayta im maxsof kaved ve-hu [ra’a la*
 she was with neckline heavy and-he saw to.her
et ha-xazia] (PD) (Web)
 ACC the-bra]
 ‘She was wearing a shirt with a low neckline and he could see her bra.’
- b. *hi [ra’ata lo et kol ha-klafim] ve-gilta*
 she saw to.him ACC all the-cards and-discovered
še-ma še-hu mastir behexlet lo
 that-what that-he was.hiding definitely not
to'em et ha-sxum še-hu himer alav. (PD) (Web)
 matches ACC the-amount that-he bet on.it
 ‘She saw his cards and discovered that what he was hiding definitely did not match the amount he had wagered.’

Invasion of privacy contexts are the most typical cases in which a possessor might be affected by a seeing event. In other contexts, the possessor is not even aware of the fact that someone is seeing the possessum, and hence the possessor is not affected by the event. In fact, the seeing event reduces to a state, much like sentences involving experiencer verbs such as *ahav* ‘loved’ (see Section 4.2 above). Invasion-of-privacy contexts, on the other hand, imply that the observer was making a conscious effort to look. This is not a unique Hebrew idiosyncrasy: the compatibility with PD of seeing verbs in invasion-of-privacy contexts has been reported for French (Diffloth 1974), Finnish (Pylkkänen 2002) and Italian (König and Haspelmath 1998).

Going back to *histakel* ‘looked’, we find that every single Google match in which this verb appears in PD had a body part or clothing item possessum (51 out of 51), whereas in OP this was only true of around half of the examples (42 out of 80).⁵ The difference is highly significant ($p < 0.001$, Fisher’s exact test). This indicates that *histakel* ‘looked’ shows a similar behavior to that of *ra’a* ‘saw’: the possessor is perceived as affected by an act of looking predominantly when the possessum is something she was trying to hide. To sum up, the volitionality of the action contributes to the acceptability of PD, consistent with the role of the transitivity of the event. However, contra Landau (1999), the difference between the volitional *histakel* ‘looked’ and the non-volitional *ra’a* ‘saw’ seems to be one of degree rather than a categorical distinction.

4.4 Additional proposed verb class restrictions

Landau (1999) identifies several other verb classes that he argues are incompatible with PD. One of these verb classes consists of subject matter verbs, such as *talk* and *discuss*. Landau cites the following judgments as evidence for the incompatibility of subject matter verbs with PD (p. 33):

- (36) a. *Gil diber al ha-avoda šel Rina.* (OP)
 Gil talked about the-work of Rina
 b. **Gil diber le-Rina al ha-avoda.* (PD)
 Gil talked to-Rina about the-work
 ‘Gil talked about Rina’s work.’

⁵ The search strings were *histakel la al* ‘he.looked to.her on’ for PD and *histakel al * šela* ‘he.looked on * her’ for OP.

- (37) a. *ha-va'ada dana ba-nisui šel rina* (OP).
 the-committee discussed in.the-experiment of Rina
 b. **ha-va'ada dana le-Rina ba-nisui*. (PD)
 the-committee discussed to-Rina in.the-experiment
 'The committee discussed Rina's experiment.'

In Landau's account, restrictions on the distribution of PD are reduced to structural constraints on movement. Under his approach, then, there must something special about the syntax projected by subject matter verbs that accounts for the contrasts in (36) and (37). The syntactic distinction that Landau draws between agentive and non-agentive perception verbs (see Section 4.3) cannot extend to the present case: subject matter verbs are always agentive, and as such should be compatible with PD. To resolve this inconsistency, he argues that the position that normally serves as the landing site for the moved possessor is occupied by a silent cognate object in subject matter verbs, but offers little evidence for this claim, making this explanation potentially circular.

