Bradley Rentz[1]

Topological Relations in Pohnpeian

[1] Department of Linguistics, University of Hawai'i at Mānoa, Honolulu, HI 96822-2217, USA, E-mail: rentzb@hawaii.edu

Abstract:
This article explores a more nuanced understanding of topological relations in the Pohnpeian language (Austronesian). The BowPed Toolkit (Bowerman and Pederson 1992. Topological relations picture series. In *Space stimuli kit 1.2: November 1992, 51*, Nijmegen: Max Planck Institute for Psycholinguistics. http://fieldmanuals.mpi.nl/volumes/1992/bowped/.) is employed as an elicitation tool with five Pohnpeian speakers. Evolutionary classification tree modeling is used as a discovery tool to find patterns in the data. The results show that the two prepositions in Pohnpeian, *nan* and *ni*, should be redefined in terms of topological relations as 'containment' and 'attachment' respectively. Likewise the meaning of some prepositional nouns are further revised.

Keywords: Pohnpeian language, topological relations, Micronesia, evolutionary classification tree modeling

DOI: 10.1515/lingvan-2016-0092

Received: November 8, 2016; **Accepted:** February 4, 2017

1 Introduction

Pohnpeian (ISO 639-3: pon) – an Austronesian language of the Oceanic subgroup spoken in the Federated States of Micronesia – is an understudied language in terms of topological relations: the linguistic coding of how two objects are related in space. The limited research on Pohnpeian spatial relations primarily focuses on the morphosyntax of relational words and does not explore the complexities of their meanings. In this paper, I present the results from the BowPed Toolkit (Bowerman and Pederson 1992) to elicit a more nuanced understanding of the topological relations in Pohnpeian, since it has been used successfully for other languages (e.g., Basque, Dutch, Ewe, Lao, Lavukaleve, Tiriyó, Trumai, Yélî Dnye, and Yukatek (Levinson and Meira 2003); Dene sųłiné, English, German, Norwegian, and Upper Necaxa Totonac (Thiering 2007); Likpe (Ameka 2007); Ronnga (Aryawibawa 2008); and Dutch (Gentner and Bowerman 2010)). In addition to presenting my qualitative results, I employ a novel method of using evolutionary classification trees to create a quantitative model of the data, which greatly aids the understanding of these complex grammatical features.

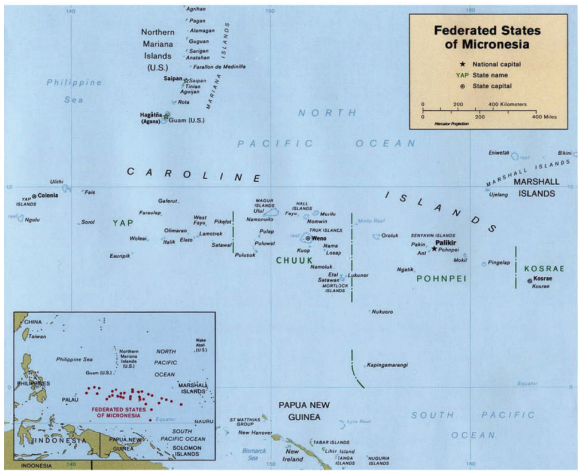
Pohnpeian is spoken primarily on the island of Pohnpei (Figure 1(a)) in the Federated States of Micronesia (FSM) (Figure 1(b)). There are about 34,000 speakers in Pohnpei (FSM Office of Statistics, Budget and Economic Management, Overseas Development Assistance, and Compact Management 2010) and approximately 12,000 in the USA based on population estimates by Hezel and Levin (2012).

The limited research on Pohnpeian topological relations has shown that it has a complex system of directional morphology including prepositions, prepositional nouns and verbs, and verbal and nominal suffixes (Rehg and Sohl 1981). In this paper I focus on the prepositions and a subset of prepositional nouns. I outline those used in this study in Section 2.2 and Section 2.3. Previous studies describe the prepositions *nan* and *ni* as 'in' and 'at, to' respectively, but this study shows that the labels 'containment' and 'attachment' better characterize the functions of the prepositions. Likewise, more nuanced meanings of some of the prepositional nouns are further demonstrated in Section 4 and Section 5.

In Section 2, I provide an overview of previous descriptions of Pohnpeian locational terms, followed by the project's method in Section 3. In Section 4, I give the project's results, discussion in Section 5, and conclusions in Section 6.

Bradley Rentz is the corresponding author.

(a)



(b)

Figure 1 Maps of (a) the Island of Pohnpei with home sections of participants marked and (b) the Federated States of Micronesia (Central Intelligence Agency 1999).

2 Background

2.1 Topological relations

In this paper, topological relations refer to "locational relations between objects that specify space in general...[and] are considered impermeable or perspective-neutral locative relations between physical objects" (Thiering 2007: 1). Furthermore, these "geometrical properties...remain constant under transformation or deformation, and so are preserved under the loss of metric angle and distance" (Levinson 2003: 71). Topological relations contrast with 'frames of reference' which are angular or directional relationships that require some sort of coordinate system to express the relationship (Levinson 1996; Levinson and Wilkins 2006), where the number of reported possible coordinate systems varies from 3 (Levinson 1996; Levinson and Wilkins 2006) to 8 (Jackendoff 1996). However, the distinction between topological relations and frames of reference is limited in that although "many topological relators express no angular or coordinate information...others do involve the vertical absolute dimension and often intrinsic [frame of reference] features...[such as, for example,] *under* (in *The dust under the rug*) compounds topological, intrinsic (under-surface, bottom) and absolute (vertical) information" (Levinson and Meira 2003: 72). The definition of topological relations is further nuanced by Vandeloise (1991), Thiering (2007) who argue for a functional approach where topological relations are often contextualized so that they cannot be adequately described by geometry alone. The present work takes a more descriptive approach to the semantics of Pohnpeian's prepositions and prepositional nouns.

One way to describe topological relationships is that they express the relationship of an object, the 'figure', that is located in space by reference to another object, the 'ground' (Talmy 1983). Alternatively, the terms 'trajector' (figure) and 'landmark' (ground) are used where the trajector is the "entity whose (trans) location is of relevance" and the landmark is the "reference entity in relation to which the location or the trajectory of motion of the trajector is specified" (Zlatev 2007: 327). Based on ample previous research, topological relations and

spatial concepts are not universal but rather constructed through discourse and can vary even within a single language community (Levinson 1996; Levinson 2003; Levinson and Meira 2003; Thiering 2007; Garza 2011; *i.a.*). Topological relations are commonly coded by adpositions, case, predicates, spatial nominals, and verbs, though other ways have been attested as well (Levinson 2003).

2.2 Prepositions

According to Rehg and Sohl (1981), Pohnpeian has two prepositions: *ni*[1] and *nan*.[2] *Ni* has an optional allomorph of *nin* when it occurs before words starting with a coronal consonant (/t/, /t̪/, /s/, /l/, /n/, or /r/).

Rehg and Sohl (1979) define *ni* as '[a]t, to' with the example sentence, *E kohla ni sidohwaho* 'He went to the store' (p. 68). This definition is slightly expanded upon in Rehg and Sohl (1981) with three more example sentences: (1) *E wahla* ***ni*** *nahso* 'He carried it to the feasthouse', (2) *E mihmi* ***ni*** *imweio* 'He is staying at my house', and (3) *E kohsang* ***ni*** *ihmw sarawio* 'He came from the church' (p. 291).

With similar brevity, *nan* is defined as '[i]n' with the example, *Kilelo mi* ***nan*** *kapango* 'That picture is in that suitcase' (Rehg and Sohl 1979: 66). *Nan* is later explained to be "used to establish a location within a space defined either in terms of two or three dimension. Consequently, 'in' is typically the most appropriate translation of *nan*...*Nan* corresponds to the English preposition 'on' only when a location within a two dimensional space is being referred to" (Rehg and Sohl 1981: 292). Two other example sentences were also provided:

(1) *E ntingihdi* ***nan*** *tehn doaropweho* 'He wrote it on that piece of paper' and

(2) *E kohla* ***nan*** *skohso* 'He went to the airport' (p. 292).

Both *nan* and *ni* can have temporal meanings as well. *Nan* is used for large periods of time or non-specific events like in *Irail pahn pwurodo nan wihk kohkohdo* 'They will return next week' or *E kin angin nan Tisempe* 'It is windy in December' (Rehg and Sohl 1981: 298). *Ni* is used for specific times as in *Soulik lemwida ni eh kilangada enio* 'Soulik got frightened at the time he saw the ghost' (Rehg and Sohl 1981: 298).

Like most fieldwork that uses English translations, these definitions do not describe the complexity of meaning (in terms of topological relations) of the prepositions, especially since their primary goal was not an in-depth study of them. In contrast Levinson (2003), Levinson and Meira (2003) have shown that specifically designed toolkits, like Bowerman and Pederson (1992), elicit localized categories of topological relations in a way that is less biased by topological categories of the researcher's L1, especially when compared with traditional elicitation techniques.

