12

Productivity vs predictability

Evidence for the syntax and semantics of Animate gender in four Northeastern-area Algonquian languages

CONOR MCDONOUGH QUINN

12.1 Introduction

The nominal gender distinction in Algonquian languages standardly known as *Animate* (AN) vs *Inanimate* (IN) is realized as pervasive selectional and/or agreement effects in key nominal-argument tracking morphology. Representative examples from Penobscot (Eastern Algonquian, central Maine, USA) showcase the AN versus IN contrast in nominal number (1a) and in verbal stem-selection (1b, underlined).

(1) **Animate vs Inanimate**

	Animate		**Inanimate**	
a.	*owa*	*phènəm*	*iyo*	*pənahpəsk*^{w}
	this.AN	woman	this.IN	rock
	'this woman'		'this rock'	
	(i)yok	*phènəm-ok*	*(i)yòlil pənahpsk-ol*	
	this.ANpl	woman-ANpl	this.INpl	rock-INpl
	'these women'		'these rocks'	
b.	*wəlíhpokəso*		*wəlíhpokat*	
	'Animate tastes good'		'Inanimate tastes good'	
	nə mohα		*nəmìčin*	
	'I eat it (Animate)'		'I eat him/her/it (Inanimate)'	

(Penobscot: Siebert 1997/PD, glosses CQ)

Algonquianist research has long grappled with the question of whether or not a (fully) predictive account for Animate/Inanimate status is possible. With a few

Gender and Noun Classification. First edition. Éric Mathieu, Myriam Dali, and Gita Zareikar (eds)

exceptions, the standard answer has been no. This is because while the AN/IN contrast appears to originate from semantic animacy, it also shows all the signs of having formalized away from it.

Hence, while humans, animals, and supernatural entities are uniformly treated as Animate, many grammatical Animates fall well outside of the core domain of familiar semantic animates (2).

(2) **Animates outside of the domain of familiar semantic animate (*Penobscot; PD*)**

èmkʷɑn	'spoon'
màkikʷ	'nasal mucus, snot, catarrh'
məlǝčess	'mitten'
kɑ̀wi	'porcupine quill'
tálɑkan	'wedge'

Furthermore, despite being describable with verbal stems translating as 'live', 'grow', and 'die' (3b, 3c)—which imply semantic animacy, while still realizing as *IN*-gender verb stems—plants are generally (i.e. with some exceptions) treated as grammatical Inanimates (3a).

(3) Plants as grammatically Inanimate living things *(Passamaquoddy-Maliseet)*

a. <pĕm´-au-sŭ-wi´-kil> 'Living things' / 'Trees and plants'
= *pemawsuwikil* = lit. '(Inanimate) things that are alive'
(Chamberlain 1899: 39)

b. *pomawsuwiw* IN verb 'It is alive' (*PMP*)
Psi-te keq hokek Laks nukcoktehson; tehpu wawikon pemawsuwik.
'Every part of Laks's body is smashed; only his backbone [IN!] is still alive. (LM)'

c. *mɑcinewiw* IN verb 'it dies, it is dead' (*PMP*)
pomikon IN verb 'it is growing, keeps growing'

In particular, (3) counters a still-common received—and possibly imposed—notion, never confirmed by native speakers, but having wide currency in pedagogical and anthropological literature: namely, that words of the kind seen in example (2) are somehow conceived of as 'alive' (or having 'power') in some culture-specific way. (See Dahlstrom 1995: 65 and Goddard 2002: 225 for further Algonquian-internal refutations of that view, and Polinsky and Jackson 1999 for a more general challenge to such claims.)

Most current accounts agree that Animate is the only marked category, with Inanimate being the default, elsewhere category (Dahlstrom 1995; Goddard 2002; Kramer 2015). The primary question is then what, if anything, might determine 'unpredictable' Animate statuses like those in example (2).

The semantic content of the lexeme seems to be significantly implicated in all accounts of Algonquian gender assignment. (Strikingly, no evidence has yet been found or offered for it ever being sensitive to phonological form.) Overall, discussions of Animate gender assignment standardly recognize areas of semantic

generalizability but emphasize an ultimately arbitrary, unpredictable character. Hence the following observations for three different Algonquian gender systems:

> . . . the gender of nouns is entirely predictable within certain semantically defined domains. . . . One might, in fact, take each noun and its near synonyms to define a semantic field with an associated gender rule. . . . [But] gender is arbitrary in other parts of the vocabulary. . . . It seems clear that . . . the animate/inanimate distinction is truly one of grammatical gender, despite the fact that gender has various semantic correlates.
>
> (LeSourd 1993: 9, re Passamaquoddy)

> The concepts animate and inanimate are semantically transparent most of the time in Miami-Illinois. . . . The indeterminacy about which nouns are animate and which are inanimate is found within the same semantic fields in all Algonquian languages: plants, body-part terms, and various culturally important objects. (Costa 2003: 205, re Myaamia)

> [Re the idea] that there are some semantic correlates in Algonquian languages between animacy and inanimacy . . . although such distinctions appear important in individual, contrastive instances, no general, single distinction such as greater 'activity' can be invoked to account globally for animacy in Arapaho, and many members of the category are simply inherited from Proto-Algonquian. (Cowell and Moss 2008: 53, re Arapaho)

Part of why accounts of Algonquian gender tend to conclude that it is ultimately not predictable is a recurrent analytical problem: positing a (straw-man) category, and then observing that Animate membership is inconsistent within it. For example, we can set up the following semantic categories and observe that in Penobscot, Animate status within them appears arbitrary:

[*utensils*]: 'Forks and knives are Inanimate, but spoons are Animate!'
[*berries*]: 'Blueberries are Inanimate, but raspberries are Animate!'
[*clothing*]: 'Pants are Inanimate, but shirts are Animate!'
[*body parts*]: 'Hands are Inanimate, but elbows are Animate!'