If affectedness is allowed to play an explanatory role, however, there is no need for such circularity. Subject matter verbs are low-transitivity, and are therefore not expected to be particularly felicitous with PD. In fact, given the right context, a highly salient possessor and a body part possessum, subject matter verbs seem to be acceptable:

- (38) *kol ha-zman medabrim li al ha-cicim*. (Web)
 all the-time they.speak to.me about the-breasts
 'People keep talking about my breasts.'

Another verb class that Landau discusses is stative location verbs, such as *contain*. He cites the following example as ungrammatical (p. 30):

- (39) **šney xadarim hexilu le-Rina et ha-rehitim*.
 two rooms contained to-Rina ACC the-furniture.
 'Two rooms contained Rina's furniture.'

In actual usage data, however, this verb is in fact attested in PD:

- (40) *hi xašva le-acma še-ha-pas ha-ze be-ezor*
 she thought to-herself that-the-strip the-this in-region
ha-xaze holex lehaxil la et ha-cicim. (Web)
 the-chest going to.contain to.her ACC the-breasts
 'She thought that that strip (of cloth) around her chest was going to hold her breasts.'

Examples like (38) and (40) are uncommon, but still grammatical. A categorical structural explanation is therefore unlikely to provide an adequate account of these facts. The relative rareness of subject matter verbs in PD is more likely to be due to the fact that talking events do not usually affect the person being talked about. Likewise, stative verbs like *contain* are low transitivity verbs, and are expected to only occur in PD when other factors in the sentence are particularly conducive to a high affectedness conceptualization of the event. It is therefore not a coincidence that (38) and (40) both refer to intimate body part possessums and have highly salient pronominal possessors.

4.5 Summary

This section has shown that the distribution of PD is influenced by three aspects of Hopper and Thompson's (1980) notion of transitivity: the number of participants in the event, the degree of eventivity and the volitionality of the action. In general, it was shown that the more transitive the event, the more likely the speaker is to use PD. König and Haspelmath's (1998) typological survey proposes a situation hierarchy (events vs. states) and syntactic relation hierarchy (transitive vs. intransitive verbs). The notion of transitivity subsumes the two hierarchies under a more general cognitive notion, and explains additional distinctions between volitional and non-volitional events, and between individual-level and stage-level predicates. Again, we see that low-transitivity PD sentences are not ungrammatical, but rare: grammatical constraints in other languages take the form of statistical preferences in Hebrew.

5 A note on unaccusativity

Borer and Grodzinsky (1986) note that unaccusative subjects make good PD possessums, whereas unergative subjects do not:⁶

⁶ Example (41b) doesn't do justice to *šaxav* 'lay', which is unnatural without a locative adverb – compare the English gloss. The sentence improves within an appropriate context (though, as a reviewer points out, the dog could very easily be an unfamiliar dog, as long as Rina is somehow affected):

(ii) *ha-kelev šaxav le-dina al ha-kvish ve-lo hiskim lazuz.* (PD)
 the-dog lay to-Dina on the-road and-not agreed to.move
 'Dina's dog lay on the road and refused to move.'

The reviewer further points out that example (42), which is supposed to be grammatical, sounds almost as unnatural as (40b) out of context.

- (41) a. *ha-maftexot naflu li.* (PD)
 the-keys fell to.me
 ‘My keys fell.’
 b. **ha-kelev šaxav le-dina.* (PD)
 the-dog lay to-Dina
 ‘Dina’s dog lay.’

According to Borer and Grodzinsky (1986), sentence (41b) is perfectly grammatical when expressed using OP:

- (42) *ha-kelev šel Dina šaxav.* (OP)
 the-dog of Dina lay
 ‘Dina’s dog lay.’