2.3 Prepositional nouns

The largest set of locational words consist of what Rehg and Sohl (1981) call 'prepositional nouns.' These nouns always occur in direct possessive constructions. The prepositional nouns found in this paper are given in the base form (3SG) and with the construct suffix in Table 1. Examples of prepositional nouns are given in (1–3).

Table 1 Prepositional nouns in this study (Rehg and Sohl 1981: 285).

Base form	Construct form	Gloss
limwah	limwahn	'next to him/her/it'
pah	pahn	'below him/her/it'
powe	pohn	'above him/her/it'
loale	loalen	'inside it (or of inward emotions for people)'
liki	likin	'outside it (or of an outward show of emotions for people)'

(1)

Lahp-o	*mwo~mwohd*	*liki-n*	*nahs-o*
guy-DIST.SG	sit~IPFV	outside-CONST	feast.house-DIST.SG

'That guy is sitting outside the feast house.'

(2)

Kep-o	*mih~mi*	*poh-n*	*tehpel-o*
cup-DIST.SG	STATV~IPFV	above-CONST	table-DIST.SG

'The cup is on the table.'

(3)

Mpwei-o	*mih~mi*	*pah-n*	*tehpel-o*
Ball-DIST.SG	STATV~IPFV	under-CONST	table-DIST.SG

'The ball is under the table.'

3 Method

3.1 Procedure

The BowPed Toolkit (Bowerman and Pederson 1992) was the stimulus used in this project. It consists of 71 decontextualized images that were placed into a pdf with each image on a separate page. Written above each image was *Iawasa* ___*?* ('Where is ___?') for singular figures or *Iahnge* ___*?* ('Where are ___?') for plural figures, where the blank was filled in with the intended figure for that image (Figure 2 shows an example slide). The toolkit is designed to elicit topological relations in a way that is not biased by the researcher's L1 and that allows the participants to use their own topological grouping patterns with the image (Thiering 2006, Thiering 2007). The toolkit is purposely decontextualized so that the largest number topological relations may arise out of the data given enough participants. The BowPed Toolkit does, however, have some flaws, such as including images that elicit frame of references but not enough for a thorough analysis. It also does not account for difference in position of the ground or figure (eg., whether an object is standing or lying on a table), which later toolkits like Hellwig and Lüpke (2001) elicit. Despite its shortcomings, the BowPed Toolkit provides an effective starting point for the study of topological relations.

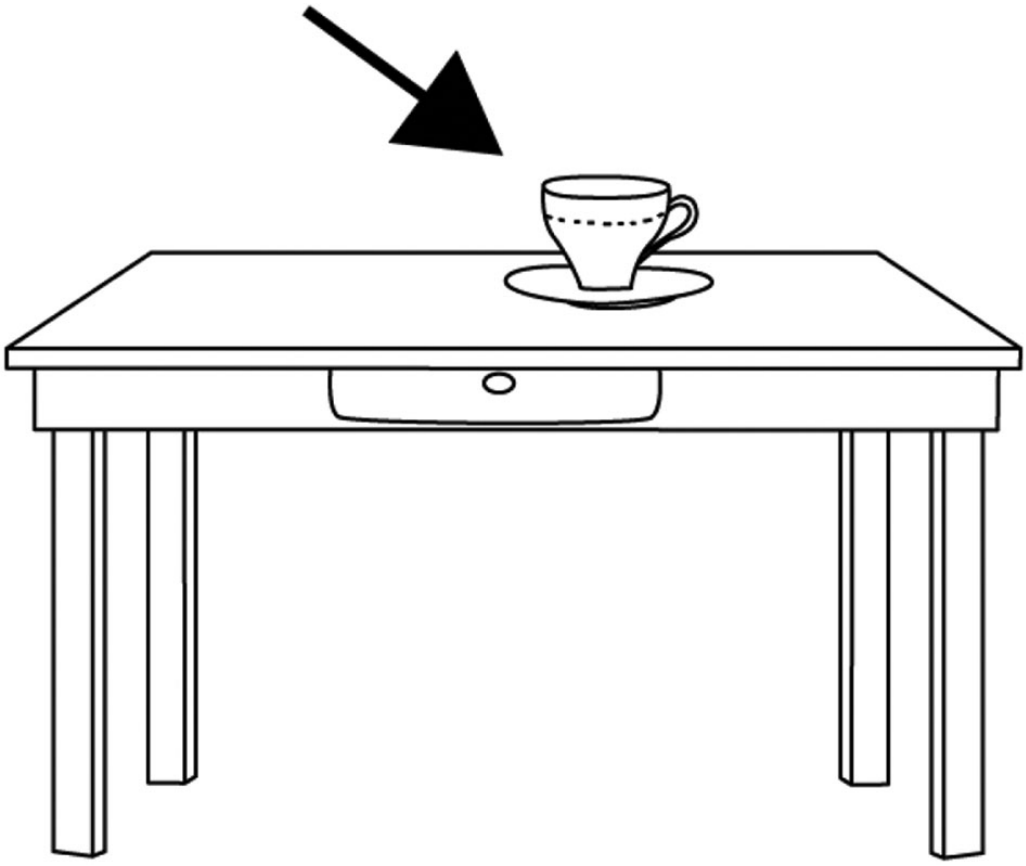


Figure 2 Example slide from the BowPed elicitation session ('Where is the cup?').

The participants were shown a question in Pohnpeian on each slide of the pdf asking them answer where the object in question is. Each participant answered the questions aloud at their own pace and advanced to the next page of questions on their own. If they were confused by the question or the picture, I helped explain what was intended. The order of the images was the same for all the participants. Each recording was saved as an uncompressed WAV file and transcribed by the author. The recording sessions lasted about 10 minutes each. The WAV files are available in the Kaipuleohone Digital Ethnographic Archive (https://scholarspace.manoa.hawaii.edu/handle/10125/33308, files BR1-015 through BR1-019).

Five L1 Pohnpeian speakers living on O'ahu, Hawai'i took part in the project. All five participants are fluent L2 speakers of English and have lived in Hawai'i for several years. Their mean age was 42 years. Information on education level was not collected. The participants varied somewhat in their original municipality in Pohnpei and most have lived in several different dialect regions of the island for extended periods and also different parts of the US before moving to O'ahu. The mobility of the participants complicated accounting for dialect variation. Table 2 lists the participants' demographics including self-identified home municipality and section on Pohnpei.[3] The home sections are mapped in Figure 1(a).

Table 2 Demographics of participants.

Participant ID	Gender	Age	Municipality	Section
M01	M	51	Kitti	Diadi
M02	M	38	Kitti	Nanmand
M03	M	27	Kitti	Nanmand
F01	F	48	Nett	Kolonia
F02	F	44	Sokehs	Palikir

3.2 Statistical modeling: Evolutionary classification trees

The data were analyzed using evolutionary classification trees with the *R* (R Core Team 2016) package *evtree* (version 1.0-0, Grubinger et al. 2014). 15 coded features for each image were used as predictor variables and the relational word produced by the participant was the dependent variable. The evolutionary tree algorithm views the entire dataset as a whole and simultaneously selects many 'parent' and 'child' splits in the data and evaluates each of them them based on a fitness function where only the splits that increase the mean quality of the population (i.e., a low misclassification error rate) are kept. The final tree is the one that has the greatest 'fitness' (see Grubinger et al. 2014 for more details). This algorithm contrasts with more traditional Classification and Regression Trees (CARTs) (Baayen 2008; Therneau and Atkinson 2015), which operate by creating nodes that maximize the 'purity' of each node according to the Gini index until they reach a node with a pre-defined minimum group size and thus tend to be overfitted. Evolutionary classification trees do not have the same biases and overfitting of CARTs and do not require pruning.[4]

Importantly, the results of the tree do not claim to be what is necessarily most salient to the speaker, but provide a statistical model to find significant patterns in the current data. The generated tree is used to show patterns in the data that might not be seen by grouping the tokens by hand.

3.2.1 Coding of images

After the data were collected, each image was coded with a series of 15 binary features based on potential salient relational features between the ground and figure. These features are not presumed to be universal in anyway and are based on perceived features that speakers may use to determine which relational word to use. Binary features were created in order to be able to model the data quantitatively. The features were chosen from the possible features that language communities have been shown to use to group topological relations (e.g., attachment, contact, containment, relative size of figure to ground, etc., (Bowerman 1996; Levinson and Meira 2003; Levinson 2003) and are used simply as a tool to describe the data. Fifteen features were used to create a surplus of distinctions to adequately distinguish all the relational words and toolkit images, while simultaneously allowing the statistical model (discussed in Section 3.2) to be able to succinctly determine which features are meaningful and which are not for the data.[5] The coding of the features was not directly influenced by the resulting data, except when the participant used a non-primary figure-ground relationship found in the

images. In such cases, the image was recoded to represent the relationship used. As a result, some images have more than one possible ground-figure relationship. Table 3 gives a description of the features used.