This problem points to a possible solution: looking very critically at (a) *exactly* what putative categories might best characterize the data, and (b) what kind of processes might give rise to them. As we will see, none of the Animates listed above are nearly as random as this initial set of categories makes them first appear.

12.2 Alternative accounts, and the current proposal

At least three major accounts have in fact pushed the 'semantic correlates' aspect much further:

> . . . [T]he animate category is best viewed as having a subset of central members, with semantic extensions connecting most of the other members of the category, [i.e., as *radial categories* (Lakoff 1987)]. The extensions connecting peripheral members with more central members are semantically motivated—that is, once you know them, they make sense—but the membership of the category is not semantically predictable. (Dahlstrom 1995: 52, re Meskwaki)

> In Penobscot at least, it appears that animacy is determined largely by analogy between individual words, rather than by one elusive, overarching semantic feature that all members

> of the class 'animate' share. . . . Animacy-assigning analogies are not random; they seem mainly (but not exclusively) to be made along the semantic lines of intrinsic function and texture. Although exceptions exist, this characterization accounts for the overwhelming majority of animate nouns attested in the language. (Quinn 2001, re Penobscot)

> In every case the nouns in ([examples]) are inanimate when used for the generic and the ordinary, and animate when used for the special or the unusual.

> The basic meaning of the animate gender is a function of the contrast with the inanimate gender. Looking at the animates by themselves and attempting to connect the dots does not reveal it. (Goddard 2002: 214, 224 respectively, re Meskwaki)

With this in mind, we work from a model that (like Goddard 2002) acknowledges the contrastive function of AN vs IN, but notes that AN membership seems much more internally consistent/constrained than a broad 'special/unusual instance of . . .' criterion would suggest. We still aim to 'connect the dots' between ANs, but again *not* by seeking one single shared feature, but instead by following (like Dahlstrom 1995 and Quinn 2001) a radial categories approach.

Dahlstrom's (1995) initial radial categories approach (Lakoff 1987) models Animate-class membership as a central-vs-peripheral prototype system. However, Dahlstrom never presented direct evidence for a particular pre-set (or persistent) set of central-member prototypes. So it may be equally likely that the Animate-assignment system is less initially-structured, and more emergent: consisting only of an ongoing process of analogical attraction. This would create *centerless clusters* rather than core-vs-periphery systems; it would be limited by not by central prototypes but just by definable constraints on what range of semantic features the analogical process may and may not attend to. Hence the observation of Quinn 2001:

> . . . . it will never be possible to describe perfectly tidy semantic 'classes' of animate nouns. But this is to be expected, since these 'classes' result from the analogical process, and do not determine them.

While we assume and use this small modification to the radial categories model, defending it is not the primary goal of this chapter.

Instead, the novel contribution here is to counter the tenacious myth that Algonquian gender is effectively unpredictable—a view that persists as the descriptive norm, despite the works just cited. To do so we marshall several sets of evidence that each motivate a renewed focus on how the overall process results in a very structured *productivity* of AN assignment (especially for loans, etc.; see Dahlstrom 1995: 61–2). That is, we present evidence that supports extending Dahlstrom's post-hoc 'once you know them, they make sense' observation in a key direction: once they make sense, *you can make more of them.*

The overall picture, then, is that synchronic competence in AN assignment is *emergent*: not a one-time fixing of an all-predictive rule that tracks one single (yet thus-far undiscovered) feature, but instead a mutable, ongoing lexicon-structuring process that builds up a set of lexical-semantic 'families' to which AN status is assigned.

These remarks here, while still preliminary, are based on an exhaustive survey of current lexicographic documentation combined with preliminary field research,

covering the following four Northeastern-area Eastern Algonquian languages: Passamaquoddy-Maliseet (*PsmMl*), Mi'kmaw (*Mw*), Penobscot (*Pb*), and Western Abenaki (*WAb*).[1,2]

For these languages at least, all evidence suggests that the Animate gender category is in fact quite dynamically synchronically productive—and far more predictable than not. Where in-depth native-speaker consultation is still possible (PsmMl, Mw), new preliminary data (examined in Section 12.4) suggests something to date barely observed in the Algonquianist literature: that speakers have a robust and largely consistent knowledge of the gender assignment of novel items—designata for which they know no pre-existing word—and of foreign-word designata in general.

Adding observations from extant Pb and WAb corpora data, we conclude that this productivity does not come from some elusive single semantic 'deep thread' that connects all grammatical Animates. Instead, it comes from a network of emergent clusters or 'families' (Section 12.1) that systematically (and quite restrictedly) attracts semantically related new members (see. Dahlstrom 1995; Quinn 2001; Wittgenstein 1953).

Two phenomena, dual animacy and variable animacy (Section 12.3.2), also support this 'family'-based model, as does the observation that Animate assignment appears to change diachronically by semantic cluster (i.e. by 'family') rather than by individual lexeme (Section 12.5). Establishing that the system is more productive than not, however, only raises further questions: particularly how to investigate it in a falsifiable fashion (Section 12.6.1), and how to model the semantic (Section 12.6.2) and syntactic (Section 12.6.3) mechanisms involved.