König and Haspelmath (1998) too note the distinction between unergative and unaccusative verbs as an independent fact about the cross-linguistic behavior of PD constructions (as part of their “syntactic relations hierarchy”). Based on these observations, the compatibility of a verb with PD has often been used as an unaccusativity diagnostic for Hebrew: if the subject of intransitive verb is compatible with PD, then the verb must be unaccusative. Indeed, there is no question that unaccusative subjects are better *on average* than unergative subjects as PD possessums. The data presented in this paper, however, raise the possibility that this may be an epiphenomenon of the tight relation between the unaccusative-unergative distinction and two of the correlates of possessor affectedness outlined in previous sections: the preference of PD for transitive events, and the infelicity of human possessums (see Gafter 2010 for a related proposal).

First, prototypical PD sentences in Hebrew and other European languages involve an event, which happened to the possessum (see Sections 4.1 and 4.2). Similarly, unaccusative verbs generally involve an event that was not initiated by their subject (Levin and Rappaport Hovav 1995): usually a change of state or location, and occasionally an appearance or disappearance. This makes them highly appropriate for PD. Unergative verbs, on the other hand, do not fit this semantic characterization very well. Some of them are stative, as in Borer and Grodzinsky’s example (41b) above. As pointed out in Section 4.2, stative verbs are disfavored in PD, regardless of the syntactic function of the possessum. Most other unergative verbs are volitional and agentive, in which case the possessum is the source rather than the target of the action, which again violates the transitivity condition. Even within the class of agentive unergative verbs, it appears that the less agentive the verb is, the better the sentence becomes:

- (43) a. *?ha-yalda lomedet lahem kol ha-zman.*
 the-child studies to.them all the-time
 ‘Their child studies all the time.’
- b. *ha-yalda mišta'elet li po kvar eser dakot.* (Web)
 the-child coughs to.me here already ten minutes
 ‘My child has been coughing here for ten minutes already.’

Another potential reason for the difference in acceptability with PD between unaccusative and unergative subjects is the fact that the subjects of unergative verbs are overwhelmingly animate. As mentioned in Section 3.1, animate possessums are highly dispreferred in PD, regardless of their syntactic function; compare (44a) to (44b):⁷

- (44) a. *naflu li ha-maftexot la-pir šel ha-ma'alit.* (Web)
 fell to.me the-keys to.the-shaft of the-elevator
 ‘My keys fell into the elevator shaft.’
- b. *??nafla li ha-menahet la-pir šel ha-ma'alit.*
 fell to.me the-boss to.the-shaft of the-elevator
 ‘My boss fell into the elevator shaft.’

Crucially, there is a small class of unergative verbs that do take inanimate subjects, the so-called Theme unergatives (Levin and Rappaport Hovav 1995; Reinhart 2002). These are typically verbs of emission (Levin 1993) – for example, *glow*, *buzz* and *reek*. As predicted by the semantic factors presented in this paper, Theme unergative verbs are in fact compatible with PD (this observation is due to Tal Siloni):

- (45) *kše-siparti le-avi nacecu lo ha-einaim.* (Web)
 when-I.told to-Avi sparkled to.him the-eyes
 ‘When I told Avi, his eyes sparkled.’

⁷ A reviewer points out that the following sentence is grammatical and attested in Google searches:

- (iii) *ha-yalda nafla li me-ha-šida.* (Web)
 the-child fell to.me from-the-dresser
 ‘My child fell down from the dresser.’

The comparison of this example with (43b) above is complicated by two factors: the word order variation, which may interact both with animacy and with PD (Melnik 2002); and the fact that (ii) involves a high relational noun (*child*), which is better in PD than less clearly relational nouns such as *boss* (Section 3.1). Clearly, a lot more work needs to be done to disentangle the effect of all of these factors.

Notably, this is not limited to body part possessums:

- (46) a. *rabotai, re'idat adama ba-rexovot, ha-bait ro'ed*
gentlemen, earthquake in.the-streets, the-house is.trembling
li. (Web)
to.me.
‘Gentleman, there is an earthquake in the streets, my house is trembling.’
- b. *axrei eize 20 dakot sixa mecalcel li*
after about 20 minutes:CS conversation rings to.me
ha-telefon. (Web)
the-phone
‘20 minutes into our conversation my phone rings.’