Table 3 Image coding features.

	Feature	Code	Description (when = 1)
1.	Contact	contact	Ground and figure touch
2.	Horizontal ground	horizontal.ground	Ground is horizontal relative to the figure
3.	3D ground	X3d.ground	Ground appears 3D to viewer
4.	Animacy of figure	animacy.fig	Animal and human figures are animate
5.	Figure is on top of ground	top.of.ground	Figure is vertically above ground regardless of contact
6.	Ground contains figure	containment	The ground contains at least most of the figure in 3D space
7.	3D figure	X3d.fig	Figure appears 3D to viewer
8.	Figure goes around ground	go.around.ground	The figure goes around the ground in either 2D or 3D space, e.g., the ring goes around the finger
9.	Ground goes around the figure	go.around.fig	The ground goes around the figure in either 2D or 3D space
10.	Vertical ground	ground.vertical	Ground is vertical relative to the figure
11.	Figure is attached to the ground	fig.attached.to.ground	The figure is connected/adhered/tied to the ground
12.	Ground is larger than figure in 2D space	ground.larger.2D.fig	The ground is significantly larger in 2D space than the figure and the figure is within that space of the ground
13.	Ground is a building/structure	ground.building	The ground is a building or structure that can contain things, e.g., cage, house, church
14.	Figure is on the side of the ground	fig.side.ground	The figure is located near the side of the ground as determined by the orientation of the ground
15.	Figure is far from the ground	fig.far.ground	The figure is a significant distance away from the ground

4 Results

299 out of a total of 355 responses from the five participants were analyzed. Responses where the participant did not refer to a meaningful ground or figure were removed as were examples where the primary locational word was a verb, such as *pidakihpene* 'to surround'; Images 15, 17, 43, 55, and 64 from the toolkit were not included in the statistical analysis for that reason.

The evolutionary classification tree generated 8 terminal nodes and had a misclassification rate of 13.7% for the dataset.[6] This error rate is quite good given the sample size and the natural variation inherent in language data. The images from Bowerman and Pederson (1992) that occur in each of the eight nodes are depicted in Figure 4–Figure 10. Some images occurred in more than one terminal node because of multiple possible figure-ground relationships encoded in the image. The model chose only 7 of the 15 available binary features[7] to use in the tree: go.around.fig, fig.attached.to.ground, top.of.ground, go.around.ground, horizontal.ground, ground.building, and ground.larger.2D.fig.

Figure 3 gives a textual description of the nodes and the most predominate relational word in each terminal node. The number inside the square brackets [ ] indicates the node number. Detailed results for each terminal node are given in the following sections under the most predominate occurring locational word. The node number from the original tree is provided along with the defining features, the percentage of occurrence for each word, the total number of tokens in the node, and the images that occur. A graphical version of the tree is given in the Appendix. The *R*-code and the collated data are available at http://hdl.handle.net/10125/42692.

Below I present the results for each Pohnpeian term, which include the terminal node that they each occur most frequently in, example sentences, and alternate uses. I conclude the section by giving general observed patterns for alternate uses of *nan*, *ni*, and *pohn/pahn* in Section 4.6.

```
[1] root
|   [2] go.around.fig in 1
|   |   [3] fig.attached.to.ground in 0: nan (n = 52, err = 7.7%)
|   |   [4] fig.attached.to.ground in 1: ni (n = 28, err = 10.7%)
|   [5] go.around.fig in 0
|   |   [6] top.of.ground in 1: pohn (n = 60, err = 13.3%)
|   |   [7] top.of.ground in 0
|   |   |   [8] go.around.ground in 0
|   |   |   |   [9] horizontal.ground in 0
|   |   |   |   |   [10] ground.building in 1: likin (n = 13, err = 30.8%)
|   |   |   |   |   [11] ground.building in 0
|   |   |   |   |   |   [12] ground.larger.2D.fig in 0: limwah (n = 7, err = 28.6%)
|   |   |   |   |   |   [13] ground.larger.2D.fig in 1: ni (n = 48, err = 18.8%)
|   |   |   |   [14] horizontal.ground in 1: pahn (n = 29, err = 20.7%)
|   |   |   [15] go.around.ground in 1: ni (n = 62, err = 8.1%)
```

Figure 3 Evolutionary classification tree structure.

4.1 Ni

Ni (previously glossed as 'at, to') was the predominate word used in terminal nodes 4, 15, and 13 (Figure 4–Figure 6). Though only explicitly included as a distinguishing feature in node 4 (Figure 4), situations in all three nodes involve various forms of the figure and ground being attached to each other. The form of attachment includes physical attachment such as an earring piercing an earlobe, a telephone hanging on the wall, a cork in the neck of a bottle, a ribbon tied to a candle, a ring on a finger, an apple on a spike, fruit and leaves on a tree branch, rain drops on a glass window, and letters on a shirt. The attachment can also include the absence of something, such a hole in a sheet or a crack in cup. On the less attached end, it can include an object leaning against a vertical surface, like a ladder leaning against a wall. Examples of these attachments are given in (4–10).

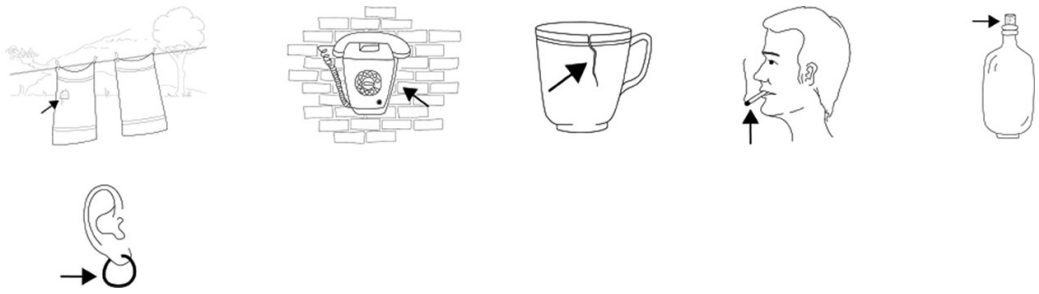


Figure 4 Features: go.around.fig = 1, fig.attached.to.ground = 1. Ni = 81%, loale = 4%, nan = 4%, pohn = 4%, n = 28 (Node 4).

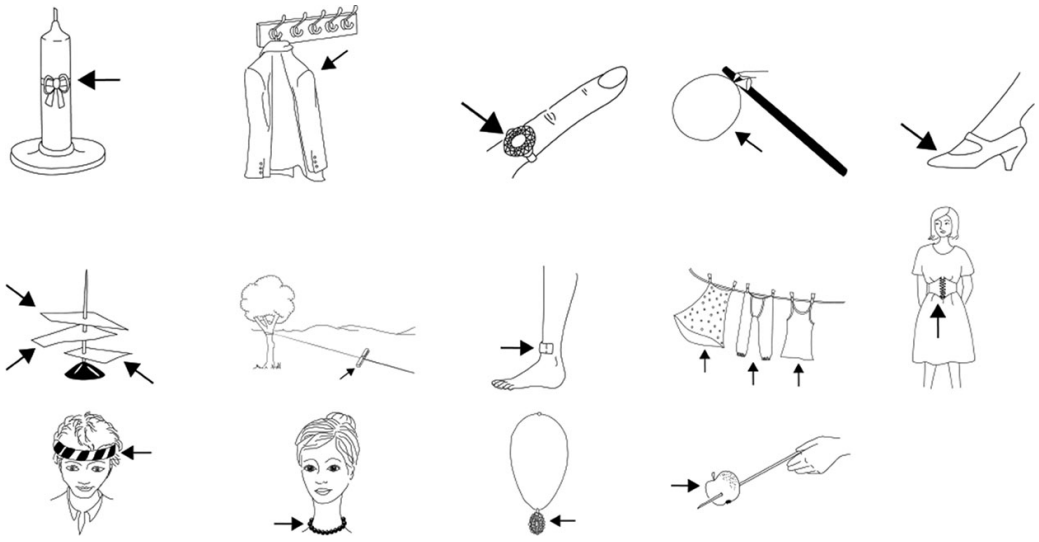


Figure 5 Features: go.around.fig = 0, top.of.ground = 0, go.around.ground = 1. Ni = 92%, nan = 5%, pohn = 3%, n = 62 (Node 15).