Up till now, the observation that Algonquian gender is closely correlated with semantic animacy, but always carries a large residue of (at least superficially) arbitrary 'unpredictable' Animates, seems to have discouraged in-depth investigation into the productivity of these types of systems.

The preponderance of evidence from the languages surveyed here, however, suggests that the phenomenon is an active, dynamic process that is only minimally understood. If we continue to neglect it, we will miss a crucial chance to investigate potentially very enlightening questions about how categorization (linguistic or otherwise) works as a cognitive process. Hence this chapter is not a comprehensive

[1] Data for each of these languages is drawn from the following sources. *Passamaquoddy-Maliseet*: dated field notes from Sipayik, Maine, USA and St Mary's, New Brunswick, Canada, as well as the *www.pmportal.org* online dictionary (cited as *PMP*). *Mi'kmaw* (*Mi'kmaq, Mi'gmaq, Mi'gmaw*): dated field notes from Listuguj, Quebec, Canada, and an unpublished MS dictionary. *Penobscot*: field notes of Frank T. Siebert, Jr, and the unpublished (in-process) *Penobscot Dictionary* MS (both archived at the American Philosophical Society, Philadelphia; the latter cited as *PD.*) *Western Abenaki*: Laurent 1884, cited as *L84*; Day 1994 and Masta 1932 also inform the survey. Brief examples of Caniba/Kennebec Eastern Abenaki (*Cn/Kn*) are from Aubery 1995, cited as *Aub95*. Many thanks to Margaret Apt (MA), MaryAnn Metallic (MAM), Roger Paul (RP), Darryl Nicholas (DN), and Victor Atwin (VA), among others, for sharing their expertise in their respective languages.

[2] For space and clarity, we simply list Animate/Inanimate status as a given. The actual evidence comes from morphological markers like those seen in example (1), i.e. mainly nominal plurals and argument-matching verbal-stem forms. Explicit part-of-speech dictionary listings have also been used.

account of Algonquian gender productivity, but a call to finally initiate a much-needed, multi-scholar research program into it.

Investigation into more predictive accounts of Algonquian gender also benefits more than just academic research. The still-standard narrative intimidates would-be learners with 'unpredictable Animates', and frustrates them by first imposing the technical terms 'Animate' and 'Inanimate' and then making them essentially meaningless/mysterious—to learners and native speakers alike. In contrast, an account that offers even partial productivity encourages learners from day one, and so could significantly help ongoing Algonquian language reclamation/revitalization efforts.

12.3 A 'family'-based model of Animate status

12.3.1 The 'family'-based approach

For all four languages surveyed, the following 'family' categories of Animates are extremely robust, being nearly exceptionless within each language.

(4) **Animate-assigning 'families' in Northeastern-area Algonquian languages**
people, animals, spirits, luminaries,
+ *representations* of these (pictures, glyphs, playing cards-gaming pieces/balls, [coins])
substantial trees (*e.g. pine but not [alder]*)
fluid containers
thorns-quills-feathers-scales/flakes
gum-swellings-substantially squishy fruits/berries-root vegetables-[bread products]
net-{rope/cord}[3] snowshoe-[footwear]-handwear
shirt-jacket/{coat} breechcloth/diaper
[wheels-disks]

Here hyphens link tentatively related 'families', while 'families' marked in plain brackets are those systematically absent in at least one language surveyed.

Both this systematic variation and concrete examples of these 'families' will be examined in data below. (Again, we stress that the glosses in example (4) are merely labels for emergent clusterings, and specifically not claims for pre-existing core prototypes.)

These groupings are *not* uniform across all Algonquian languages, though there does seem to be more overlap than not: see especially Bloomfield 1962: 28–36 for Menominee, and Goddard 2002 for Meskwaki and Algonquian overall. Instead, 'families' appear to come and go across time and region. Hence for example there are two robust AN categories (5) totally absent in the four Northeastern-area languages surveyed, but found in another Algonquian language, Munsee.

[3] This {rope/cord} grouping is tentative: it is robust overall, but a strong set of exceptions suggests that a more precise characterization is still needed. Significant variation also obtains for items in the {coat} category. For these two we use curly brackets to indicate a 'family' whose exact status remains unclear.

(5) **Distinct AN categories in Munsee (O'Meara 1996: 405, 401, 538; 35, 414, 414)**

[*books-paper-documents*]	*pámbiil* 'book, paper, letter', *mbáypul* 'Bible', *noos-pépul* 'newspaper'
[*vehicles*]	*amóoxol* (AN/IN) 'boat', *ahtamóombiil* 'car', *káal* 'train car'

As we will see in Section 12.5, this kind of *systematic* variation underscores what we miss when we hew to a model that treats 'unpredictable' Animate assignment as a largely arbitrary, lexeme-by-lexeme process.

12.3.2 Evidence for 'family' effects: dual animacy and variable animacy

Two phenomena support a view of semantic 'family'-based Animate assignment over a more arbitary system.