To sum up, while unaccusative verbs are indeed much more common than unergative verbs in PD sentences, this again appears to be a statistical tendency rather than an inviolable constraint, and may be a reflection of more general principles. First, most unergative subjects are animate, which disqualifies them as PD possessums. Second, events described by unergative verbs are usually non-agentive and low-transitivity, which makes them less likely to be used with PD. This alternative explanation calls into question accounts that derive the difference in behavior between unaccusative and unergative verbs from their different structural properties (Borer and Grodzinsky 1986; Landau 1999).

6 Conclusion

Previous work has characterized the distribution of PD in terms of the types of verbs it can occur with. These verb class restrictions have been taken to stem from differences in the syntactic structures associated with those verbs. Borer and Grodzinsky (1986) argue that only verbs that have an underlying internal argument – transitive, passive and unaccusative verbs – are compatible with PD. Landau’s (1999) possessor raising account extends this approach considerably. According to Landau, non-agentive perception verbs (*see*) and subject matter verbs (*discuss*) are incompatible with PD because the syntactic structures they project lack a position that can host the moved possessor. While these authors acknowledge that PD conveys an “implication” of affectedness (Landau 1999: 3), this fact does not play an explanatory role in their account of the distribution of the construction.

This paper has shown that the affectedness condition on PD is motivated independently of the particular verb used with the construction. It is unlikely

that the quantitative patterns described in this paper can be reduced to differences in syntactic structure. For instance, it is hard to see how the statistical preference for body part possessums could be encoded as a restriction on landing sites for syntactic movement. Affectedness must therefore constitute an inherent part of the meaning of the construction. Accounts that build affectedness into the grammar of the construction, such as Pylkkänen (2002), are therefore both more parsimonious and more empirically adequate than accounts based on verb classes.

More broadly, the present work illustrates how cross-linguistic comparative studies can inform the study of language-internal variation. The variation between the Hebrew possessive dative construction and the ordinary possessive construction has been linked to the notion of possessor affectedness. This notion is not clearly defined, however, making it difficult to determine for a given state of affairs whether speakers are likely to use PD to describe it. We turned to similar constructions in related languages in search of clearly identifiable factors associated with affectedness. The key assumption was that the shared cognitive notion of affectedness may be grammatically encoded in a different way in each of those languages.

This paper showed that Hebrew imposes few grammatical constraints on PD, aside from the broad affectedness condition. Related languages, on the other hand, enforce the affectedness condition by regulating specific properties of the event. German, for instance, requires the PD possessor to be animate, and French requires the possessum to be a body part. Under the assumption that these grammatical restrictions reflect the conventionalization of various aspects of the cognitive notion of affectedness, we expected these restrictions to surface as statistical tendencies in Hebrew. Indeed, the Hebrew PD turned out to be much more common with animate than inanimate possessors, and favored body part possessums more than any other kind of possessum, mirroring the German and French restrictions.

The Hebrew PD joins other documented examples of “hard” grammatical constraints in one language mirroring “soft” statistical tendencies in another (Givón 1979; Aissen 1999; Bresnan et al. 2001). This similarity between grammatical facts and statistical tendencies underscores the theoretical assumption that grammatical conventions evolve out of prevalent discourse patterns (Ariel 2008; Bybee and Hopper 2001). The fact that the same constraint is categorical in one language and stochastic in another is difficult to capture through formal restrictions on syntactic movement, and motivates formal architectures that explicitly allow constraints to be either categorical or statistical, such as Stochastic Optimality Theory (Boersma and Hayes 2001; Boersma 1998).

