

Rapid Word Collection and *The Combine*

Kevin Warfel
Danny Rorabaugh

28 September, 2023

Word-collection strategies

- Flypaper method
- Translating a word list
- Working from a blank dictionary
- Elicitation - Using picture books
- Finding words in texts
- **Rapid Word Collection** workshops using *semantic domains*



Why Rapid Word Collection?

Why should you be interested?

- jump start a dictionary
- help translators
- standardize orthography

Rapid Word Collection (RWC)

- uses semantic domains to facilitate collection
- enables groups of speakers to work together
- consistently results in 10,000+ raw words
 - ~7,000 entries with one or more senses
 - most productive method in two weeks
- allows speakers to harvest important cultural vocabulary that might be missed via other methods (e.g., word list)
- “No other method both collects and classifies words at the same time.” - Ron Moe

RWC Workshop Data

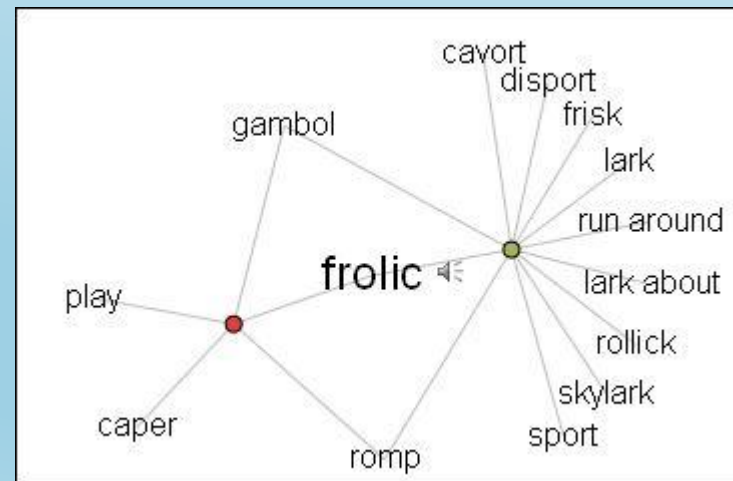
Language	Country	Year	# words
Buli	Ghana	2012	14,747
Lotud	Malaysia	2013	5,063
Kaansa	Burkina Faso	2014	11,500
Madhya-Purbiya	Nepal	2014	22,940
Syuba	Nepal	2014	12,608
Thulung	Nepal	2014	11,072
Koorete	Ethiopia	2014	11,889
Gwama	Ethiopia	2014	6,641
Bissa Barka	Burkina Faso	2015	13,802
Kemedzung	Cameroon	2015	13,685
Shilluk	South Sudan	2015	15,117
Kabwa	Tanzania	2015	8,880
Chungmboko	Cameroon	2015	12,264
Naami	Cameroon	2016	18,890
Dangaleat	Chad	2016	13,084
Ikizu-Sizaki	Tanzania	2017	13,504
Gusiilaay	Senegal	2017	12,485
Acoma	USA	2017	9,321
Djimini	Ivory Coast	2018	13,702
Crow	USA	2018	14,863
Haisla	Canada	2020	3,940
Karon	Senegal/Gambia	2020	8,713

Why semantic domains?

Research shows that

- we relate words to other words
- we organize words in a giant network
- words tend to cluster in groups called *semantic domains*

A semantic domain is like a **family of closely related words** that are linked in hierarchical and horizontal ways.



Implications of semantic domains

- Ability to rapidly recall words that belong to a particular semantic domain
- “sun”
 - light, moon, shine, sunbeam, sunrise, ...
- 4-5 people working together
 - goes more quickly, and the list is longer
- Repeat for each semantic domain
 - very effective method for capturing the building blocks of a language

Semantic Domains used in RWC

- Universe, creation
- Person
- Language and thought
- Social behavior
- Daily life
- Work and occupation
- Physical actions
- States
- Grammar

Each semantic domain includes a

- number

5.3.5 Clothes for special people

Use this domain for special clothes worn by special people. It is necessary to think through all the various special types of people in the culture and think of any special clothes worn by them. Several examples are given below, but there are many others.