The first is *dual animacy*: where the same stem shows meaning differences according to its use as either an AN or an IN. The crucial observation is that seen in example (6): the semantics of the AN use of the dual-animacy stems tie readily to established 'families' of Animates such as FEATHER/FAN (6a), and so on:

(6) **Dual animacy in Penobscot** (PD)

a. *wə̀təhɑkan*	AN:	'fin' (see AN *wə̀ləkʷan* 'wing': FEATHER/FAN)
	IN:	'paddle, oar'
b. *kəlósəwɑkan*	AN:	'speech wampum; large, long belt of wampum, as used for intertribal treaties' (GLYPH; CORD)
	IN:	'a word, the word; speech, talk'
c. *pə̀ko*	AN:	'chewing gum' (cf. *màkikʷ* 'nasal mucus, snot, catarrh': GUM/CLOT)
	IN:	'gum, pitch (in sap form, or when used to pitch a canoe)'[4]
d. *[p]skʷ[h]ám[]ɑkan*	AN:	'lacrosse [*or generic*] ball: BALL'
	IN:	'lacrosse game'

Dual animacy has been noted for other Algonquian languages: e.g. Costa 2003: 208 for Myaamia reports pairs like AN *ahkihkwa* 'drum' vs IN *ahkihkwi* 'kettle'; and AN *mihtekamina* 'June bug' vs IN *mihtekamini* 'acorn' (see also Déchaine, this volume, and Dahlstrom 1995).[5] That the designata for AN uses of dual-animate stems reliably

[4] The GUM/CLOT grouping seems to consist strictly of designata that form very densely/thickly cohesive lumps. This is reflected in PsmML *sakoliyat* 'plaster', literally 'that (Animate) which turns hard' and loanwords *weks* 'wax' and *ocis* 'cheese' as ANs (PMP), with AN/IN variation attested (as perhaps expected) for various terms for clay; whereas Pb *kəlamótikan* 'adhesive, glue, mucilage' is consistently IN (hence the IN use of *pə̀ko* here), as are all terms for dung across these languages.

[5] Myaamia is a very conservative Algonquian language in preserving singular gender markers: *-a* 'Animate', *-i* 'Inanimate'. Note too that the first pair here is an additional example of systematic cross-Algonquian Animate assignment variation (Section 12.5), since kettles (as fluid containers) are Animate in all of the Northeastern-area languages surveyed; and for all but Mi'kmaw, drums are Animate as well.

track language-internal Animate 'families' strongly suggests that Animate status remains closely tied to lexical semantics, and is not purely formal.[6]

The second phenomenon is *variable animacy*: where the same stem shows AN/IN variation. Crucially, attested cases of this kind appear to track exactly the margins/fuzzy edges of established 'families'. So, for example, in Penobscot *mskíhkəwimin* 'strawberry' is generally AN, but is attested as IN (plural) *mskihkəwíminal* 'strawberries' for one speaker (SDMC: 998), as well as IN plural for one diminutive form, *mskihkəwíminsal* (PD).

This is as expected. In Penobscot, smaller berries (e.g. *sàhtal* 'blueberries') are uniformly IN; larger soft berries (and soft-peeled fruits: e.g. *mìnsəss* 'raspberry' and *ččikəne* 'apple') are uniformly AN. Strawberries would therefore be generally predicted to be AN (as they are): but the much smaller wild berry may fall below that threshold—hence especially the IN diminutive attestation.

If the system consists of emergent 'families', then we expect variability around the margins of robust 'families': but we also have a modicum of falsifiability, in that the model predicts gender variation only for 'borderline designata' of this kind. (We should also note that logically, some reports of variable animacy might just cases of dual animacy where documentation has missed the precise designatum difference.)

It is worth noting in passing that the dual animacy pattern (and thus this kind of gender system overall) shows suggestive similarities to the English mass vs count system. We observe that the English mass vs count distinction has similar properties: it is semantically based, productive, largely predictable, and meaning-constraining—but at the same time it is still tied closely to idiosyncrasy in lexicalization.

Hence for example the contrasts in (7), where the count-noun uses of *speech* and *drive* emerge from the predictably count-noun sense of '(unit) instance of…', but then develop unpredictable narrowings of designatum from there.

(7) **English count vs mass semantic distinctions**

a. COUNT *(a) speech* (= lecture!)
   MASS *speech* (= general act) [*I heard a speech* ≠ *I heard speech*]

b. COUNT *(a) drive* (= golf!)
   MASS *drive* (= motivation!) [*she's got a drive* ≠ *she's got drive*]

Algonquian dual animacy reflects a similar pattern: exact lexical meaning is as unpredictable as history so often can make it—but the range of *possible* unpredictable meaning is nonetheless sharply semantically constrained by the system. Comparative investigation into the productivity of Algonquian gender vs English mass/count may thus offer insights into how to model both types of system. (For a discussion of mass-count itself in Algonquian languages, see Bale and Coon 2014 for Mi'kmaw; Mathieu

[6] Perhaps the most celebrated forms of dual animacy has not been recognized as such: in at least some Algonquian languages, Inanimates, when spoken of as personified entities, can receive grammatical Animate treatment (Bloomfield 1962: 28; Goddard 2002: 202–9; see especially the latter for critical caveats). Since persons are Animates par excellence, this is simply another, very clear instance of Animate status tracking the categorical affiliation of the precise designatum.

2012a, 2012b for Ojibwe; Gillon 2009 for Innu; and Wiltschko 2012 and Wiltschko and Ritter 2015 for Blackfoot.)[7]

12.4 Evidence for dynamic synchronic productivity: Passamaquoddy-Maliseet and Mi'kmaw

Perhaps the most striking data in favour of the 'family' based model of Animate status remains preliminary and anecdotal—but quite suggestive: speakers of Passamaquoddy-Maliseet and Mi'kmaw show consistent, productive knowledge of the Animate status of English words, and/or newer or unfamiliar designata.

