Productivity vs. predictability: evidence for the syntax and semantics of Animate gender in four Northeastern-area Algonquian languages

Conor McDonough Quinn
University of Maine
conor.mcdonoughquinn@maine.edu

Gender, Class, and Determination | University of Ottawa, 19 Sept. 2015

- Finding this presentation....

http://www.conormquinn.com/Quinn2015ProductivityVsPredictability0919draft.pdf

or just navigate via

www.conormquinn.com

to Downloads...

1. Background

- *Animate* vs. *Inanimate*: a nominal gender distinction in Algonquian languages, realized as selectional and/or agreement effects in key nominal-argument tracking morphology:

nominal-side:	[demonstratives], numerals, plurals	*[+ "adjectives"]*
verbal-side:	"Finals" (= affixal verbs), verbal arg-indexing morphology	

(1) Animate vs. Inanimate (*Pb: Siebert 1996/PD, glosses CQ*)

Animate		**Inanimate**	
owa phenəm	'this woman'	*iyo pənahpəsk*^{w}	'this rock'
(i)yok phenəmok	'these women'	*(i)yòlil pənahpskol*	'these rocks'
wəlíhpokəso	'Animate tastes good'	*wəlíhpokat*	'Inanimate tastes good'
nàmohα	'I eat h/her/it (Animate)'	*nəmìčin*	'I eat it (Inanimate)'

- *Abbreviations:*

AN = Animate
IN = Inanimate

{AN, IN} verb used here for a verb/Final selecting {AN, IN} absolutive argument...
....for non-Algonquianist convenience/clarity

Question: Is a (fully) predictive account for Animate/Inanimate status possible?

Answer: **No. (Or so the literature still generally says, with a few exceptions.)**

Question: **Why not?**

Answer: **Because AN/IN appears to originate from semantic animacy, but shows all the signs of having formalized away from it:**

(a) Many grammatical Animates fall well outside of the core domain of familiar semantic animates (*all Penobscot*):

èmkʷαn	'spoon'
màkikʷ	'nasal mucus, snot, catarrh'
məlǝčess	'mitten'
kɑ̀wi	'porcupine quill'
tálαkan	'wedge'

(b) Most plants are Inanimate...despite being describable with terms translating as 'live', 'grow', and 'die':

pemawsuwikil <pĕm´-au-sŭ-wi´-kil> 'Living things' / 'Trees and plants' (*PsmMl: Chmb1899:39*)

cf. *pomawsuwiw* IN verb 'it is alive'

Psi-te keq hokek Laks nukcoktehson; tehpu wawikon <u>pemawsuwik</u>.

Every part of Laks's body is smashed; only his backbone [IN!] is still <u>alive</u>. (LM)'

pomikon IN verb 'it is growing, keeps growing'

mehcinewiw IN verb 'it dies, it is dead'

(Note that (b) challenges the frequent speculation (never confirmed by native speakers, but having wide currency especially in second-language teaching) that (a)-type words are somehow conceived of as "alive" in some culture-specific way. See Dalhstrom 1995:65, Goddard 2002:225 for more challenges to this claim.)

→ Animate is generally agreed to be *marked* category, with Inanimate as the *elsewhere*. So, what determines "unpredictable" Animates?

→ Discussions of Animate gender assignment standardly recognize areas of semantic generalizability, but emphasize an ultimately arbitrary, unpredictable character:

- *"...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 verious 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 1994:205, re Myaamia)*

- *"[Re the idea] that there are some semantic correlates in Alg lgs betwee 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 & Moss 2008:53, re Arapaho)*

Common descriptive problem: positing a (straw-man) category, and then complaining that Animate membership is inconsistent within it.

[*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!"

(cp. Dahlstrom 1995:56-57 re [*"power"*])

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.

(*We'll see that all of the above Animates are not nearly as random as they might first appear.*)

→ At least three accounts push 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, re Meskwaki)

- **Today's proposal:** we acknowledge the contrastive function of AN vs. IN, but note that AN membership seems far 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 a radial categories approach.

- Dahlstrom 1995 presumes [central vs. peripheral] membership. Could just be ongoing process of analogical attraction---creating *centerless clusters* rather than core-vs.-periphery systems---limited by not by protoypes but just by definable constraints on *how* that analogical process may proceed.

(Cf. 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.")

- New here: a renewed focus on how this process results in a very structured *productivity* of AN assignment (esp. for loans, etc.; cp. D1995:61-62). I.e. we extend the "*once you know them, they make sense*" observation in a key direction: "*once they make sense, you can make more of them*".