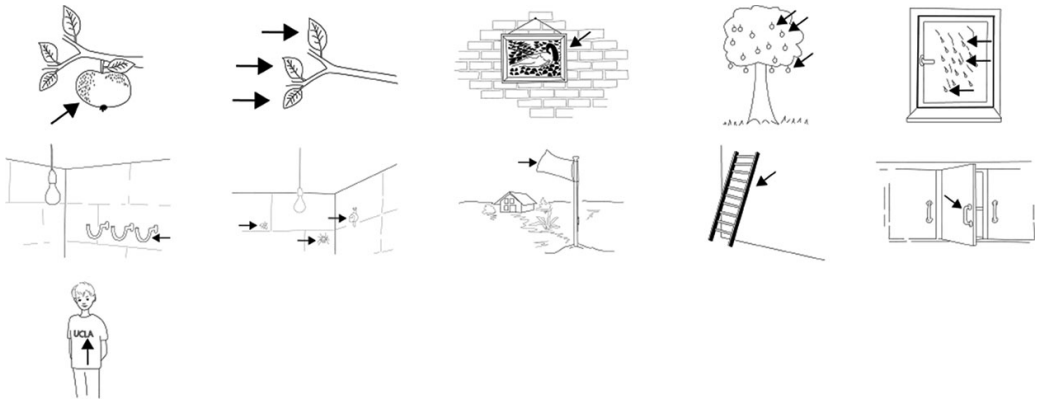


Figure 6 Features: go.around.fig = 0, top.of.ground = 0, go.around.ground = 0, horizontal.ground = 0, ground.building = 0, ground.larger.2d.fig = 1. Ni = 81%, nan = 13%, pohn = 6%, n = 48 (Node 13).

(4)

Kapwud-kau	*mih~mi*	*ni*	*tuhkeh-o*
leaves-DIST.PL	STATV~IPFV	LOC	tree-DIST.SG

'The leaves are on the tree.'

(5)

Sikah-o	*mih~mi*	*ni*	*ewe-n*	*ohl-o*
cigarette-DIST.SG	STATV~IPFV	LOC	mouth-CONST	man-DIST.SG

'The cigarette is in the man's mouth.'

(6)

Delepwohn-o	*mih~mi*	*ni*	*dihd-o*
telephone-DIST.SG	STATV~IPFV	LOC	wall-DIST.SG

'The telephone is on the wall.'

(7)

Doaropweh-kau	*mih~mi*	*ni*	*kisin meteh-o*
paper-DIST.PL	STATV~IPFV	LOC	spike-DIST.SG

'The papers are on the spike.'

(8)

Rihng-o	*mih~mi*	*ni*	*peh-n*	*lih-o*
bag-DIST.SG	STATV~IPFV	LOC	finger-CONST	woman-DIST.SG

'The ring is on the woman's finger.'

(9)

Pwahl-o	*mih~mi*	*ni*	*kep-o*
crack-DIST.SG	STATV~IPFV	LOC	cup-DIST.SG

'The crack is in the cup.'

(10)

Kehndakeh-o	*mih~mi*	*ni*	*dihd-o*
ladder-DIST.SG	STATV~IPFV	LOC	wall-DIST.SG

'The ladder is against the wall.'

None of the participants used *ni* as 'at', as was proposed in the Rehg and Sohl (1981) example: *E mihmi ni imweio* 'He is staying at my house'. Such a usage might be more contextualized and thus not elicited by the Bowerman and Pederson (1992) toolkit. Likewise, the toolkit did not elicit motion uses.

4.2 Pohn

Pohn (previously glossed as 'above') appears predominately in terminal node 6 (Figure 7). The defining feature is top.of.ground = 1, meaning that the figure is vertically above/on the ground (either with or without contact) (11–13).

Examples with contact:

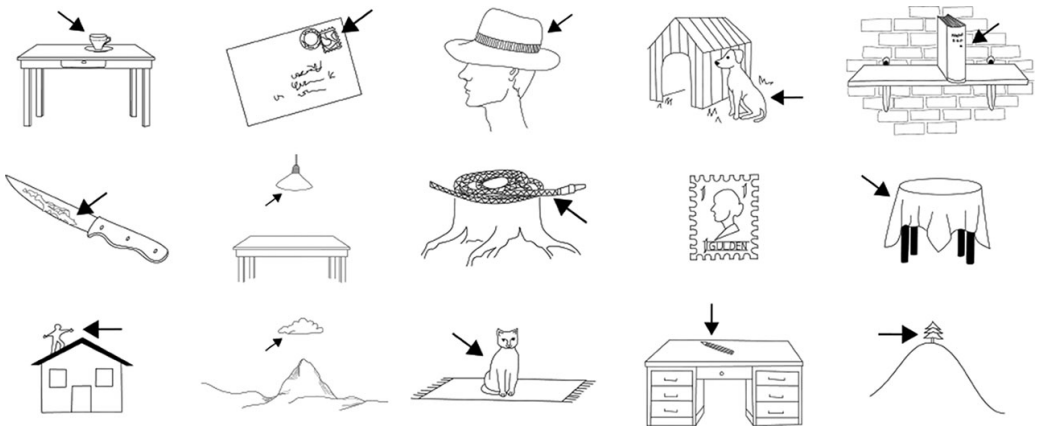


Figure 7 Features: go.around.fig = 0, top.of.ground = 1. Pohn = 87%, nan = 8%, ni = 5%, n =60 (Node 6).

(11)

Kep-o	*mih~mi*	*poh-n*	*tehpel-o*
cup-DIST.SG	STATV~IPFV	above-CONST	table-DIST.SG

'The cup is on the table.'

(12)

Kahto-o	*mih~mi*	*poh-n*	*pingin likou-o*
cat-DIST.SG	STATV~IPFV	above-CONST	rug-DIST.SG

'The cat is on the rug.'

Example without contact:

(13)

Depwek-o	*mih~mi*	*poh-n*	*nahna-o*
cloud-DIST.SG	STATV~IPFV	above-CONST	mountain-DIST

'The cloud is over the mountain.'

The use of *pohn* decreases significantly when the ground goes around the figure such as when the ground completely contains a 2D figure. In such cases *nan* (14) is the most common, followed by *ni* when there is also some sort of attachment (15), then *pohn* (16).

(14)

Lahp-o	*mih~mi*	*nan*	*sdamp-o*
person-DIST.SG	STATV~IPFV	in	stamp-DIST.SG

'The person is on the stamp.'

(15)

Lahp-o	*mih~mi*	*ni*	*sdamp-o*
person-DIST.SG	STATV~IPFV	LOC	stamp-DIST.SG

'The person is on the stamp.'

(16)

Lahp-o	*mih~mi*	*poh-n*	*sdamp-o*
person-DIST.SG	STATV~IPFV	above-CONSTR	stamp-DIST.SG

'The person is on the stamp.'

4.3 Nan

Nan (previously glossed as 'in') is used most commonly in terminal node 3 (Figure 8) where the ground goes around the figure and where the figure and ground are not attached. This can be summarized mostly as the ground containing the figure in some way. This is most easily seen when the ground is 3D and contains the figure (17 and 18).

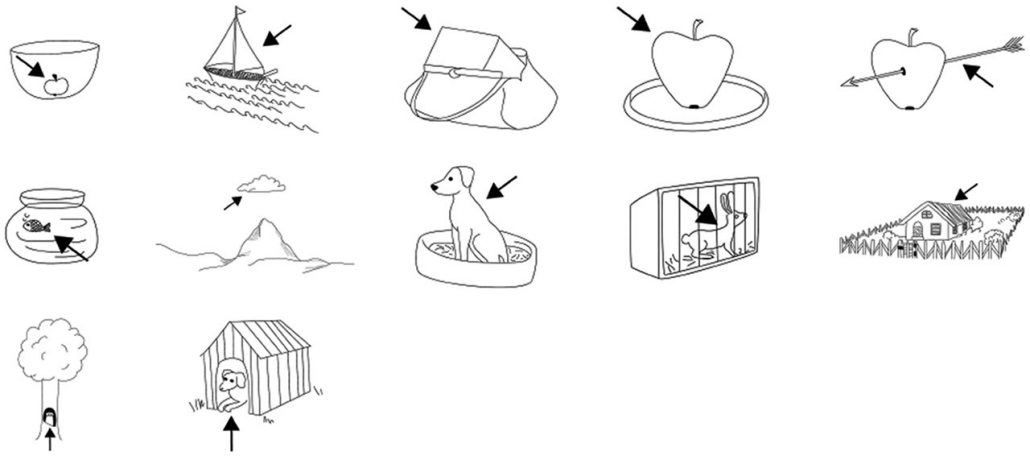


Figure 8 Features: go.around.fig = 1, fig.attached.to.ground = 0. Nan = 92%, pohn = 4%, ni = 2%, loale = 2%, n = 52 (Node 3).

(17)

Apel-o	*mih~mi*	*nan*	*pwohl-o*
apple-DIST.SG	STATV~IPFV	in	bowl-DIST.SG

'The apple is in the bowl.'

(18)

Depwek-o	*mih~mi*	*nan*	*lahng*
cloud-DIST.SG	STATV~IPFV	in	sky-DIST.SG

'The cloud is in the sky.'

It can also be used if a 2D ground is larger than the figure and contains it in that space, though in such cases *pohn* may also be used. In (19) and (20), the ground, rug, is 2D, much larger than the figure, and contains the figure in 2D space so both *nan* and *pohn* may be used, since the cat is both contained by the rug and above it.