Hence for (8a), a Passamaquoddy speaker's instant ability to select the AN plural *-ok* for English forms *thermos* and *syringe* tracks the fact that fluid containers and needle-like objects both form robust Animate 'families' in that language. Similarly, in (8b), two speakers' immediate knowledge of the AN status of a visible but unfamiliar designatum (dates) tracks the general treatment of fleshy, soft-peeled fruits as Animate. Example (8c) documents the same effect for a speaker of Mi'kmaw. Example (8d) tracks again a Maliseet speaker's immediate identification of a fluid container as AN; and (8e) two different Maliseet speakers' active use of the soft-fleshy-fruit 'family' as a trigger for AN status.

(8) Synchronic productivity of Animate assignment in Passamaquoddy-Maliseet and Mi'kmaw

a. Q: [*re plurals*] What about "thermos"? A: Thermos-*ok*.
Q: And "syringe"? A: Syringe-*ok*.
(*Psm: 200605-MA*)

b. [*speaker looking directly at basket of unfamiliar fruit*]: What are those?
[Before CQ can even answer ("Dates"), speaker 1 starts discussing with speaker 2, referring to them using <u>*AN*</u> morphology.]
(*Psm: 200605_-ML*)

c. Q: [*re plurals*] What about "dates"?
A: [*w/ no hesitation; <u>AN</u> pl.*] Dates-*ig*.
(*Mw: 201407_-MAM*)

d. Q: [(*in discussion of wòt [AN] vs. yùt [IN] 'this'), and indicating a can of soda*] And this?
A: wòt.
(*Ml: 20150729-RP*)

e. Q: [*handing over bag of dried apricots*] Can you describe these [illegible]m
A: Kespahtekil [= <u>*IN*</u> *dried ones*] . . . [*seeing that they're apricots*] . . .
. . . .no, I should say "<u>*kespahsicik*</u>" [= <u>*AN*</u> *dried ones*]

[7] One could probably also make comparable observations for stability of aspect/Aktionsart in otherwise idiomatic verb-particle constructions; see especially Wiltschko and Ritter (2015) for discussion of parallels between animacy, boundedness, and Inner Aspect.

Q: So, [*AN pl.*] "*apricots*-ək"? A: Yes.
Q: And "*dates*-ək"? A: Yes.

[Conversation continues: another speaker volunteers as directly related example: "*cherries*-ək" . . . noting, "You can't say [*IN pl.*] *cherries*-əl", and confirming correctness of researcher-offered "*peaches*-ək".]
(*Ml: 20150729-DN/VA*)

More formal experiments than these are of course still needed—particularly to test out 'families' beyond just the two demonstrated here (fluid containers, soft/fleshy fruits). But these initial observations are representative of a clear effect: speakers seem to show a robust and largely consistent knowledge of the gender assignment of novel, recent, and foreign (in these cases, mainly English) word designata.[8]

12.5 The 'family'-based model's synchronic processes in relation to variation and (systematic) diachronic change

A further contribution of the 'family'-based model is that it explains why diachronic change in gender assignment across Algonquian proceeds not simply on an individual lexeme-by-lexeme basis, but by semantic cluster.

Hence we observe, for example that SHOE (and boots, etc.) is quite generally IN across Algonquian, and remains so in Western Abenaki and Caniba/Kennebec (E. Abenaki). But from Penobscot east/northward, it is consistently AN.

(9) SHOE as IN and AN across Algonquian
IN: WAb *Mkezenal* [= *IN* pl.] 'Shoes; moccasins'
(*L84:26*)
Pôlsal 'Boots'
Cn/Kb *ne makesenar* [= *IN* pl.] 'mes souliers'
(Aub95: 482 and most other Algonquian languages)
AN: Pb *màhksən* 'shoe' PsmMl *pkoson* 'shoe' Mw *mg'sn* 'shoe'

This appears to be an area-specific innovation, perhaps motivated by analogy to terms for 'snowshoe', which are generally AN across Algonquian.

Such consistent diachronic semantic-cluster changes are not rare. BREAD and comparable products are consistently AN in Western Abenaki, Penobscot, Passamaquoddy-Maliseet, and consistently IN in Mi'kmaw (10).

(10) BREAD as AN and IN across Northeastern-area Algonquian
AN: WAb *Pkuazigan* 'bread' (*L84: 29*) Cn/Kb *pk8ésigan* 'Pain (noble [= *AN*])' (*Aub95*)

[8] Contrast this with Baran's (2003) report of significant interspeaker variation in the gender assignment of English loanwords in Polish. Note too that there is ample evidence against the claim (suggested by a reviewer, and occasionally made for Algonquian languages outside this survey) that AN status might be a general treatment for culturally novel items. Passamaquoddy, however, treats as IN a wide range of post-contact foods (cabbage, lettuce, etc.) and other technology (cars, computers, televisions, etc.), and similar examples can be readily found for the other languages discussed here.

Pb *àpɑn* 'bread' *sokálapɑn* 'cake' PsmMl *opan* 'bread' *sukolopan* 'cake'
(cf. *innumerable other baked-good terms: WAb Pata 'A tart, a pie'; PsmMl kolahkossok 'crackers'…, WAb Abônak 'Cakes' (L84: 30); WAb Kalakonak 'Biscuits (sea biscuits[])]' (L84: 30))*
IN: Mw *pipnaqan* 'bread' Mw *ke'k* 'cake' Mw *petaqan* 'pie'

Here it is unclear if Mi'kmaw has innovated away from the WAb-Pb-PsmMl cluster to its southwest, or vice versa: some Algonquian languages also treat this area as AN (e.g. Menominee (Bloomfied 1962: 31–2)), others as IN. Again, the key observation is the consistent diachronic change of a whole semantic class, rather than just an individual lexeme.