Overall picture: synchronic competence in AN assignment is *emergent*: not the setting of a single rule, but a structure (constantly) built up.

2. Today's interesting news...

- Exhaustive survey of current lexicographic documentation + preliminary field research, covering four Northeastern-area Eastern Algonquian languages:

Passamaquoddy-Maliseet	[field notes (Sipayik; Sakomawi Malihk), *pmportal.org* online dictionary]
Mi'gmaq	[field notes, unpublished ms. dictionary, + published dictionaries]
Penobscot	[Siebert/Speck field notes, in-process *Penobscot Dictionary* digital ms.]
Western Abenaki	[Laurent 1884, Masta 1932, Day 1994, inter alia]

- 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.

- For the two languages for which in-depth native-speaker consultation is still possible (Passamaquoddy-Maliseet and Mi'gmaq), preliminary data suggests 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 (= English) word designata in general.

- Combined with observations from extant Penobscot and Western Abenaki corpora data, we conclude that this productivity indeed 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" (as per esp. Dahlstrom 1995 and Quinn 2001, cp. also Wittgenstein 1953) that systematically (and quite restrictedly) attracts semantically related new members.

→ The still-standard view of Algonquian gender as "ultimately/largely arbitrary" must be abandoned.

Otherwise we will miss a crucial chance to investigate potentially very enlightening questions about how categorization (linguistic or otherwise) works as a cognitive process.

3. The "family"-based model

For all four languages surveyed, the following "family" categories of Animates are extremely robust, being *nearly exceptionless*.

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}
shirt-jacket/coat{}
[wheels-disks]
snowshoe-[footwear]-handwear
breechcloth/diaper

- "Families" marked in [*brackets*] are ones *systematically* absent in at least one language surveyed; to be discussed below.

- Acquirers of Alg systems may set up a core prototype for each group (= Dahlstrom), or simply accrete new ANs into emergent groupings based on shared semantic features of their designata. (See §9 for possible constraints on what kinds of features can(not) trigger accretion.)

- These groupings are NOT uniform across all Algonquian languages, though there does appear to more overlap than not. *(For comparable clustering observations, see esp. Bloomfield 1962:28-36 for Menominee, and Goddard 2002 for Meskwaki and Algonquian in general.)*

-"Families" come and go: note a robust AN category totally absent in the 4 surveyed, but found in another Alg lg, Munsee (O'Meara 1996: 405, 401, 538; 35, 414, 414):

[*books-paper-documents*]	*pámbiil* 'book, paper, letter', *mbáypul* 'Bible', *noospépul* 'newspaper'
[*vehicles*]	*amóoxol* (NA/NI) 'boat', *ahtamóombiil* 'car', *káal* 'train car'

- Examining individual exceptions: superficiality of unexpected IN status for Pb/PsmMl 'potato', and Mq 'apple'

(a) 'potato' *IN* in Pb *ápəčetes* and PsmMl *početes*
vs. *AN* for Mq *tap'tan*, WAb *Padatesak* (L84:30) (*+ root vegetables in general*)

- Pb literal expression directly from an *IN verb*:

àpəte 'it (weather, any INAN object) is hot, warm'

+ Medial (nominal incorporant) *-əče-* '(small) round object (?)'
(cp. *-əča-* in WAb *mk^{w}əčasikan* <Mkuejazigau> [sic] (L84:59) 'Roast meat')
+ nominalizing diminutive *-s*. (*PsmMl form directly cognate.*)

[Cp. *àpəte* w Medial *-ahpsk-* 'rock; round obj': *apáhpskəte* 'it is a hot rock, it is hot (flat cooking rock); it is a hot stove'; IN 'stove' [lit. "hot round object"]

In short: across Alg lgs, the 'potato' designatum is typically AN. The exceptional IN case seen here is precisely one that names the designatum via an explicitly IN-verb-based metaphorical framing.

(b) 'apple' *IN* in Mq *wen'ju'su'n* vs. *AN* for Pb *ččìkəne*, PsmMl *cikòn*, WAb *Aples* (L84:31)

- Mq literal expression is (a description) based directly on an *IN noun*:

su'n 'cranberry' (= expected+attested as IN across all 4 lgs)
+ prenominal modifier *wen'ju'-* 'French; foreign'.

- Using literal IN-based constructions as naming strategies can overrule an AN status that would otherwise be expected (and that is attested whenever the lexeme doesn't involve this factor).