(19)

Kaht-o	*mih~mi*	*nan*	*pingin likou-o*
cat-DIST.SG	STATV~IPFV	in	rug-DIST.SG

'The cat is on the rug.'

(20)

Kahto-o	*mih~mi*	*poh-n*	*pingin likou-o*
cat-DIST.SG	STATV~IPFV	above-CONST	rug-DIST.SG

'The cat is on the rug.'

A slight exception to the 2D ground is that *pohn* is typically not used with *sehd* 'ocean, sea' or *madau* 'ocean beyond the barrier reef' but only *nan* as in (21). These words for sea can also be used where *nan* is bound to the word: *sehd* > *nansed*, *madau* > *nanmadau*. The difference in meaning between the bound and unbound forms have not been fully explored.

(21)

Pwoht-o	*mih~mi*	*nan*	*sehd-o*
boat-DIST.SG	STATV~IPFV	in	ocean-DIST.SG

'The boat is in the ocean.'

4.3.1 Loale

Loale (previously glossed as 'inside') was only used twice (once in node 3 and once in node 4) by any of the participants. It was used in a way that seems to be a hyponym of *nan* to mean inside some container/structure (22). However, its use in (23) shows that the figure does not have to be completely inside the ground to be able to use it. It is not completely clear how its usage differs from *nan* and more data are needed.

(22)

Kidi-o	*mih~mi*	*loale-n*	*ihmw-o*
dog-DIST.SG	STATV~IPFV	inside-CONST	house-DIST.SG

'The dog is inside the house.'

(23)

Kehpei-o	*mih~mi*	*loale-n*	*pwohtel-o*
cork-DIST.SG	STATV~IPFV	inside-CONST	bottle-DIST.SG

'The cork is in the bottle.'

4.4 Pahn

Pahn (previously glosses as 'under') occurred predominately in terminal node 14 where the primary features are top.group = 0 and horizontal.ground = 1. The combination of these two features more generally correspond to where the figure is under a horizontal ground regardless of contact (24) [without contact] and (25) [with contact]. It is effectively the opposite of *pohn*.

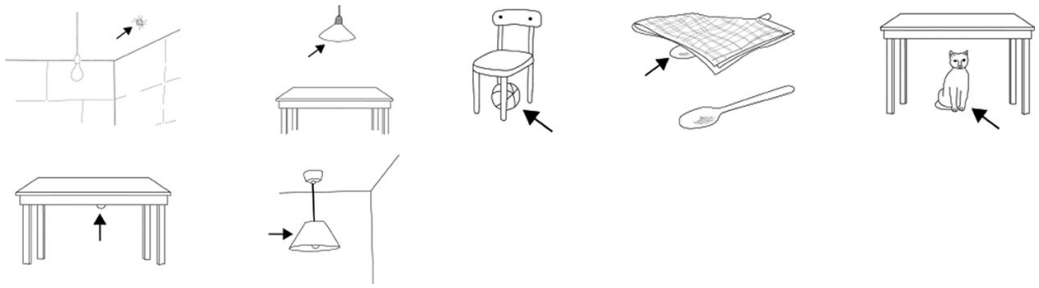


Figure 9 Features: go.around.fig = 0, top.of.ground = 0, go.around.ground = 0, horizontal.ground = 1. Pahn = 79%, ni = 13%, nan = 7%, n = 29 (Node 14).

(24)

Kaht-o	*mih~mi*	*pah-n*	*tehpel-o*
cat-DIST.SG	STATV~IPFV	under-CONST	table-DIST.SG

'The cat is under the table.'

(25)

Kisinpwil-o	*mih~mi*	*pah-n*	*tehpel-o*
chewing.gum-DIST.SG	STATV~IPFV	under-CONST	table-DIST.SG

'The chewing gum is under the table.'

4.5 Likin, limwah, dohsang

The last three locational words, *likin*, *limwah*, and *dohsang* are the only members of terminal node 10 (Figure 12) and *likin* is the node's most frequent member. *Limwah* is also the most prominent member of node 12 (Figure 10). These three words do not represent topological relations but rather are part of the few images in the BowPed Toolkit that elicited frames of reference responses. While not the primary focus of this paper, they will be discussed briefly.

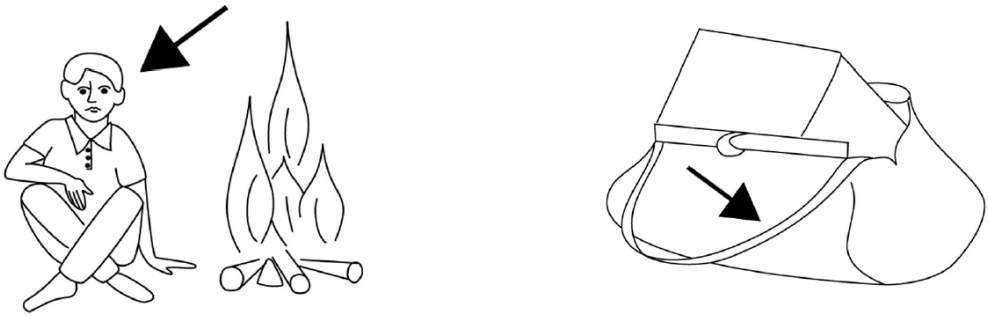


Figure 10 Features: go.around.fig = 0, top.of.ground = 0, go.around.ground = 0, horizontal.ground = 0, ground.building = 0, ground.larger.2d.fig = 0. Limwah = 71%, ni = 29%, n = 7 (Node 12).

Limwah previously has been translated as 'next to'; however, this definition is not completely accurate since *limwah* was used in both Images 38 and 49 shown in Figure 11.

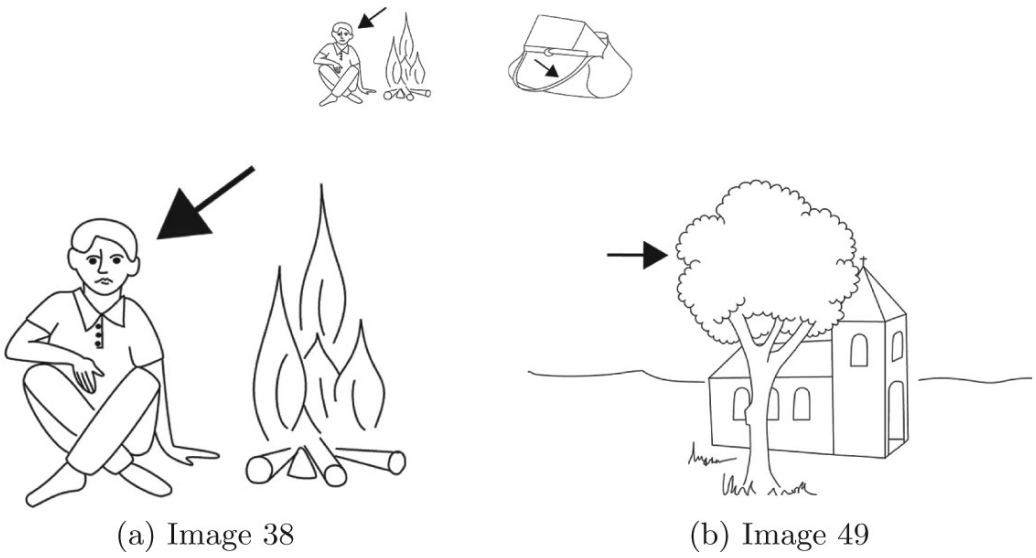


Figure 11 Two images for *limwah* (Bowerman and Pederson 1992).

Given the responses for Image 49, it is not clear if *limwah* means instead something like 'near' or if it means 'next to' based on the orientation of the ground relative to the figure (the tree is next to the church since the tree is near the side of the church as opposed to its front or back) instead of a frame of reference relative to the speaker's position. Given this dataset, the nature of the frame of reference system employed is unclear and future research is needed.

Likin is used when the ground is some kind of enterable structure and the figure is not contained in it. Its meaning is synonymous with 'outside' given the current data.[8] Examples of *likin* are found in (26–28).

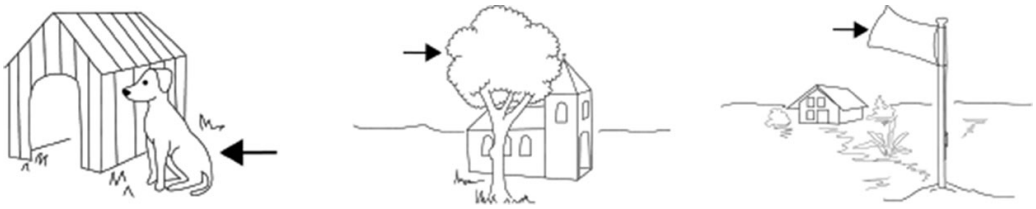


Figure 12 Features: go.around.fig = 0, top.of.ground = 0, go.around.ground = 0, horizontal.ground = 0, ground.building = 1. Likin = 69%, limwah = 23%, ni = 8%, dohsang = 8%, n = 13 (Node 10).

