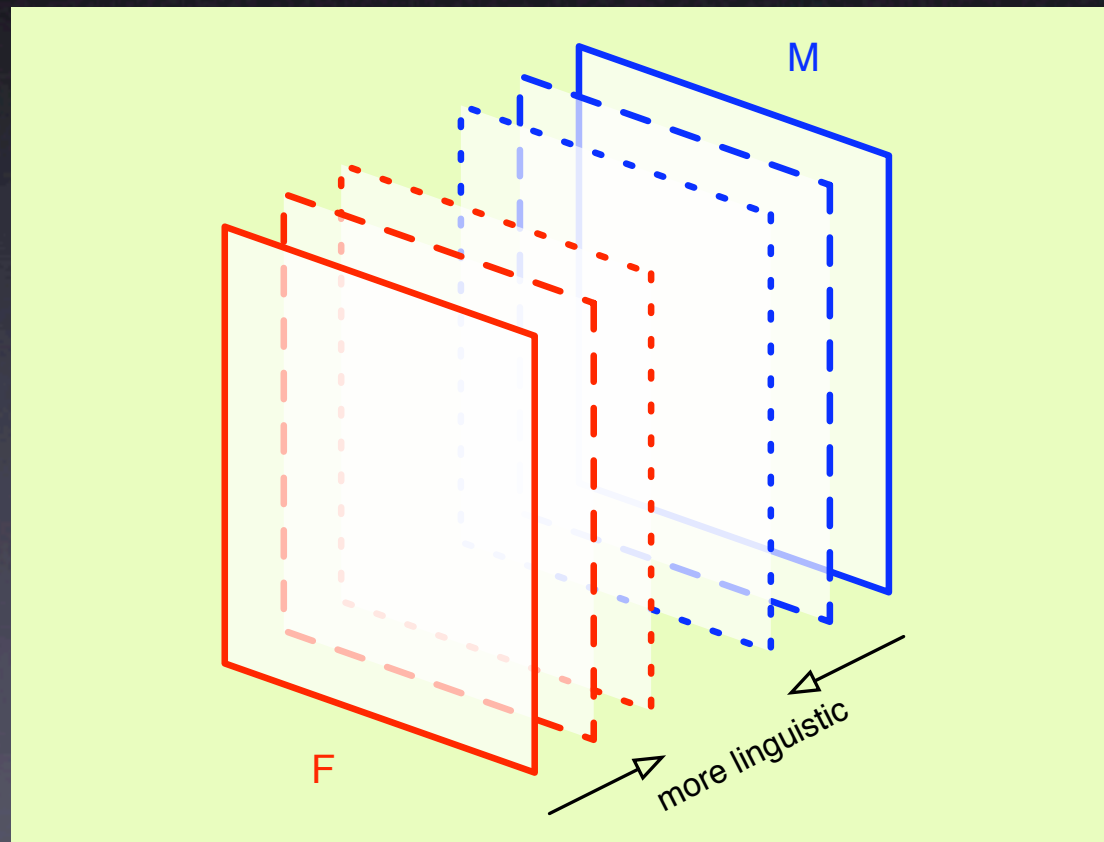


Units of composition

2009-09-01

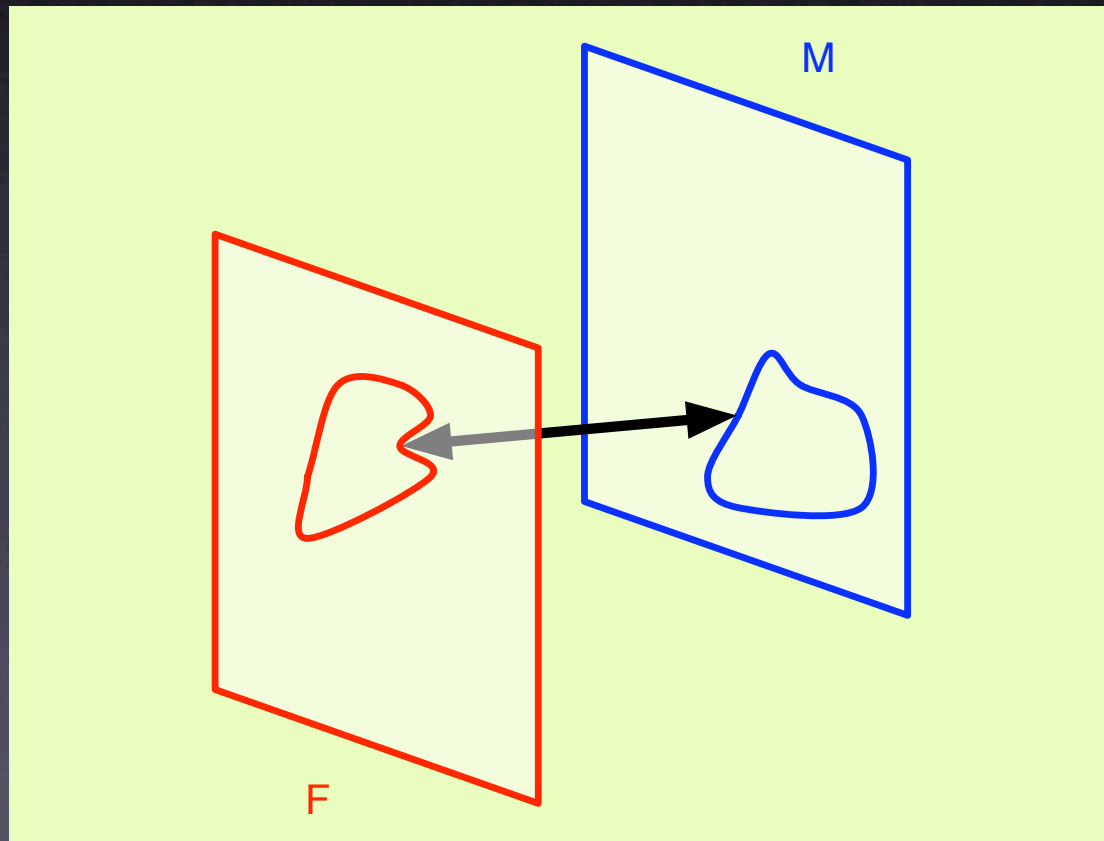
Form and Meaning

- Linguistics is about abstract levels mediating Forms and Meanings.