- *Superficial counter-example:* Pb personal name *αpátehtek* 'The Leaner' is based on IN verb stem, uses IN argument marker *-k*. But in textual attestation, it is otherwise treated entirely as a grammatical AN (i.e. verbal marking, participation in obviation system, etc.)

- *Suggestion*: explicit personification can coerce AN treatment---as independently reported across Algonquian for IN nouns when personified, as often happens in mythological texts (Wolfart 1996:399, Dahlstrom 1995:57-58, Goddard 2002:202-210, C&M2008:52-53).

Upshot: Animacy assignment involves multiple competing factors.

- Personification can coerce AN status, while explicit IN-lexeme-based framing of non-person designata can block expected AN status.

Key: Gender in (a, b) is not directly predictable from the designatum---but it is still predictable from the literal designatum of the actual wordform being acquired.

4. Evidence for "family" effects: dual animacy and variable animacy

- Dual animacy: meaning differences in AN vs. IN use of the same stem track established "families"

Pb *wətəhɑkan* AN: 'fin" (cf. AN *wələkʷan* 'wing': FEATHER/FAN) IN: 'paddle, oar'

Pb *pəko* AN: 'chewing gum' (cf. *mɑkikʷ* 'nasal mucus, snot, catarrh': GUM)
IN: 'gum, pitch (in sap form, or when used to pitch a canoe)'
(cp. *kələmótikan* 'adhesive, glue, mucilage')

Pb *apeskʷ[h]ám[]ɑkan* AN: 'lacrosse [*or generic*] ball: BALL' IN: 'lacrosse game'

Pb *kəlósəwɑkan* AN: 'speech wampum; large, long belt of wampum, as used for intertribal treaties' (GLYPH/CORD)
IN: 'a word, the word; speech, talk'

(cp. Costa 1994:208: AN *ahkihkwa* 'drum' IN *ahkihkwi* 'kettle'; AN *mihtekamina* 'June bug' IN *mihtekamini* 'acorn')

- Variable animacy: AN/IN variation for the same apparent lexeme tracks the margins/fuzzy edges of established "families" (cp. Costa 1994:208)

Pb *mskíhkəwimin* 'strawberry': generally *AN*, but [*IN pl.*] *mskihkəwíminal* for 1 spkr (SDMC:998)
cp. also one case of [*IN pl.*] dim. *mskihkəwíminsal* (PD)

- Speculation: in Pb, smaller berries (e.g. blueberries) are uniformly IN; larger, squish-prone berries (and soft-peeled fruits) are uniformly AN. Strawberries generally predicted to be AN (as they are): but the much smaller wild berry may fall below that threshold (hence esp. the IN dim. attestation).

- Variability around the margins of robust "families" is expected; we predict to occur only for "unstable designata" of this kind.

- (Some cases of variable animacy might just cases of dual animacy where documentation has missed the precise designatum difference.)

- Dual animacy shows suggestive similarities to the English *mass* vs. *count* system.

- English *mass/count* is [*productive, largely predictable, meaning-constraining*], but still tied closely to *idiosyncrasy in lexicalization*:

COUNT	*(a) speech* (= lecture!)	vs.	MASS	*speech*
COUNT	*(a) drive* (= golf!)	vs.	MASS	*drive* (= motivation!)

[*she's got a drive ≠ she's got drive*]
[*she's got a game ≠ she's got game*]

- 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. *(Cp. aspect stability in verbal idioms)*

5. Evidence for dynamic synchronic productivity: Passamaquoddy-Maliseet and Mi'gmaq

- Only some preliminary/anecdotal---but quite suggestive---data re Animate status of English words (and/or newer or unfamiliar designata).

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? (*Psm: 200605_-ML*)
[Before CQ can even answer ("Dates"), speaker 1 starts discussing w/speaker 2,
referring to them using <u>*AN*</u> morphology.]

c. Q: [*re plurals*] What about "dates"? A: [*w/ no hesitation; <u>AN</u> pl.*] Dates-*ig*. (*Mq: 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 for me? (*Ml: 20150729-DB*)

A: Kespahtekil [= *IN dried ones*]...[*seeing that they're apricots*]...
....no, I should say "*kespahsicik*" [= *AN dried ones*]

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 CQ-offered "*peaches*-ək".]

Upshot: Formal experiments will be needed to take this beyond mere anecdote, but speakers of PsmMl and Mq do seem to show a robust and largely consistent knowledge of the gender assignment of novel items---designata for which they know no pre-existing word---and of recent and/or foreign (mainly English and/or French) word designata in general.