(26)

Kidi-o	*mih~mi*	*liki-n*	*ihmw-o*
dog-DIST.SG	STATV~IPFV	outside-CONST	house-DIST.SG

'The dog is outside the house.'

(27)

Tuhke-o	*mih~mi*	*liki-n*	*ihmw sarawi-o*
dog-DIST.SG	STATV~IPFV	outside-CONST	house sacred-DIST.SG

'The tree is outside the church.'

(28)

Pilaik-o	*mih~mi*	*liki-n*	*ihmw-o*
flag-DIST.SG	STATV~IPFV	outside-CONST	house-DIST.SG

'The flag is outside the house.'

Dohsang was used only once and means 'far away from' as seen in (29).

(29)

Pilaik-o	*doh-sang*	*ihmw-o*
flag-DIST.SG	to.be.far.away-from	house-DIST.SG

'The flag is far away from the house'.

4.6 Alternate uses

Besides the most clear cut situations, *nan, ni*, and *pohn/pahn*[9] all have overlapping uses. These overlapping uses arise in situations where more than one possible type of figure-ground relationship is possible. Since the images were decontextualized, the participants were free to choose from any of the available relationships.[10] For example, in Image 4 (Figure 13), the ribbon is tied around the middle of the candle. Half of the responses for this relationship were *ni* and the other half *nan*. Those who responded *ni* emphasized the attachment of the ribbon to the candle. The *nan* responses emphasized the containment of the ribbon by the candle, especially since the ribbon is in the middle of the candle and the candle is much larger than the ribbon.

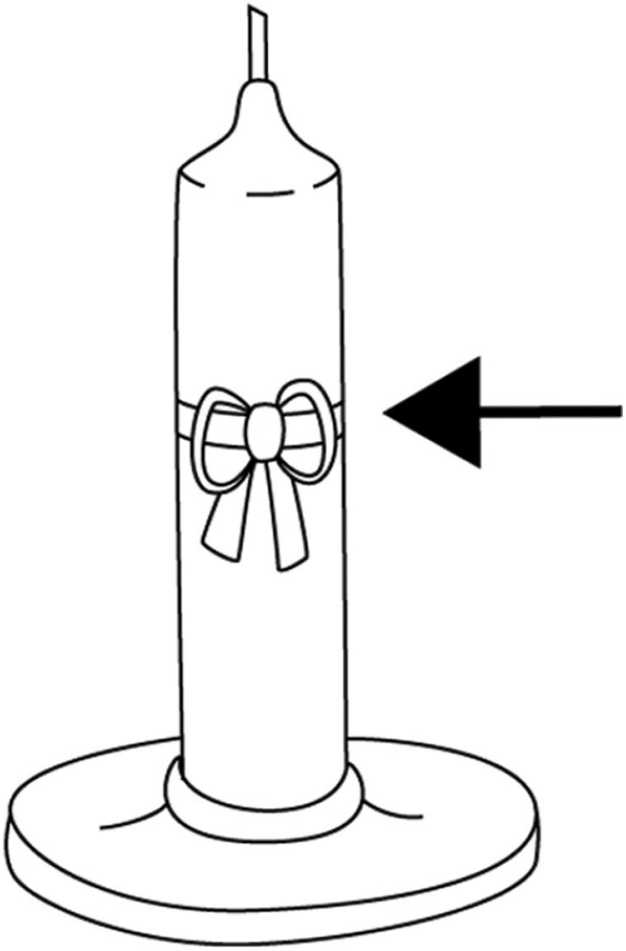


Figure 13 Image 4 (Bowerman and Pederson 1992).

Another example of alternate usages is Image 28 (Figure 14), where a person appears on a stamp. 50% of the responses used *ni* emphasizing that the person is attached to the stamp. 25% used *pohn* to emphasize that the person appears somehow above/on top of the stamp. The other 25% used *nan* to indicate that the person is contained by the stamp.

Figure 14 Image 28 (Bowerman and Pederson 1992).

An example with *pahn* occurs in Image 63 (Figure 15), where a lamp hangs from the ceiling. 50% of the responses used *pahn* emphasizing that the lamp is below the ceiling. 25% used *nan* indicating the lamp is contained by the ceiling and the other 25% used *ni* indicating that the lamp is attached to the ceiling.

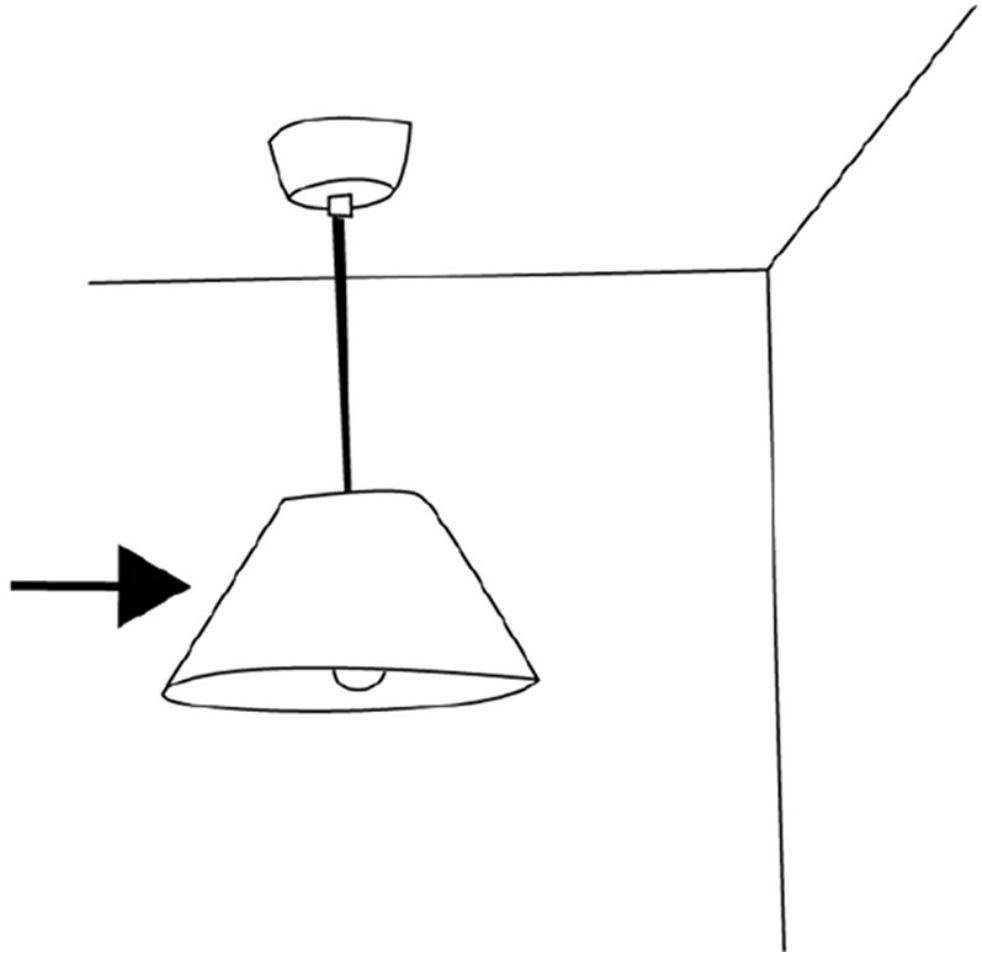


Figure 15 Image 63 (Bowerman and Pederson 1992).

Further evidence for the overlapping uses is that *ni, nan*, and *pohn/pahn* tend to co-occur in the same nodes (Figure 4–Figure 9), though with different frequencies. *Ni* is the only one to occur in nodes without *nan* or *pohn/pahn* (Figure 10 and Figure 12).

Figure 16 depicts some of the overlapping uses of *nan, ni, pohn*, and *pahn*.