The exact same distribution of AN status is also found for coins (and comparable objects):

(11) COIN as AN and IN across Northeastern-area Algonquian
AN: WAb *Mdala sansak* 'Ten cents' (*L84: 46*)
Pb *nəkʷətákiso* 'a silver dollar', *(s)sentak* 'pennies'
PsmMl *kaltolu(hk)*, *kawtolu* '(coin) quarter-dollar',
tensens '(coin) dime; ten cents', *payopsens* 'nickel',
sens 'cent, penny, coin', *sumalkin* 'copper penny; half-penny (Maliseet)'
IN: Mw *tlansu* '25 cents', *galgie, galtie* 'quarter (coin)', *sumalgi* 'cent'

Here the WAb-Pb-PsmMl cluster may be equating coins to representations of people (i.e. the images) or simply akin to gaming pieces, both of which are robust Animate "families". Again, the striking feature is the consistent semantic-class-based treatment. An instance of a precisely reversed distribution (IN for the WAb-Pb-PsmMl cluster, AN for Mi'kmaw) is the case of KNIFE:

(12) KNIFE as IN and AN across Northeastern-area Algonquian
IN: WAb *Nsakuakw* 'A knife' (*L84:28*) Pb *nsèhkʷakʷ* 'knife'
PsmMl *mihqotanis* 'knife' Mw *wa'qan* 'knife, blade'; *wa'qanji'ij* 'jacknife'
AN: Mw *tlawo'q* 'butcher knife; [hunting knife]' Mw *awa'qi'gn* 'crooked knife'

This apparently Mi'kmaw-specific innovation (for substantial/sharp knives only; hence IN *wa'qan*) may be modelled on the more extensive AN 'family' of THORN-like objects. Compare this again to the regional innovation of SHOE as Animate, possibly from the nearly pan-Algonquian treatment of snowshoes as Animate.

Two final examples of the same are LEAF and BELT/SCARF (13). Here note especially the systematic shifts in gender for directly cognate forms like Pb *kspìson* vs Mw *gispisun*:

(13) **LEAF and BELT as IN and AN Across Northeastern-area Algonquian**
a. LEAF
IN: WAb *Wanibagw* 'A leaf' (L84: 32)
Pb *mìpi* 'leaf' PsmMl *mìp* 'leaf'

AN: Mw *ni'pi* 'leaf'
Mw *wi'gatignipgw* 'lettuce'
Mw *sug̲luanipgaji'jit* 'k.o. leaf for medical purposes'

b. BELT/SCARF
IN: WAb *Kwutguabizon* 'A girdle; a belt' (L84: 26)
Pb *k^{w}ətək^{w}ápison* 'belt'
Pb *kspìson* 'waist broad belt, waist sash'
PsmMl *k(o)spisun* 'belt, sash'
WAb *Kchi-moswa* 'A shawl' (*L84: 27*)
Pb *ssal* 'shawl'
PsmMl *wiwonekonosut* 'shawl, wrap . . . '
WAb *Nôpko[w]an* 'The neck tie; [collar]' (*L84: 26*)
Pb *kihkə́sk^{w}epi* 'collar'
Pb *kkə̀sk^{w}epi* 'scarf, kerchief; handkerchief'
PsmMl *wiwonoskopun* 'scarf'
but AN!: PsmMl skahp 'scarf'

AN: Mw *gispisun* 'belt'
Mw *wijipoti* 'money belt (with pocket)'
Mw *sa'l* 'shawl'
Mw *qotaqanigjipilaqan* 'muffler, scarf'
Mw *ugqotaqanigjipilaqan* 'necktie, neck wrapper, collar'
Mw *qotaqanigjipilo'qon* 'necktie'

These two further Mi'kmaw innovations—'leaf' perhaps modelled on the QUILL-FEATHER-FLAKE 'family', BELT/SCARF perhaps tied to the strong tendency in all four languages for cords/binders to be AN—track what one might already guess: Mi'kmaw is quite generally the most historically distinct/divergent of the four surveyed languages. The development of the AN assignment system reflects this clearly.

Again, the crucial observation is that Mi'kmaw's divergences from its neighbours in AN assignment—bearing in mind that most of the Animate 'families' are still robustly shared among all four languages—are not simply random: they are evidently quite systematic, and semantically clustered. Cases of systematic, clustered gender shifts across languages, regions, and even dialects suggest that the determination of the formal AN property applies primarily over semantic groupings within the lexicon rather than solely over individual lexemes.

This observation also suggests a new potential area of investigation: historical-comparative research in Algonquian languages might benefit from comparing not just shared/cognate lexemes, but also shared/'cognate' patterns of Animate assignment (as examined above for four languages), i.e. what we might call 'categorizational isoglosses'.[9]

[9] As Algonquian is a large and well-documented family, this encourages work like Polinsky and Van Everbroeck (2003), which models diachronic reanalysis of gender in a heavily morphological/phonological gender assignment system (i.e. from Latin to Old French)—but now also exploring systems where productive gender seems to be almost exclusively semantically determined.

12.6 Outstanding issues

12.6.1 Falsifiability

In many ways, formal-linguistic understanding of Algonquian gender systems remains still in its infancy, centuries after the earliest attested documentation by missionary grammarians. The de facto near-consensus that Animate status is ultimately unpredictable has, with some notable exceptions, not encouraged in-depth examination of Algonquian gender productivity, or further questioning as to what predictivity and productivity in such systems might actually be.