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

- "Family"-based model accounts well for why diachronic change in gender assignment across Algonquian proceeds not simply on an individual lexeme-by-lexeme basis, but by *semantic cluster*.

- SHOE (etc.) is generally IN across Algonquian, remains IN in WAb and Caniba/Kennebec (E. Abenaki)...but from Penobscot east/northward, it's consistently AN:

IN: WAb *Mkezenal* [= *IN* pl.] 'Shoes; moccasins' (L84:26; cp. *Potsal* 'Boots')
Cn/Kb *ne makesenar* [= *IN* pl.] 'mes souliers' (Aub:482) *+ most other Algonquian....*

AN: Pb *màhksən* 'shoe' PsmMl *pkoson* 'shoe' Mq *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?

- BREAD and comparable products are consistently AN in WAb, Pb, PsmMl, and consistently IN in Mq:

AN: WAb *Pkuazigan* 'bread' (L84:29) Cn/Kb *pk8ésigan* 'Pain (noble [= *AN*])' (Aub95)

Pb *àpɑn* 'bread' *sokálapɑn* 'cake' PsmMl *opan* 'bread' *sukolopan* 'cake'

+ 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: Mq *pipnaqan* 'bread' Mq *ke'k* 'cake' Mq *petaqan* 'pie'

It's unclear here if Mq has innovated away from the WAb-Pb-PsmMl cluster, or vice versa: some Algonquian lgs go for AN (e.g. Menominee), others IN.

- COIN is consistently AN in WAb, Pb, PsmMl, and consistently IN in Mq:

AN: WAb *Mdala sansak* 'Ten cents' (L84:46) Pb *nəkʷətáḱiso* '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 (Ma)'

IN: Mq *tlansu* '25 cents', *galgie, galtie* 'quarter (coin)', *sumalgi* 'cent'

WAb, Pb, PsmMl may be equating coins to representations of people (i.e. the images) or akin to gaming pieces?

- KNIFE is generally not AN---though *THORN*-like items (harpoons) typically are---but it is in Mq:

IN: WAb *Nsakuakw* 'A knife' (L84:28) Pb *nsèhkʷakʷ* 'knife'
PsmMl *mihqotanis* 'knife' + Mq *wa'qan* 'knife, blade'; *wa'qanji'ij* 'jacknife'

AN: Mq *tlawo'q* 'butcher knife; [hunting knife]' Mq *awa'qi'gn* 'crooked knife'

Mq innovation perhaps modeled on AN thorn/spike "family"? (Cp. again SHOE innovation possibly from common AN 'snowshoe'?)

- LEAF is generally not AN, but is in Mq:

IN: WAb *Wanibagw* 'A leaf' (L84:32) Pb *mìpi* 'leaf' PsmMl *mìp* 'leaf'

AN: Mq *ni'pi* 'leaf' Mq *wi'gatignipgw* 'lettuce' Mq s*ugtluanipgaji'jit* 'k.o. leaf for medical purposes'

Mq 'leaf' AN innovation perhaps modeled on AN quill/feather/flake "family"?

- BELT/SCARF is generally not AN, but is systematically so in Mq:

IN: WAb *Kwutguabizon* 'A girdle ; a belt' (L84:26) Pb *kʷətəkʷápison* '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 *kihkə́skʷepi* 'collar'
Pb *kkə̀skʷepi* 'scarf, kerchief; handkerchief' PsmMl *wiwonoskopun* 'scarf'
but AN!: *PsmMl <u>skahp</u> 'scarf'*

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

Another (mainly!) Mq innovation, perhaps tied to the strong tendency in all 4 lgs for cords/binders to be AN.

- As one might guess, Mq is quite generally the most historically distinct/divergent of the 4 surveyed lgs, and the development of the AN assignment system reflects this clearly.

- The crucial observation is Mq's divergences from its neighbors in AN assignment---bearing in mind that most of the "families" ARE still robustly the same---are not simply random: they are evidently quite systematic, and semantically clustered.

- Such cases of clustered gender shifts across languages, regions, and even dialects suggests that the determination of the formal Animate property applies primarily over semantic groupings within the lexicon, rather than solely over individual lexemes.

- Historical-comparative research in Algonquian languages would therefore benefit from comparing not just shared/cognate stems, but also shared/"cognate" Animate assignment, i.e. "categorizational isoglosses".