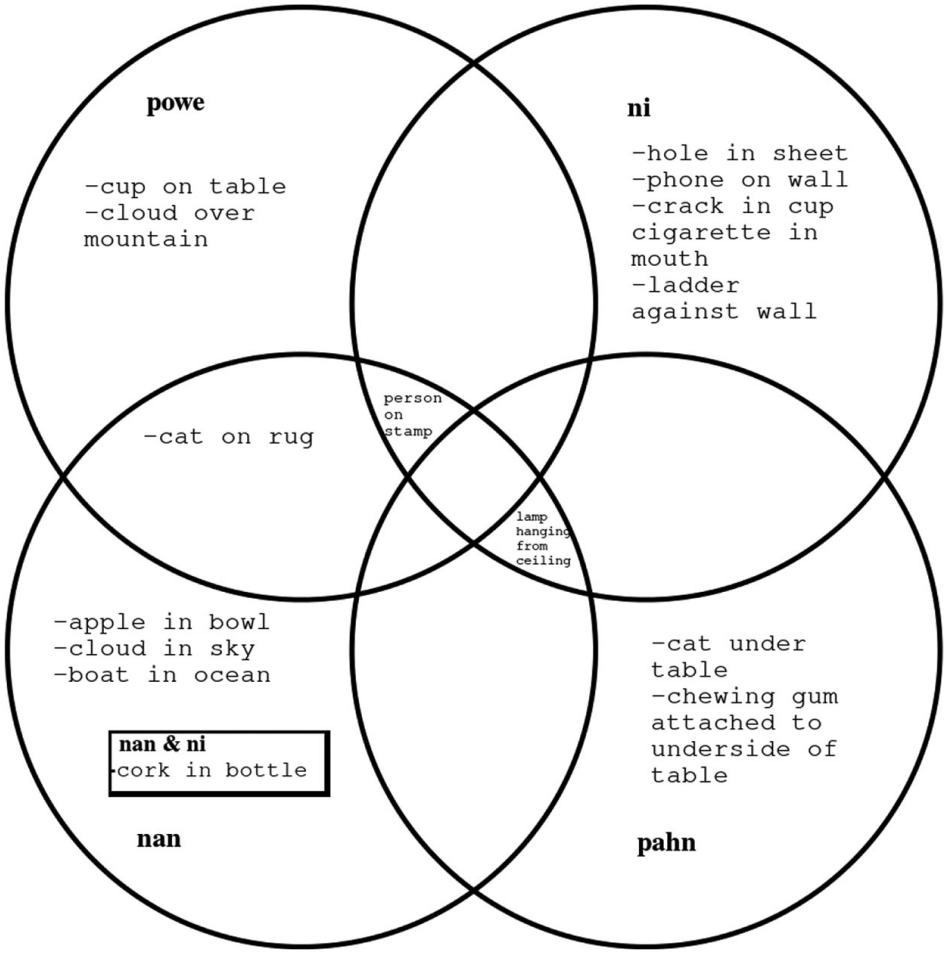


Figure 16 Overlapping uses of *nan, ni, pohn*, and *pahn*.

5 Discussion

The use of binary features to code the ground-figure relationship in each stimulus image was sometimes difficult to code; in some instances, it was fairly ambiguous where the exact cut-off was for a image being '0' or '1' for a given feature. However, since the primary goal of the binary features was to help find patterns in the data, such ambiguities or occasional irregularities did not necessarily hurt the model's outcome. The use of the evolutionary classification tree grouped the 71 images into 8 terminal groups. Out of those 8 groups, *ni* was used predominantly in 3 (nodes 4, 15, and 13) and *nan, pohn, pahn, likin*, and *limwah* in only 1 (nodes 3, 6, 14, 10, and 12 respectively). The nodes where a single word is predominate show its more common usage. Since *ni* occurs predominantly in three separate terminal nodes, a broader super-category 'generalized attachment' was inferred. Likewise, when a word is not the predominate one in a terminal node, it gives some insight into competing thought/categorization processes; different aspects of the figure-ground relationship may have been more salient to some speakers than they were to other speakers.

The use of the evolutionary classification tree model was a useful tool for finding patterns in the data. Some of those patterns may have been found by grouping the data by hand, but the model efficiently created 8 groups of the images (the terminal nodes) with a relatively low error rate. The model not only saved time in the analysis, but also effectively modeled the variation found in the data by showing the percentage of occurrence for each locational word in the terminal nodes. For example, from this model one can see that *ni, nan*, and *pohn* often occur together in the terminal nodes, but the percentage of usage of each one varies based on the node's features. This being said, the generalizability of the tree model has not been tested. It is a common critique that some tree models (especially CARTs) tend to overfit the training data set and may not generalize well (see Baayen 2008; Grubinger et al. 2014; Hothorn et al. 2006; and Alves Torgo 1999; *i.a.*), though the evolutionary classification tree algorithm tends to out-perform other decision tree models. Testing the generalizability of the model was not one of the explicit goals of this paper, rather its effectiveness of finding meaningful patterns within the current dataset, which it successfully did. Other models such as Random Forests may provide more

generalizability but are harder to interpret (see Bae 2008; Dilts 2013; Hill and Jones 2014; and Klavan et al. 2015) and provide few insights into its decision making process.

Given the results of this study, the previous definitions for the locational words posited by Rehg and Sohl (1979), 1981 need to be revised. Based on this study, the following amendments are proposed:

1. *Nan*: Containment (in 3D or 2D space); in, on. *Lahpo mihmi nan ihmwo*. 'That person is in the house.' *Kahto mihmi nan pingin likou*. 'That cat is on the rug.'[11]
2. *Ni*: Attachment; on, in (as in a crack or hole in something), against (as in against the wall). *Rihngo mihmi ni pehn liho*. 'That ring is on the woman's finger.' *Pwahlo mihmi ni kepo*. 'The crack is in the cup.'
3. *Powe*: Above, on, on top of (irrespective of contact). *Liho mihmi pohn ihmwo*. 'The woman is on top of the house.' *Depweko mihmi pohn ihmwo*. 'The cloud is above the house.'
4. *Pah*: Under, on the bottom of, underneath, on the underside of (irrespective of contact). *Kahto mihmi pahn tehpelo*. 'The cat is under the table.' *Kisinpwilo mihmi pahn tehpelo*. 'The chewing gum is on the underside of the table.'

The revised definitions are only applicable to the topological relations observed in this study and do not include meanings involving motion, which were not elicited.

6 Conclusions

Pohnpeian topological relations vary significantly from those of other studied languages. The most basic relational distinctions observed in this study are containment (*nan*) and attachment (*ni*), followed by above (*pohn*) (vertically above regardless of contact) and below (*pahn*) (vertically below regardless of contact). The results also showed that locational word usage is somewhat fluid in decontextualized situations where more than one type of topological relation (in the Pohnpeian system) is possible. In those situations speakers choose one of the possible relational words, though based on the current data, some responses are more probable than others.

The use of the evolutionary classification tree model for this dataset has been advantageous as an exploratory tool. It successfully described patterns in the data by reducing the number of factors from 15 to 7 by creating 8 terminal nodes. The model also greatly reduced the time it would have taken to group the ~300 tokens by hand and helped visualize relationships in the data that may not have been noticed otherwise. Based on this study, I suggest this method as a useful tool for documentary linguists to apply to under-documented languages to better understand other complex linguistic phenomena.

Acknowledgments

I would like to thank Andrea Berez-Kroeker, Gary Holton, Bradley McDonnell, Kenneth Rehg, Robert Andreas, Carisma Jano, and two anonymous reviewers for their help and comments with this project. All remaining faults are my own. Most of all I would like to thank Linda Kihleng-Albert, Maurina Ludwig, Dusty Santos, Joshua Fredrick, and Robinson Fredrick for sharing their knowledge and allowing me to write about their language. Kalahngan en kupwuromwail!

Notes

[1]The International Phonetic Alphabet (IPA) values for the orthographic symbols used for Pohnpeian are as follows: <a> = /ɐ/, <e> = /ɛ/, <oa> = /ɔ/, <mw> = /m^w/, <ng> =/ŋ/, <pw> = /p^w/, <t> = /t̪/, <d> = /t/, <Vh> = /V:/. Other symbols such as <p> and <u> have their expected IPA values. Exceptions to the Leipzig Glossing Rules are as follows: CONST, construct-genitive; PROX, proximal: first degree of proximity (near speaker); MED, medial: second degree of proximity (near listener); DIST, distal: third degree of proximity (away from listener and speaker); STATV, stative.

[2]*Sang* 'from' and *ong* 'to' may also function as prepositions but are most commonly used as verbal suffixes. Rehg and Sohl (1981) occasionally called them prepositions but did not include them with *ni* and *nan*. Rehg (p.c. 2016) further clarified that *sang* and *ong* when used as verbal suffixes are only loosely bound to the verb. *Sang* was encountered as a verbal suffix in this study and neither were used as prepositions.

[3]There are five municipalities (*wehi* in Pohnpeian) on Pohnpei which represent autonomous paramount chiefdoms in Pohnpeian society (Hanlon 1988; Petersen 1982). Each *wehi* is made up of several *kousapw* 'sections'. The Kitti *wehi* has at least two dialect regions that are the

most divergent from the others. A detailed dialect study of Pohnpeian has yet to be done, though all dialects of Pohnpeian on Pohnpei are mutually intelligible.

[4]In selecting a statistical tree-based model for this data set, CARTs, conditional inference trees, and evolutionary classification trees were tried. The evolutionary classification tree model had a lower misclassification error rate that the other models in addition to a lower level of tree complexity (8 terminal nodes vs. 9) and was selected because of its better predictive power coupled with less complexity.

[5]15 binary features allows for 2^{15} or 32,768 combinations. This excessive amount of possible combinations gives room for the model to choose a subset of meaningful features.

[6]The CART model had a 15.1% misclassification rate with 9 terminal nodes.

[7]The model reduced the number of possible combinations to 2^7 (128) from 2^{15} (32,768).