Acknowledgments: I am particularly grateful to Mira Ariel for general guidance and extensive feedback on previous versions of this work. I thank Julia Horvath and Tal Siloni for helpful discussion. This work also benefited from comments from Ibtisam Ammouri, Mark Baltin, Larry Horn, Itamar Kastner, Idan Landau, Hillel Taub-Tabib and two anonymous reviewers, as well as audiences at the Linguistics Colloquium at Tel-Aviv University in 2009 and the Workshop on Variation and Change in Argument Realization in Naples in 2010. Most of the research presented in this paper was conducted as part of the author's Master's thesis in Linguistics in Tel Aviv University (Linzen 2009).

References

- Adler, Meni & Michael Elhadad. 2006. An unsupervised morpheme-based HMM for Hebrew morphological disambiguation. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*. acl.ldc.upenn.edu/P/P06 (accessed 28 Feb 2014).
- Aissen, Judith. 1999. Markedness and subject choice in Optimality Theory. *Natural Language and Linguistic Theory* 17(4). 673–711.
- Ariel, Mira. 1990. *Accessing noun phrase antecedents*. New York: Routledge.
- Ariel, Mira. 2008. *Pragmatics and grammar*. Cambridge: Cambridge University Press.
- Ariel, Mira, Elitzur Dattner, John Du Bois & Tal Linzen. submitted. Pronominal datives: The royal road to argument status.
- Berman, Ruth. 1982. Dative marking of the affectee role: Data from Modern Hebrew. *Hebrew Annual Review* 6. 35–59.
- Boersma, Paul. 1998. *Function phonology: Formalizing the interaction between articulatory and perceptual drives*. Amsterdam: University of Amsterdam dissertation.
- Boersma, Paul & Bruce Hayes. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32(1). 45–87.
- Borer, Hagit. 1999. Deconstructing the construct. In Kyle Johnson and Ian Roberts (eds.), *Beyond principles and parameters*, 43–89. Dordrecht: Kluwer.
- Borer, Hagit & Yosef Grodzinsky. 1986. Syntactic cliticization and lexical cliticization: The case of Hebrew dative clitics. In Hagit Borer (ed.), *Syntax and semantics* 19, 175–217. New York: Academic Press.
- Bresnan, Joan, Shipra Dingare & Christopher D. Manning. 2001. Soft constraints mirror hard constraints: voice and person in English and Lummi. In Miriam Butt & Tracy Holloway King (eds.) *Proceedings of the LFG '01 conference*, 13–32. Stanford, CA: CSLI Publications.
- Bybee, Joan & Paul Hopper (eds.). 2001. *Frequency and the emergence of linguistic structure*. Amsterdam & Philadelphia: John Benjamins.
- Dąbrowska, Ewa. 1998. How metaphor affects grammatical coding: The Saxon genitive in computer manuals. *English Language and Linguistics* 2(1). 121–127.
- Diffloth, Gérard. 1974. Body moves in Semai and French. In Michael La Galy, Robert Fox & Anthony Bruck (eds.), *Papers from the tenth regional meeting, Chicago Linguistic Society, April 19–21 1974*, 128–138. Chicago: Chicago Linguistic Society.

