Vocabulary and semantics

Sketch Acquisition Workshop

January 2025

What's in a word?

- Word-like units are the child's initial target for verbal communication.
  - Children on average begin to produce words at 12 months.
  - You will see clear change across the span of the sketch corpus.

- But what counts as a word in a language is going to differ across different typologies.
  - Typology of language means the lexicon has different relationships to other domains of language.

- Luckily, children's initial productions will be fairly simple (though may be hard to decipher), with predicted development across time.

Vocabulary and semantics

Box 9. Key areas of focus in child language: Lexicon and semantics.

Core	
(i)	What is the distribution of word classes and semantic fields across development?
(ii)	What is the relative distribution of nouns to verbs across development?
(iii)	When do children begin to use other word classes?
(iv)	Do you observe any patterns in the truncation of words? Which ones and at which ages? (See Section 6.1 on word structure)
Extension	
(v)	What additional observations can be made about children's semantic development, given the language and the data?

Word classes and semantic fields

- Children's vocabulary won't be so extensive during the recording sessions; coding them for syntactic class and semantic field gives some general developmental information.

(7) Inuktitut:

Animals: *aiviq* 'walrus', *umimmak* 'muskox', *uviluq* 'mussel', *tuktuk* 'caribou'

Vehicles: *anartauti* 'septic truck', *haanta* 'all-terrain vehicle', *qajariaq* 'canoe'

Food/Drink: *panirtitaq* 'bannock', *puijiviniq* 'seal meat', *misiraq* 'whale oil'

Clothing: *amauti* 'parka for carrying baby', *atigi* 'parka', *pualuq* 'mitten'

Baby Words: *aataaq* 'hurt', *amaama* 'bottle, suckle', *ammu* 'sleep', *vuvu* 'vehicle'

People: *ajak* 'maternal aunt', *akkak* 'paternal uncle', *najak* 'sister of boy'

Locations: *avani* 'there', *maani* 'here', *paani* 'up there', *kanani* 'down here'

- Can start with the semantic fields from the MB-CDI.
- But there will be important culturally-specific terms.
- Some fields will differ in size based on cultural considerations (e.g., kin).

Word classes

- Distribution of syntactic categories in CDL and in children's speech has long been of interest to researchers.
  - The difference between children's linguistic behaviour and their input is revealing about constraints on learning.

- One common topic in this space has been so-called ***noun-bias*** in acquisition.
  - Robust evidence in many languages (Dutch, English, French, Hebrew, Italian, Korean, & Spanish (Bornstein et al., 2004).
  - Less robust evidence in other languages (e.g., Tseltal, Casillas et al., 2024).

Lee & Allen (2023) *Lang. Doc. Con.* (Inuktitut Sketch)

Table 31. Average proportion of word class in children's speech.

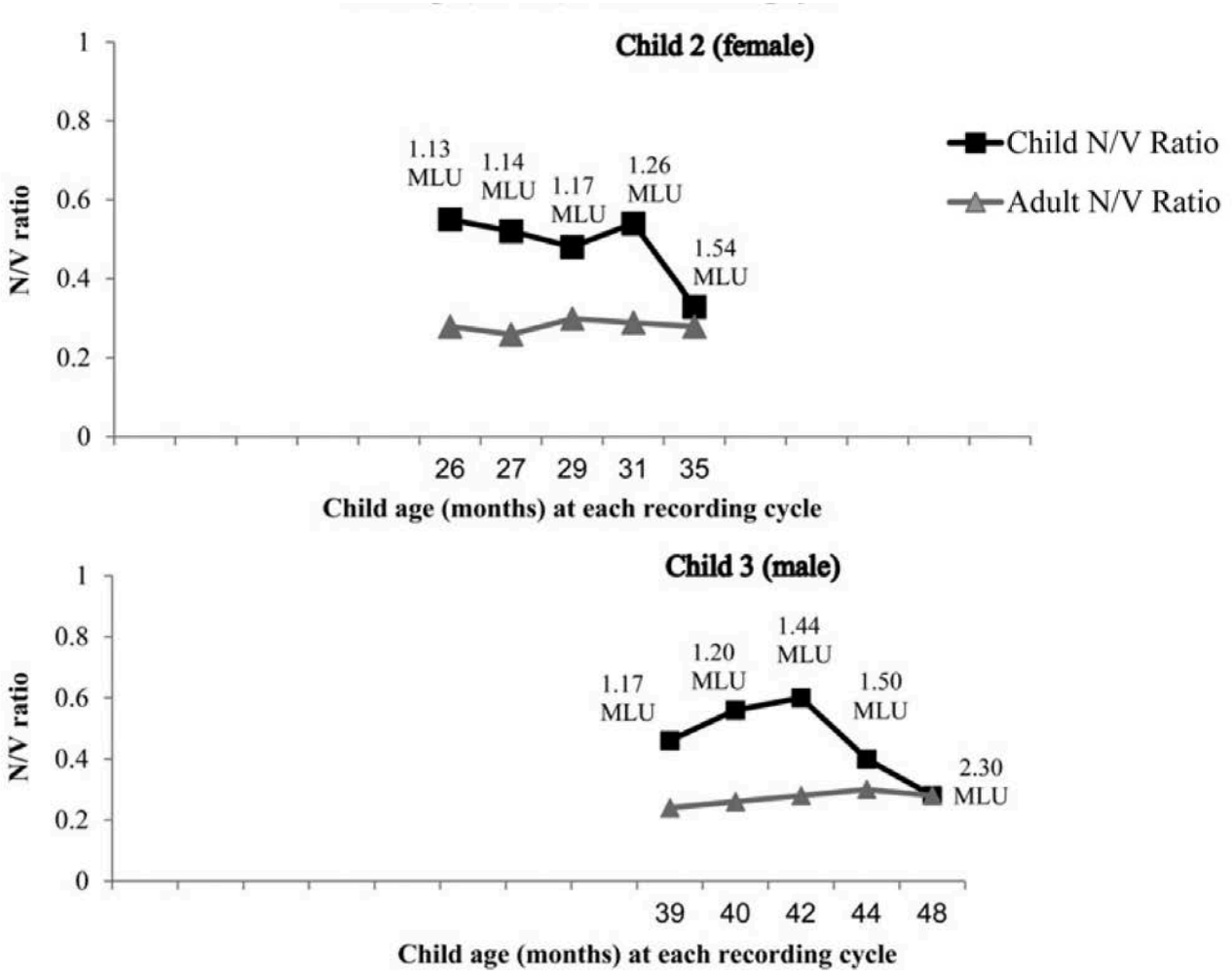
Group	Types				Tokens			
	VR	NR	LR	IACT	VR	NR	LR	IACT
1;4	0.23	0.40	0.03	0.34	0.26	0.62	0.01	0.39
1;10	0.08	0.42	0.07	0.42	0.08	0.55	0.16	0.32
2;4	0.34	0.44	0.04	0.18	0.20	0.51	0.08	0.21
2;10	0.37	0.46	0.03	0.14	0.24	0.51	0.04	0.20
3;4	0.44	0.34	0.04	0.18	0.38	0.36	0.06	0.20