7. The implacable methodological problem: clever explanations vs. built-in confirmation bias.

In the languages surveyed, eggs are not categorized as Animate. But the term for 'nit, louse egg' is:

PsmMl	*konasis*	'louse egg, nit'
Mq	*gna'ji'j*	'nit'
Pb	*nɑ̀phis*	'nit'

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

- But are these clever solutions too easy to find? Can't we 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 in at all in actual speaker knowledge of their Animate assignment system?

- Similarly, having constructed a map of the "families" like this one...

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}
shirt-jacket/coat{}
[wheels-disks]
snowshoe-[footwear]-handwear
breechcloth/diaper

...how do we know we're not just constructing a really pretty (self-) delusion? *We basically don't.*

- Esp. for closed data sets (WAb, Pb), we can keep refining the above set further and further until it all "fits" extremely well. But we don't know if that actually models anything real.

"Note, however, that the categories of "special gender" include (or could include) just about everything except abstract nouns and maybe structures...." (Goddard 2002:200, re B1962's categorization of Menominee ANs)

How 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 (Goddard 2002:207)?

Particularly when we have so many cases where our statistical n is (unavoidably) very small?

So far: no good answer. But the productivity effects still strongly suggest that a coherent categorization process is in fact occurring.

- The inherent challenge for straightforward falsification shouldn't drive us away from this phenomenon. It should push us to look closer, to see what analytical/experimental strategies might ultimately capture valid results.

- Some hope lies in the predictive power for still-spoken languages (PsmMl, Mq). 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 Newtonian physics model (vs. a quantum physics one): still not capturing exactly what's going on, but refining the observations/generalizations enough to make the next breakthrough possible.

8. Towards a formal model of the semantic component

So: a system neither perfectly predictable, nor all that *un*predictable. Since both pre-existing and totally novel nouns (or nominal designata) overall quite reliably track the established "families".

- AN status emerges from at least two factors within the (ongoing) acquisition process: (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:

(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 consistntly assign gender to new lexical items, including ones in/from other languages. *(W/fuzziness feeding inherent variation and change.)*

9. Towards a workably restrictive model of the semantic component

- What, if anything, constrains this emergent categorizational process? 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.

- ANY analogical relation could support a "family" relation. But process looks more constrained:

NEVER: color, material composition, weight, shape...
"one Animate designatum is (prototypically) {purple, wooden}...so all {purple, wooden} objects are Animate"

INSTEAD: *telos* = what is the object definitionally FOR = fluid container {pot, cup, lung}
some kinesthetic properties = stickyblob {gum, clay, clot}; pokingspike {thorn, quill, harpoon})

- System ignores a lot of analogical possibilities...and sensitive to only a certain few.

- Speculation: may be about salient (and definitional) interactional properties of a designatum, i.e. its inherent attraction to the attentional "foreground" (cf. Goddard 2002's "special/unusual"?).

- Hence these specific "families" are especially diachronically and synchronically robust:

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 (tracked attentionally similarly to prototypical semantic animates)

- Not a "common thread" analysis: just seeking cognitive limits on what range of properties might be relevant to Animate state, rather than a single determining property.

(Cp. "syntactically relevant semantics" in functional elements overall)

- More semantic constraints:

a. ANs must be concrete entities, never abstract. *(Cp. I-E formal gender flexibility here.)*
b. Semantic animates must be AN. *(Except most plants = least saliently animate, most perceptually "background" living things)*

10. Towards a formal model of the syntactic component

No decisive arguments here...but data fits well with AN gender on *n*, nominalizing a category-neutral Root (Ferrari 2005, Kramer 2009+, Acquaviva 2009, etc.). Data seem to demand an element (or process) that

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/can't be Animate.

d. *provides an interface component to the formal syntax:*
Massive attention paid by morphosyntax to this property of nominals: both agreement/concord-type effects, and also selectivity of stems themselves---we could reasonably claim that Algonquian morphosyntax never interacts with a noun but through its gender.

11. What we still need...

Historically, "unpredictable" ANs have seen little in-depth empirical research, having been lumped into the familiar phenomenon of arbitrary formal gender. A close look at available data, however, suggests that to properly understand these systems, we still need at least the following:

a. A tenacious 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 applied, 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 Alg lgs as possible, esp. given that most of these systems are threatened (and esp. since AN-system shift with degree of exposure to it).

Final note: Pedagogical/revitalization significance

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 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 morphosyntactic 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 learnes find their feet.

http://www.conormquinn.com/Quinn2015ProductivityVsPredictability0919draft.pdf

and/or

www.conormquinn.com

References *TBA*