- Fried, Mirjam. 1999. From interest to ownership: A constructional view of external possessors. In Doris Payne & Immanuel Barshi (eds.), *External possession*, 473–504. Amsterdam & Philadelphia: John Benjamins.
- Gafter, Roey. 2010. A Prominence Account of the Hebrew Possessive Dative Construction. ms., Stanford University. http://linguistics.stanford.edu/documents//RoeyGafterQP1_final.pdf (accessed 28 Feb 2014)
- Givón, Talmy. 1979. *On understanding grammar*. New York: Academic Press.
- Halevy, Rivka. 2013. The dative in Modern Hebrew. In Geoffrey Khan (ed.), *Encyclopedia of Hebrew language and linguistics*. Leiden: Brill.
<http://pluto.mscc.huji.ac.il/~halevyne/dativeinMHHalevy.pdf> (accessed 28 Feb 2014).
- Haspelmath, Martin. 1999. External possession in a European areal perspective. Doris Payne & Immanuel Barshi (eds.), *External possession*, 109–136. Amsterdam & Philadelphia: John Benjamins.
- Havers, Wilhelm. 1911. *Untersuchungen zur Kasussyntax der indoeuropäischen Sprachen*. Straßburg: Trübner.
- Hopper, Paul & Sandra Thompson. 1980. Transitivity in grammar and discourse. *Language* 56(2). 251–299.
- Hopper, Paul & Elizabeth Traugott. 2003. *Grammaticalization*. Cambridge: Cambridge University Press.
- Horvath, Julia & Tal Siloni. 2008. Active lexicon: adjectival and verbal passives. In Sharon Armon-Lotem, Gabi Danon & Susan D. Rothstein (eds.), *Current issues in generative Hebrew linguistics*, 105–134. Amsterdam & Philadelphia: John Benjamins.
- Jelinek, Eloise & Richard A. Demers. 1983. The agent hierarchy and voice in some coast Salish languages. *International Journal of American Linguistics* 49(2). 167–185.
- König, Ekkehard & Martin Haspelmath. 1998. Les constructions à possesseur externe dans les langues d'Europe. In Jack Feuillet (ed.), *Actance et valence dans les langues de l'Europe*, 525–606. Berlin & New York: Mouton de Gruyter.
- Kuno, Susumo. 1987. *Functional syntax: anaphora, discourse and empathy*. Chicago: University of Chicago Press.
- Landau, Idan. 1999. Possessor raising and the structure of VP. *Lingua* 107(1). 1–37.
- Levin, Beth. 1993. *English verb classes and alternations*. Cambridge, MA: MIT Press.
- Levin, Beth & Malka Rappaport Hovav. 1995. *Unaccusativity: At the syntax-semantics interface*. Cambridge, MA: MIT Press.
- Linzen, Tal. 2009. *Hebrew Possessive Datives: The effects of affectedness*. Tel Aviv: Tel Aviv University MA thesis.
- Linzen, Tal. 2010. Hebrew statistical linguistics using a morphologically analyzed blog corpus. Paper presented at the Israeli Seminar on Computational Linguistics 2010, Tel Aviv, Israel.
- Melnik, Nurit. 2002. *Verb-initial constructions in Modern Hebrew*. Berkeley, CA: University of California at Berkeley dissertation.
- Meltzer-Asscher, Aya. 2011. Adjectival passives in Hebrew: evidence for parallelism between the adjectival and verbal systems. *Natural Language and Linguistic Theory* 29(3). 1–41.
- Neumann, Dorothea. 1996. The dative and the grammar of body parts in German. In Hilary Chappell & William McGregor (eds.), *The grammar of inalienability: A typological perspective on body part terms and the part-whole relation*, 745–782. Berlin & New York: Mouton de Gruyter.

- Nichols, Johanna. 1988. On alienable and inalienable possession. In William Shipley (ed.), *In honor of Mary Haas: From the Haas festival conference on native American linguistics*, 475–521. Berlin & New York: Mouton de Gruyter.
- Payne, Doris L. & Immanuel Barshi. 1999. External possession: What, where, how, and why. In Doris Payne & Immanuel Barshi (eds.), *External possession*, 3–29. Amsterdam & Philadelphia: John Benjamins.
- Pylkkänen, Liina. 2002. *Introducing arguments*. Cambridge, MA: MIT dissertation.
- Reinhart, Tanya. 2002. The theta system – an overview. *Theoretical Linguistics* 28(3). 229–290.
- Seiler, Hansjakob. 1983. *Possession as an operational dimension of language*. Tübingen: Gunter Narr.
- Shibatani, Masayoshi. 1994. An integrational approach to possessor raising, ethical datives and adversative passives. *BLS* 20. 461–485.
- Siloni, Tal. 2001. Construct states at the PF interface. *Linguistic Variation Yearbook* 1(1). 229–266.
- Silverstein, Michael. 1976. Hierarchy of features and ergativity. In R. M. W. Dixon (ed.), *Grammatical categories in Australian languages*, 112–171. Canberra: Australian Institute of Aboriginal Studies.
- Wierzbicka, Anna. 1988. *The semantics of grammar*. Amsterdam & Philadelphia: John Benjamins.