What special clothes are worn by police or soldiers?

uniform, dress uniform, fatigues, camouflage, helmet, beret, cap, gasmask, braid,

What kinds of uniforms are there?

police uniform, military uniform, nurse's uniform, band uniform

What special clothes are worn by religious people?

habit (nun), cowl (monk), robe

What special clothes are worn by the king?

purple cloth, crown, royal robes

What special clothes are worn by doctors?

greens, surgical suit, white coat, nurses cap

Each semantic domain includes a

- domain label

5.3.5 Clothes for special people

Use this domain for special clothes worn by special people. It is necessary to think through all the various special types of people in the culture and think of any special clothes worn by them. Several examples are given below, but there are many others.

What special clothes are worn by police or soldiers?

uniform, dress uniform, fatigues, camouflage, helmet, beret, cap, gasmask, braid,

What kinds of uniforms are there?

police uniform, military uniform, nurse's uniform, band uniform

What special clothes are worn by religious people?

habit (nun), cowl (monk), robe

What special clothes are worn by the king?

purple cloth, crown, royal robes

What special clothes are worn by doctors?

greens, surgical suit, white coat, nurses cap

Each semantic domain includes a

- description

5.3.5 Clothes for special people

Use this domain for special clothes worn by special people. It is necessary to think through all the various special types of people in the culture and think of any special clothes worn by them. Several examples are given below, but there are many others.

What special clothes are worn by police or soldiers?

uniform, dress uniform, fatigues, camouflage, helmet, beret, cap, gasmask, braid,

What kinds of uniforms are there?

police uniform, military uniform, nurse's uniform, band uniform

What special clothes are worn by religious people?

habit (nun), cowl (monk), robe

What special clothes are worn by the king?

purple cloth, crown, royal robes

What special clothes are worn by doctors?

greens, surgical suit, white coat, nurses cap

Each semantic domain includes a

- series of elicitation questions

5.3.5 Clothes for special people

Use this domain for special clothes worn by special people. It is necessary to think through all the various special types of people in the culture and think of any special clothes worn by them. Several examples are given below, but there are many others.

What special clothes are worn by police or soldiers?

uniform, dress uniform, fatigues, camouflage, helmet, beret, cap, gasmask, braid,

What kinds of uniforms are there?

police uniform, military uniform, nurse's uniform, band uniform

What special clothes are worn by religious people?

habit (nun), cowl (monk), robe

What special clothes are worn by the king?

purple cloth, crown, royal robes

What special clothes are worn by doctors?

greens, surgical suit, white coat, nurses cap

Each semantic domain includes a

- (optional) sample words

5.3.5 Clothes for special people

Use this domain for special clothes worn by special people. It is necessary to think through all the various special types of people in the culture and think of any special clothes worn by them. Several examples are given below, but there are many others.

What special clothes are worn by police or soldiers?

uniform, dress uniform, fatigues, camouflage, helmet, beret, cap, gasmask, braid,

What kinds of uniforms are there?

police uniform, military uniform, nurse's uniform, band uniform

What special clothes are worn by religious people?

habit (nun), cowl (monk), robe

What special clothes are worn by the king?

purple cloth, crown, royal robes

What special clothes are worn by doctors?

greens, surgical suit, white coat, nurses cap

Prerequisites: best-practice RWC workshop

1. A motivated language community

Prerequisites: best-practice RWC workshop

1. A motivated language community
2. A functional orthography

Prerequisites: best-practice RWC workshop

1. A motivated language community
2. A functional orthography
3. A sufficient number of literate people

Prerequisites: best-practice RWC workshop

1. A motivated language community
2. A functional orthography
3. A sufficient number of literate people
4. A sufficient number of bilingual participants

Prerequisites: best-practice RWC workshop

1. A motivated language community
2. A functional orthography
3. A sufficient number of literate people
4. A sufficient number of bilingual participants
5. A workshop coordinator

RWC workshop roles:

- Consultant
- Coordinator
- Logistics Manager
- Team leaders
- Typists
- Language experts