The present survey raises more questions than it answers. First is the issue of how a 'family'-based model is falsifiable at all—how it can avoid confirmation bias and circular definitions. For example, in the languages surveyed, eggs are not categorized as Animate. But the term for 'nit, louse egg' consistently is (14).

(14) **'nit, louse egg'**

PsmMl	*konasis*	'louse egg, nit'
Mw	i*gna'ji'j*	'nit'
Pb	*nə̀phis*	'nit'

A clever solution comes from noting that nits are quite distinctive among eggs in prototypically occurring glued tightly to human hairs. They are thus comparable to a burr or sticktight: designata that robustly attest as ANs. (The Pb etymon suggests this literally, being based on the Root *nəp-* 'hooked on, attached'.)

But are these clever solutions too easy to find? Can we not tell a story like this for anything? And how can we ever be sure a story like this (or some psychological equivalent) plays any role at all in actual speaker knowledge of their Animate assignment system?

Similarly, having constructed the local map of Animate 'families' in (4) above, how do we know we are not just constructing a very pretty (self-) delusion? Especially for closed data sets (WAb, Pb), we can keep refining the above set further and further until it all 'fits' extremely well. But we do not know if this actually models anything real. Examining Bloomfield's (1962) categorization of Menominee Animates, Goddard observes comparably that "…the categories of 'special gender' include (or could include) just about everything except abstract nouns and maybe structures…" (Goddard 2002: 200). How, then, do we falsify claims for particular 'family' categories, and for each instance of membership therein? What prevents us from arranging and re-arranging till everything fits (ibid.: 207)—particularly when we have so many cases of putative 'families' where our statistical *n* is (unavoidably) very small?

So far, we have no comprehensive answer. But the productivity effects reported in Section 12.4 (alongside the similar diachronic effects seen in Section 12.5) still strongly suggest that a coherent categorization process is in fact occurring. The inherent challenge for straightforward falsification should not drive us away from this phenomenon. It should push us to look closer, to explore what analytical/experimental strategies might ultimately capture valid results.

Some hope lies in the model's predictive power for still-spoken languages (PsmMl, Mw). We may never get 100% coverage, but a model with 85% (or higher) accuracy for novel data is, at worst, a useful pedagogical tool. And at best, it might at least be a Newtonian physics model (vs a quantum physics one): still not capturing exactly what is going on, but refining the observations/generalizations enough to make the next breakthrough possible.

12.6.2 Modelling the semantics of Animate assignment

An equally pressing question is how to formally model the semantic and syntactic components of this system. A comprehensive model for either is beyond the scope of this chapter, so we instead focus on identifying key issues any such model will have to address and account for.

As we have seen, Animate status assignment seems clearly grounded in the semantics of the designatum, forming a system that is never perfectly predictable—but also never blankly *un*predictable: since both pre-existing and totally novel nouns (or nominal designata) overall quite reliably track the established 'families'.

This particular combination of the elusive and the clear suggests that it may be most effective to model the Animate-assignment system as an emergent phenomenon, one developing in part out of learning strategies that attend to at least two factors within the ongoing process of acquisition: (a) exemplar data from the speech community, and (b) individual-level attempts to make immediate and ongoing sense of this input. This latter cognitive component is restrictive and organizational, following approximately this process:

(15) a. Simply accept the bulk of of Animate assignments as they are encountered.
b. From these, create a post-hoc set of (radial-categorizational) 'families' based on designata semantics.
c. As the emergent patterns in (b) are reinforced (and/or revised), productively assign Animate in accordance with those patterns.

This path results in exactly what we see: a system that can accept initial (and ongoing) opaque gender assignments, even as it can also productively and consistently assign gender to new lexical items, including ones in/from other languages. (The constrained fuzziness inherent to such a system would also feed precisely the kinds of synchronic variation and diachronic change we observe.)

This aspect of the model being so inherently open-ended, it is crucial to ask what, if anything, constrains this emergent categorizational process. Here it is helpful to reiterate that the common property of AN assignment systems is not a single feature shared among all ANs, but instead, the *limitations* on what designata properties the analogical system can attend to.

In principle, any analogical relation/semantic feature could support a 'family' relation. But in empirical observation, the process looks more constrained. We never see Animate status tracking colour, material composition, weight, or shape. Hence we never need to make statements like "*One Animate designatum is (prototypically) {purple, wooden}… so all {purple, wooden} objects are Animate*".

As Quinn (2001) notes, we do however find Animate 'families' often tracking *telos*, i.e. what the object is definitionally *for* (e.g. *fluid container* {pot, cup, lung}), and/or a narrow range of certain kinesthetic properties (e.g. *sticky blob* {gum, clay, clot}; *poking spike* {thorn, quill, harpoon}).[10] Beyond that, however, this system ignores an enormous range of analogical possibilities, and is evidently sensitive to only a certain few.

Observing this narrow range, we might even preliminarily speculate (and we emphasize that this is purely speculation) that Animate-assigning analogies may be about salient (and definitional) interactional properties of a designatum: i.e. its inherent attraction to the attentional 'foreground'. (Note the affinity of this view to Goddard 2002's 'special/unusual case of' criterion.)