Appendix. Methodology

The appendix describes the data used in the rest of the paper (Section A.1), and discusses the challenges of distinguishing the Hebrew possessive dative from related syntactic constructions (Section A.2).

A.1 Sources of data

This paper presents three types of data: corpus frequency counts, example sentences found in Google searches, and invented examples with grammaticality judgments. Most of the invented examples are drawn from published papers; examples without references are original judgments.

The quantitative data are based on the Israblog Corpus (Linzen 2010), a 165-million word corpus that consists of texts in various registers of Modern Hebrew, written by thousands of authors of different ages between 2005 and 2008. The corpus was extracted in September 2008 from www.israblog.co.il, a blog hosting site. It was morphologically analyzed and disambiguated (part-of-speech tagged) using the BGUTagger morphological analyzer (Adler and Elhadad 2006).

We searched the first 500,000 sentences of the corpus for possessive sentences with direct object possessums, using the following frames:

(47) a. POSSESSIVE DATIVE

V *le-X* *et* Y

V to-X ACC Y

b. ORDINARY POSSESSION

V *et* Y *šel* X

V ACC Y of X

We automatically filtered out sentences with 50 common verbs which select the dative preposition *le-*, such as *natan* ‘give’, since a dative-marked noun phrase in these sentences is almost certainly a goal argument of the verb rather than a PD. Additionally, we excluded a handful of sentences in which a single possessive relation was expressed using both of the possessive constructions simultaneously. For example:

(48) *mi axal li et ha-dag šeli?* (Web)

who ate to.me ACC the-fish my

‘Who ate my fish?’

Finally, we manually eliminated other irrelevant search results. The final sample consisted of 1124 (20%) hand-verified PD sentences and 4547 (80%) OP sentences.

Hebrew has a third construction that is occasionally used to express possession, called the construct state (Siloni 2001; Borer 1999). This construction is largely limited to formal registers of the language. In the construct state construction, the possessum and possessor are simply juxtaposed:

(49) *efšar le-cayer et [yad ha-yeled] al niyar lavan.* (Web)

possible to-draw ACC hand the-child on paper white

‘One can draw the child’s hand on white paper.’

In some cases, the juxtaposition is accompanied by a morphological change in the possessum (the first noun). Since this construction has very limited productivity in colloquial Hebrew, we have chosen to limit our attention to the two productive constructions illustrated in (47).

A.2 Distinguishing PD from other dative constructions

Hebrew has a family of related dative constructions, which are often not easy to tell apart. There is significant variability in how researchers carve them apart. For

example, Halevy (2013) argues that the dative-marked noun phrase in (50a) is a possessive dative, but in (50b) it is a “dative of interest”:

- (50) a. *ro'im la et ha-taxtonim.*
 see:IMPERSONAL to.her ACC the-underwear
 ‘One can see her underwear. (= Her underwear is showing)’
 b. *hu haras le-ruti et ha-mexonit.*
 he wrecked to-Ruti ACC the-car
 ‘He wrecked Ruti’s car.’

Halevy does not propose a principled way to draw the line between these two constructions, so it is not clear whether they need to be distinguished to begin with.

Another competing construction is the ethical dative (Berman 1982; Borer and Grodzinsky 1986; Halevy 2013). The noun phrase marked as an ethical dative is “an onlooker perceived as being intensely affected by, or as having a strong emotional stake in, the state of affairs described in the sentence” (Halevy 2013). For example (Halevy’s gloss):

- (51) *be-emca ha-štur hu nixnas li.*
 in-middle the-class he walked.in to.me
 ‘In the middle of the class he walked in [in his impudence].’