Phases of a best-practice RWC workshop

1. Preparation phase
2. Training phase
3. Word-collection phase
4. Cleanup phase

Preparation

Word Collection

4-6 months

3
days

10 days

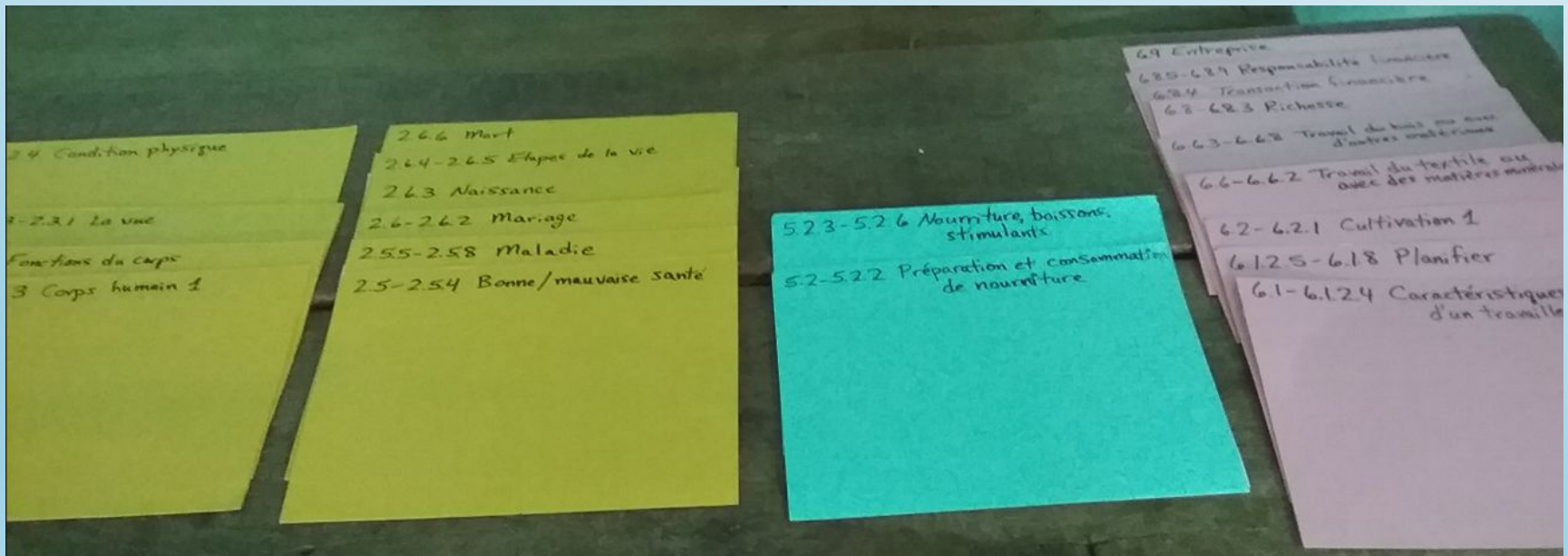
5 days

Training

Initial Cleanup

Word-collection phase

- semantic domains organized into folders



Word-collection phase

- semantic domains organized into folders
 - closely related domains

2.5	Healthy
2.5.1	Sick
2.5.1.1	Recover from sickness
2.5.2	Disease
2.5.2.1	Malnutrition, starvation
2.5.2.2	Skin disease
2.5.2.3	Stomach illness
2.5.2.4	Tooth decay
2.5.3	Injure
2.5.3.1	Amputate
2.5.3.2	Poison
2.5.4	Disabled
2.5.4.1	Blind
2.5.4.2	Poor eyesight
2.5.4.3	Deaf
2.5.4.4	Mute
2.5.4.5	Birth defect

Word-collection phase

- semantic domains organized into folders
 - closely related domains
- 6 groups



Word-collection phase

- semantic domains organized into folders
 - closely related domains
- 6 groups
 - 4-6 people per group



Word-collection phase

- semantic domains organized into folders
 - closely related domains
- 6 groups
 - 4-6 people per group
 - each working with a different folder
 - group members call out words



Word-collection phase

- semantic domains organized into folders
 - closely related domains
- 6 groups
 - 4-6 people per group
 - each working with a different folder
 - group members call out words
 - Typist enters the words, using The Combine
 - Gloss is decided by group, entered by Typist