The number of word types at each age is as follows: 1;4=25, 1;10=40, 2;4=75, 2;10=165, 3;4=140. The number or word tokens at each age is as follows: 1;4=67, 1;10=184, 2;4=272, 2;10=466, 3;4=363.

N roots > V roots

Taverna & Waxman (2020) *Jn. Ch. Lang.* (Wichi)

- 3 children recorded 5 times across 12-month period.
- Noun-Verb ratio initially favours nouns but becomes more even around age 3+ years.
- Linked to MLUm – once children had MLU => 1.5, children begin adding more verb stems, verb affixes, and noun affixes to their vocabulary.
- Early nouns lay foundation for acquisition of verbs and other relational terms (Gleitman, 1990).



Word truncation

- Word truncation will be common, particularly in morphologically complex languages.

Agglutinative languages children use isolated morphemes

(5) Inuktitut: Child (Jini 1;4) and her sister

Sister: *maa-li-ruk.*
kiss-POL-IMP.2SG.SBJ>3SG.OBJ
'Kiss her.'

Child: *maa.*
kiss
'Kiss.'

Sister: *maa-pait.*
kiss-PAR.2SG.SBJ>3SG.OBJ
'You kissed her.'

Child: *maa.*
kiss
'Kiss.'

Sister: *maa-tau-laur-langa.*
kiss-PASS-POL-IMP.1SG.SBJ
'Kiss me.' [lit: let me be kissed]

In more fusional languages ch. may isolate syllables across morphemes boundaries

(6) K'iche' Mayan (Pye 1983: 587): A1 Tiya:n (2;2)

Adult: *la: utz kawiloh*
'Do you like it?'

Child: *jah, loh (= jah, kinwiloh)*
'Huh? I like it.'

Adult: *kawarik*
'He's sleeping.'

Child: *lik (= kawarik)*
'He's sleeping.'

Lexical concordance (Pye, 2021, *JCL*)

- Pye (2021) suggests lexical concordance sets (below: Northern Pame), see: http://pyersqr.org/minimal/

(4) A lexical concordance for a two-year-old Northern Pame speaker

Word	Begin Time	Child's Utterance	Adult Target	Translation
daʔtsəlʔ	03:54.9	kiʔil	daʔtsəlʔ	It bit.
daʔuaɲ	20:56.8	ane	daʔuaɲ	It threw it.
danɜ̀s	12:44.2	nas	danɜ̀s	orange
dapaj	05:15.5	paj	dapaj	tomato
	16:26.2	paj	dapaj	tomato
	16:30.6	paj	dapaj	tomato
	07:21.6	pa	dapaj	tomato

- On the basis of concordance sets argues one could analyse: (i) lexicon, (ii) segmental inventories, (iii) prosody, (iv) morphosyntax, and (v) verb arguments.

Extensions?

- Will depend on the language…
  - e.g., in Intuktitut Lee and Allen (2022) analysed categories of noun roots (including demonstratives).

- Other components of semantics may be difficult beyond semantic fields.
  - e.g., the scope of lexical meanings (Bowerman, 1980).

- Interesting typological features will be revealed through lexicon.
  - Tone
  - Morphology (next meeting).

Some further readings…

Pye, C. (2021). Documenting the acquisition of indigenous languages. ***Journal of Child Language, 48,*** 454 – 479.

Taverna, A., & Waxman, S. (2020). Early lexical acquisition in the Wichi language. ***Journal of Child Language, 47,*** 1052 – 1072.