This model would begin to explain why the following specific 'families' are especially diachronically and synchronically robust:

(16) a. semantic animates—but systematically generally *not* plants—and representations thereof
b. things that stick in you (thorns, burrs), on you (gum), or that you get stuck in (nets)
c. gaming pieces, balls (tracked attentionally similarly to prototypical semantic animates)
d. fluid containers (prototypical handling requires intense attentional tracking to avoid spilling)

Namely, these are precisely the sorts of designata that reliably demand foregrounded attention from among the overall perceptual 'background' of inanimate objects.

This, then, is not a new kind of 'one common thread' analysis, but instead is an attempt to establish some cognitive limits on what *range* of properties might be relevant to Animate status rather than a single determining property.

We observe at least two more semantic constraints on Animates. They must be concrete entities, never abstracts: contrast this with Indo-European gender, which cross-cuts the concrete-/abstract-noun distinction. And semantic animates, excepting plants (which are precisely the least saliently animate, and most perceptually 'background' living things) must always be grammatical Animates. These two points underline the view that this system, while clearly having a strong formal component, continues to be very heavily grounded in some form of core conceptual semantics.

12.6.3 Modelling the syntax of Animate assignment

The observations here also support a particular range of models of the formal-syntactic aspect of the system. Specifically, the phenomena reported here fit well with current models placing AN gender on a light element *n*, that nominalizes a category-neutral Root or Root complex (Ferrari 2005; Acquaviva 2009; Kramer 2009,

[10] We use '*telos*' here as a clarificatory relabelling of what Quinn (2001) originally calls 'intrinsic function'. In particular, this property may be usefully connectable to the TELIC quale component of Pustejovsky's (1995: 76–7) Generative Lexicon model.

2015; Acquaviva, this volume; but see also Déchaine, this volume, for an alternative model with a much higher syntactic locus for AN, based on data from Plains Cree, a Central Algonquian language). Overall, the data seem to demand an element (or process) that

(17) a. *allows one single stem to potentially lexicalize as either gender (very often with a different designatum):*
As in the dual animacy effects noted earlier.
b. *establishes a configuration that functions as a local maximal domain for lexicalized meaning:*
As per (a), the [AN/IN+stem] constituent *as a whole* is the seat of specific/idiomatic meaning.
c. *imposes (like English mass-count) a consistent semantic restriction on possible designata of each:*
As per robust semantic restrictions on what designata can/cannot be Animate.
d. *provides an interface component to the formal syntax:*
Massive attention is paid by morphosyntax to this property of nominals: both agreement/concord-type effects, and also selectivity of stems themselves—so we could reasonably claim that Algonquian morphosyntax never interacts with a noun but through its gender.

To capture (17a–c) in particular, effective formal-syntactic models of Animate assignment will likely be those that allow a gender element/feature to be separated from the lexical stem itself, while still forming a tightly local semantic constituent with it. (And while not a part of this survey, the role that the Animate vs Inanimate contrast may play in individual/count vs. collective/mass contrasts, may also be an important part of a complete syntactic model.)[11]

12.7 Conclusion

With these questions of falsifiability and semantic and syntactic modelling still very much open, it is clear that Algonquian animacy systems call out for far more substantial investigation. Again, 'unpredictable' Animates have historically seen little in-depth empirical research, having often been lumped into the familiar phenomenon of more extensively arbitrary formal gender. (It is unsurprising that there has not been much work on the productivity of patterns still widely assumed to be unpredictable.)

[11] Specifically, Goddard (2002: 212–14) gives extensive Meskwaki examples contrasting Animate individuals vs Inanimate collectives. A distinct claim of use of Animate count vs Inanimate mass in Ojibwe is made by Mathieu (2012a, 2012b). Adding to this two further points—Fraurud's (1996) ontological model in which semantic animates are the prototypes for individual entities overall, and Kraaikamp's (2012) observation that Dutch masculine/common (≈ human/animate) gender assignment is associated with a high degree of individuation, and neuter (≈ non-human/inanimate) with a low one—we see again that set of syntactic and semantic phenomena associated with Animates circles around inherent or prototypical prominence/foregrounding/Figure-hood.

A close look at available data, however, suggests that to even begin to properly understand these systems, we still need at least the following:

(18) a. A rigorous process of proposing, testing, and refining identifications of language-specific AN 'families' (and the links between them, language-internally and comparatively).
b. A clear/better way to counter the confirmation bias inherent to the above methodology.
c. Formal experimental methods application, e.g. systematically testing
– how speakers handle novel designata, and/or loanwords.
– where the precise margins of robustly identified categories are, using concrete designata.
d. Solid surveys of this sort across as many Algonquian languages as possible, especially given that most of these systems are threatened (and especially since Animate-assignment systems may possibly shift with richness/degree of lexemic exposure).

Of these, point (18c) is perhaps the most urgent. This chapter itself demonstrates that we so far have only indirect evidence (lexicographic and initial/anecdotal field observations), but no direct experimental data on Animate-assignment productivity for native speakers of *any* Algonquian language—even as the main windows of opportunity for such work are closing rapidly.

We might finally add once more that this model/understanding could contribute well not just to cross-linguistic theoretical analysis of nominal gender, but also to current-day practical revitalization efforts.

Heretofore, Algonquian gender has been commonly presented as a largely arbitrary, brute-force-memorized system. Recognizing that there is in fact much more predictability than unpredictability radically facilitates learning both correct lexemic gender and the complex morpho-syntactic phenomena that build off of it, and demystifies what is generally experienced as a baffling and intimidating obstacle for would-be new speakers. Nearly all Algonquian languages are severely threatened, so this alternative framing of their gender systems could significantly help beginning learners find their feet.