The result is the beginnings of a dictionary and a thesaurus, with each entry consisting of a word, a gloss, and one or more semantic domain tags.

baninge *Heliconia humilis* (ch. sém. : 1.5.3 - Herbes, lianes, plantes grimpantes.)
bannga rolhier d'Abyssinie (ch. sém. : 1.6.1.2 - Oiseau.)
banḡanlaga apiaceae (ch. sém. : 2.5.7.5 - Médecine traditionnelle.)
banḡanlaga kulo *Steganotaenia araliacea* (ch. sém. : 1.5.3 - Herbes, lianes, plantes grimpantes.)
banḡoloma métier à tisser (ch. sém. : 6.6.1.4 - Tisser.)
banḡibanḡi blanc et noir (ch. sém. : 8.3.3.3.5 - Couleur des animaux, tache.)
baramba yenle moucherotte à ventre roux (ch. sém. : 1.6.1.2 - Oiseau.)
bari empêcher (ch. sém. : 3.3.4.4 - Empêcher.)
bari lere na surprendre quelqu'un (ch. sém. : 3.4.1.3 - Surprise.)
bariko barrique (ch. sém. : 6.6.7.2 - Transport de l'eau, 5.2.3.7.1 - Préparation alcoolisée.)
bashi bache (ch. sém. : 7.2.4.1.1 - Véhicule.)
bateri batterie (ch. sém. : 7.2.4.1.1 - Véhicule, 6.6.8.1 - Travail associé à l'électricité.)
batemu baptême (ch. sém. : 4.9.5.4 - Cérémonie religieuse, 4.2.2.1 - Cérémonie.)
batisi Baptiste (ch. sém. : 4.9.7.2 - Christianisme.)
bo₁ 1) bambou de Chine 2) bambou géant (ch. sém. : 1.5.1 - Arbres, 6.5.3 - Matériaux de construction.) 2) *Oxytenanthera abyssinica* (ch. sém. : 6.6.4.2 - Vannerie.)
bo₂ suffire (ch. sém. : 8.1.7 - Assez.)
bokagara dialecte de Bokala (ch. sém. : 9.7.1.5 - Nom de langue.)
Bokala 1) clan Bokala (ch. sém. : 9.7.1.3 - Nom de clan.) 2) village Bokala (ch. sém. : 9.7.2.3 - Nom de ville.)

Translators can make use of this even in this rudimentary form.

Headword	Lexeme Form	Glosses (Eng)	Semantic Domains
Show All	Show All	Show All	Person in authority
mambo	mazimambo	king	4.5.1 - Person in authority; 4.6.1 - Ruler
mpfumu₁	mpfumu	chief	4.5.1 - Person in authority; 4.6.1 - Ruler
mpfumu₂	mpfumu	chief	4.5.1 - Person in authority; 4.6.1 - Ruler
ntongi	tongi	an authority	4.5.1 - Person in authority; 4.7.6 - Judge, render a verdict
ntsogoleri	tsogoleri	leader	4.5.1 - Person in authority; 7.2.5.3 - Guide; 7.2.5.1 - Go first
nyakutongi	nyakutongi	leader	4.5.1 - Person in authority
nyakwawa	nyakwawa	ruler province	4.6.1.2 - Government official; 4.5.1 - Person in authority

SIL's word-collection tools:

- *Collect Words* tool
 - in FLEx
- *The Combine*
 - a web app

“some assembly required”

Alternatives to the best-practice formula

- too few speakers available
- speakers aren't available for 10 full days
- incorporation of audio recordings
- collecting words in multiple dialects
- glossing words in multiple languages

References

Boerger, Brenda, and Verna Stutzman. 2018. “Single-event Rapid Word Collection workshops: Efficient, effective, empowering.” *Language Documentation & Conservation* 12 (2018): 147–193. <http://hdl.handle.net/10125/24766>, accessed June 1, 2020.

Moe, Ronald. 2007. “Dictionary Development Program”. SIL Forum for Language Fieldwork. SIL.org. https://www.sil.org/system/files/reapdata/74/26/20/74262098499698801429138168340553232050/SILForum2007_003.pdf, accessed June 9, 2020.