This definition is of limited use: as pointed out by several authors, the possessor in a possessive dative sentence is also perceived as being affected by the state of affairs described in the sentence (Berman 1982). Borer and Grodzinsky (1986) propose a distributional diagnostic for distinguishing between possessive and ethical datives. According to them, the dative-marked noun phrase in an ethical dative construction must be a pronoun, whereas in the possessive dative it can also be a lexical noun phrase:

- (52) *rak al tatxil le-hitrakex li / *le-mixal.*
 only do.not start to-soften.up to.me/*to-Mixal
 ‘Don’t you get soft on me / *Mixal now!’

In the interest of excluding sentences that are potentially ambiguous between PD and ethical dative, Borer and Grodzinsky restrict their discussion of PD to sentences in which the dative-marked noun phrase is not a pronoun. The same strategy is adopted by Landau (1999). However, following this approach would severely impact the quantitative aspects of this paper, since the vast majority of

naturally occurring PD sentences have a pronominal possessor. As discussed in Section 2, this fact follows from the affectedness meaning component encoded in PD: salient possessors, typically referred to with pronouns, are more likely to be perceived as affected by an event. Restricting our queries to full noun phrases would reduce the amount of data we have by a factor of 20. In addition, this would make the Google searches impossible: since the dative preposition *le-* is fused with the noun following it, forming a single orthographic word, there is no way to search for a verb followed by any dative-marked noun.

In practice, the risk of misclassifying an ethical dative as a PD seems to be minimal. An informal survey of dative sentences from the corpus revealed that ethical datives such as (51) are much more rare in actual usage than possessive datives. Even if there were a way to unambiguously identify ethical datives and exclude them from further analysis, it would probably have a negligible effect on the statistical results.

Furthermore, note that in contrast with examples (51) and (52), the frames that are in the focus of our investigation, shown in (47), always contain a potential possessive relation between the dative marked argument and the verb's object. There is a strong bias for interpreting the dative marked argument as the possessor of the direct object:

- (53) *mixal šavra li et ha-xalon.*
 Mixal broke to.me ACC the-window

In example (53) the PD reading “Mixal broke my window” is overwhelmingly preferred to a putative ethical dative reading “Mixal broke the window (not necessarily mine) in her impudence”. This makes it unlikely for speakers to choose to use an ethical dative in such frames.

Finally, it is not clear whether the lexical noun phrase condition on ethical datives is even empirically supported. For example, the dative-marked noun phrase *le-rani* ‘to Rani (proper name)’ in sentence (54) seems to fall under the ethical dative category:

- (54) *kulam mitxatnim le-rani kol ha-zman ve-rak*
 everyone gets.married to-Rani all the-time and-only
hu niš'ar levad.
 he stays alone
 ‘Everyone Rani knows goes and gets married on him but he’s still single.’

It may be the case that ethical datives are not required to be pronominal, but show a strong statistical tendency to be pronominal, much in the same way that

possessive datives do. The preference for pronominal noun phrases may be stronger in ethical datives than in the PD, possibly since it is more difficult to figure out how the ethical dative is affected without the possessive link tying the possessor to the event, so a more salient entity is needed. At any rate, it is not obvious that this is a grammatical constraint as argued by Borer and Grodzinsky (1986). Since the ethical dative is not the focus of this paper, we do not pursue this matter any further; in any case, it seems safe to conclude that in practice the various competing dative constructions do not cast doubt on the results of this paper.

In summary, we remain agnostic about the distinctions between the various dative constructions. We use an operational definition of PD: a PD is any case in which a dative marked argument of the verb is interpreted as the possessor of another object of the verb. Even if there is reason to distinguish the different dative constructions, the chance of accidentally misclassifying a sentence as an instance of PD is minimal in practice, and the benefits of using an expansive definition of PD outweigh the risks.