[8]A slight caveat to this is that in English (at least the author's variety), for a figure to be outside a ground it typically implies the ability of the figure to also be located inside the ground. With Image 49, one would typically not say that the tree is outside the church since it would be unusual that it would occur inside it given the size of the tree relative to the church.

[9]This co-occurrence happens equally with *pohn* and *pahn* though in opposite situations since they are opposites. *Pohn* and *pahn* never co-occur with themselves.

[10]Rehg (p.c. 2016) pointed out that in normal conversation, there might not be such variation, because the figure-ground relationship is much clearer based on context.

[11]The revised meanings of *nan* and *ni* can also be shown by their temporal meanings, discussed in Section 2.2, where *nan* can mean 'in an extended period of time', which contrasts with *ni* which describes a single temporal point or event.

References

Ameka, Felix K. 2007. The coding of topological relations in verbs: the case of Likpe (Sɛkpɛlé). *Linguistics* 45(5/6). 1065–1103. DOI: https://-doi.org/10.1515/ling.2007.032.

Aryawibawa, I. Nyoman. 2008. *Semantic typology: Semantics of locative relations in Ronnga (ISO 639-3: ROR)* University of Kansas MA thesis.

Baayen, Harold R. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Bae, Choongsoon. 2008. *Analyzing random forests* University of California, Berkeley dissertation.

Bowerman, Melissa. 1996. The origins of children's spatial semantic categories: Cognitive versus linguistic determinants. In John J. Gumperz & Stephen C. Levinson (eds.), *Rethinking linguistic relativity*, 145–176. Cambridge: Cambridge University Press.

Bowerman, Melissa & Eric Pederson. 1992. Topological relations picture series. In *Space stimuli kit 1.2: November 1992, 51*, Nijmegen: Max Planck Institute for Psycholinguistics. http://fieldmanuals.mpi.nl/volumes/1992/bowped/.

Central Intelligence Agency. 1999. Map of the Federated States of Micronesia. http://www.lib.utexas.edu/maps/islands_oceans_poles/micronesia_pol99.jpg.

Dilts, Philip. 2013. *Modelling phonetic reduction in a corpus of spoken english using random forests and mixed- effects regression*. University of Alberta dissertation.

FSM Office of Statistics, Budget and Economic Management, Overseas Development Assistance, and Compact Management. 2010. FSM-wide census of population & housing: Pohnpei State. http://www.sboc.fm.

Garza, Alejandra Capistrán. 2011. Locative and orientation descriptions in Tarascan: Topological relations and frames of reference. *Language Sciences* 33. 1006–1024. DOI: https://doi.org/10.1016/j.langsci.2011.06.005.

Gentner, Dedre & Melissa Bowerman. 2010. Why some spatial semantic categories are harder to learn than others: The typological prevalence hypothesis. In J. Guo, E. Lieven, N. Budwig, S. Ervin-Tripp, K. Nakamura & S. Ozcaliskan (eds.), *Crosslinguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin*. Psychology Press Festschrift Series, 465–480. New York Psychology Press: https: //books.google.com/books?id=W-B4AgAAQBAJ.

Grubinger, Thomas, Achim Zeileis & Karl-Peter Pfeiffer. 2014. evtree: Evolutionary learning of globally optimal classification and regression trees in R. *Journal of Statistical Software* 61(1). 1–29. DOI: https://doi.org/10.18637/jss.v061.i01.

Hanlon, David. 1988. *Upon a stone altar: A history of the Island of Pohnpei to 1890*. Honolulu: University of Hawai'i Press https://scholarspace.manoa.hawaii.edu/handle/10125/25823.

Hellwig, Birgit & Friederike Lüpke. 2001. Caused positions. In Stephen C. Levinson & N. J Enfield (eds.), *Manual for the field season 2001*, Nijmegen: Max Planck Institute for Psycholinguistics. http://fieldmanuals.mpi.nl/volumes/2001/caused-positions/.

Hezel, Francis X. & Michael J. Levin. 2012. Survey of Federated States of Micronesia migrants in the United States including Guam and the Commonwealth of Northern Mariana Islands (CNMI). http://www.healthypacific.org/uploads/1/0/0/6/10065201/combined_final_report.pdf.

Hill, Daniel W. & Zachary M. Jones. 2014. An empirical evaluation of explanations for state repression. *American Political Science Review* 108(3). 661–687. DOI: https://doi.org/10.1017/S0003055414000306.

Hothorn, Torsten, Kurt Hornik & Achim Zeileis. 2006. Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics* 15(3). 651–674.

Jackendoff, Ray. 1996. The architecture of the linguistic-spatial interface. In Paul Bloom, Mary A. Peterson, Lynn Nadel & Merrill F. Garrett (eds.), *Language and space*, 1–30. Cambridge: The MIT Press.

Klavan, Jane, Maarja-Liisa Pilvik & Kristel Uiboaed. 2015. The use of multivariate statistical classification models for predicting constructional choice in spoken, non-standard varieties of estonia. *SKY Journal of Linguistics* 28. 187–224.

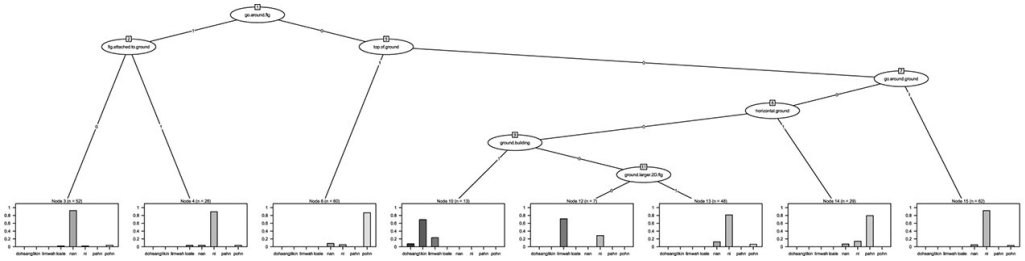
Levinson, Stephen & Sérgio Meira. The Language Cognition Group. 2003. 'Natural concepts' in the spatial topological domain – adpositional meanings in crosslinguistics perspective: an exercise in semantic typology. *Language* 79(3). 485–516. http://www.jstor.org/stable/4489463.

Levinson, Stephen C. 1996. Language and space. *Annual Review of Anthropology* 25. 353–382. http://www.jstor.org/stable/2155831.

Levinson, Stephen C. 2003. *Space in language and cognition: Explorations in cognitive diversity*. Cambridge: Cambridge University Press.

Levinson, Stephen C. & David P. Wilkins. 2006. The background to the study of the language of space. In Stephen C. Levinson & David P. Wilkins (eds.), *Grammars of space*, 1–23. Cambridge: Cambridge University Press.
Lichtenberk, Frantisek. 1985. Possessive constructions in Oceanic languages and Proto-Oceanic. In Andrew K. Pawley & Lois Carrington (eds.), *Austronesian linguistics at the 15th Pacific Science Congress*, 93–140. Canberra: The Australian National University.
Petersen, Glenn. 1982. *One man cannot rule a thousand: Fission in a Ponapean chiefdom*. Ann Arbor: The University of Michigan Press.
R Core Team. 2016. R: A language and environment for statistical computing (version 3.3.1). http://www.R-project.org.
Rehg, Kenneth L. & Damian G. Sohl. 1979. *Ponapean-English Dictionary*. PALI language texts: Micronesia. Honolulu: University of Hawai'i Press.
Rehg, Kenneth L. & Damian G. Sohl. 1981. *Ponapean reference grammar*. PALI language texts: Micronesia. Honolulu: University of Hawai'i Press.
Talmy, Leonard. 1983. How language structures space. In *Spatial orientation: Theory, research and application* New York: Plenum Press, 225–282.
Therneau, Terry M. & Elizabeth J. Atkinson. 2015. An introduction to recursive partitioning using the RPART routines.
Thiering, Martin. 2006. Topological relations in an athapaskan language. *University of Alberta Working Papers*. http://hdl.handle.net/11858/00-001M-0000-0012-B81E-6.
Thiering, Martin. 2007. *The construction of topological space*, University of Alberta dissertation.
Thiering, Martin. 2011. Figure-ground reversals in language. *Gestalt Theory* 33(3/4). 245–276.
Thiering, Martin. 2015. *Spatial semiotics and spatial mental models: Figureground asymmetries in language*. Berlin: De Gruyter Mouton.
Torgo, Luís Fernando Raíhno Alves. 1999. *Inductive learning of tree-based regression models*. Universidade do Porto dissertation.
Vandeloise, Claude. 1991. *Spatial prepositions: A case study from French*. Chicago: The University of Chicago Press.
Zlatev, Jordan. 2007. Spatial semantics. In Dirk Geeraerts & Hubert Cuyckens (eds.), *The Oxford handbook of cognitive linguistics*, 318–350. Oxford: Oxford University Press.

Appendix



Full evolutionary tree model