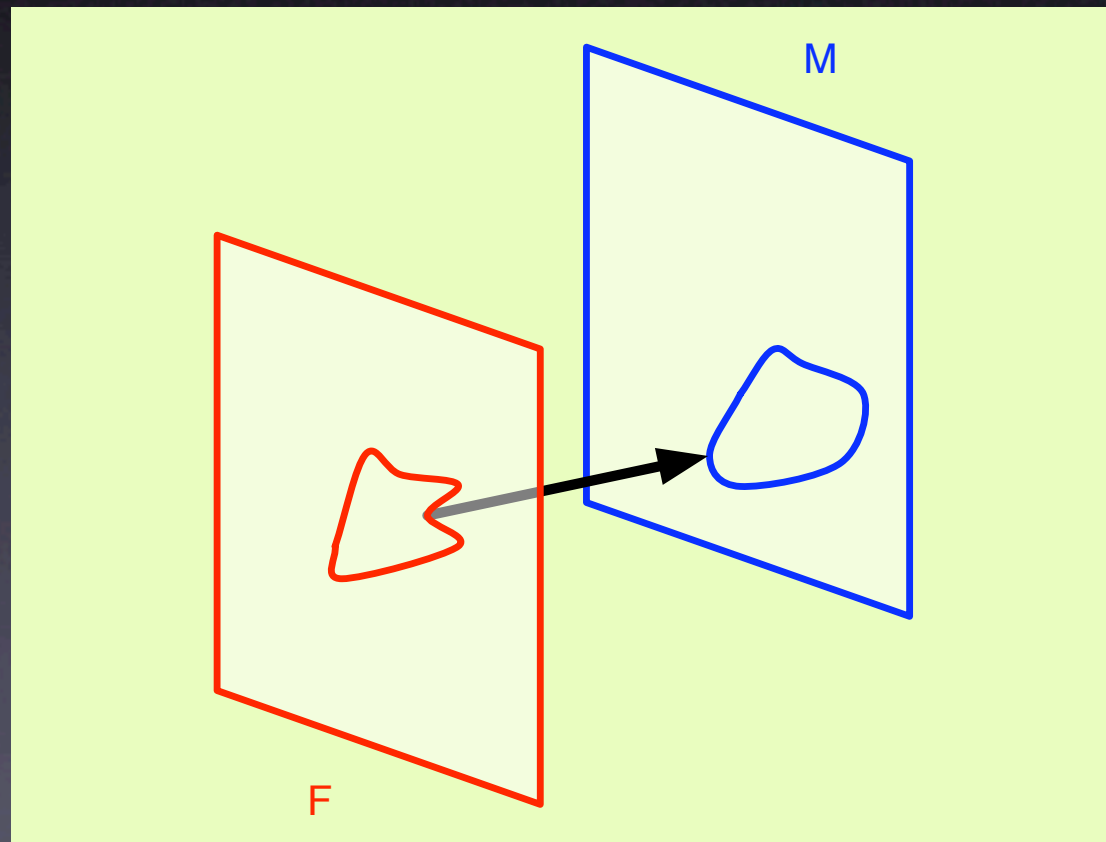
An utterance

- An utterance associates a Form with a Meaning.



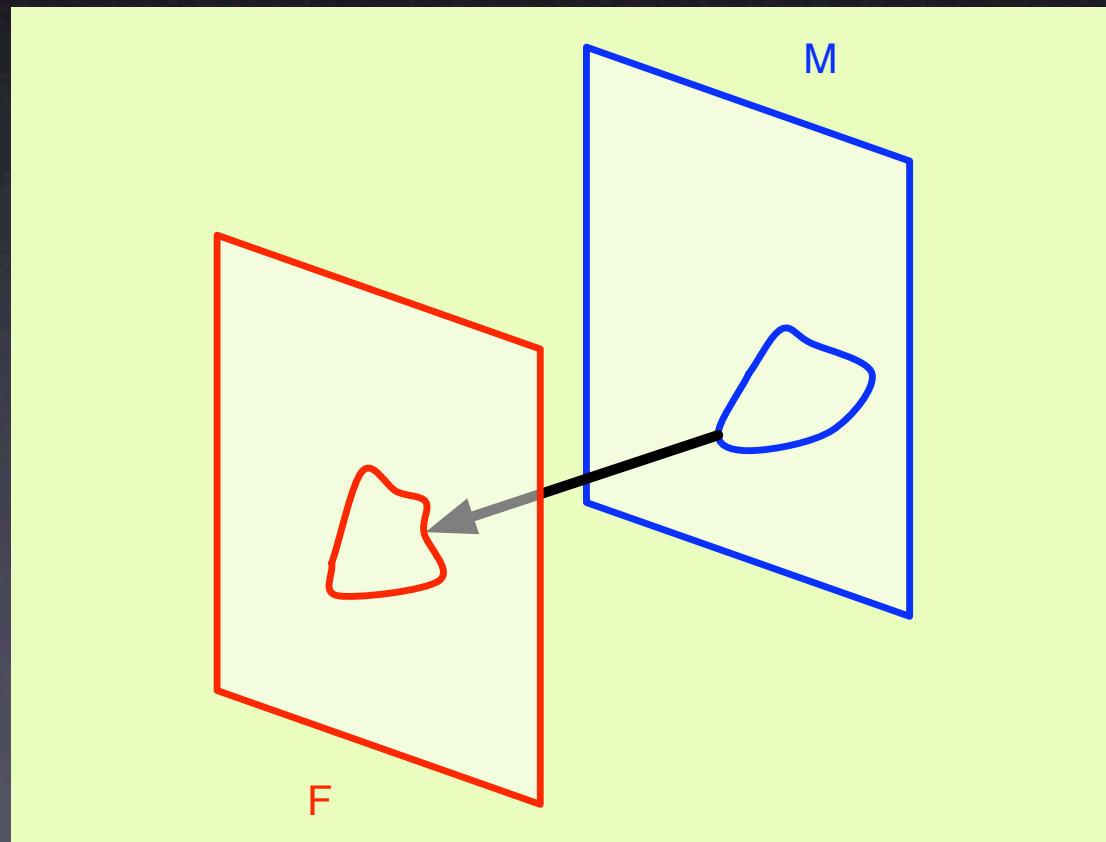
Generalization (productivity): comprehension

- Novel utterances can be comprehended.



Generalization (productivity): production

- Novel utterances can be produced.



The generalization problem

The generalization problem

- How can speakers produce utterances they've never heard or said before?

The generalization problem

- How can speakers produce utterances they've never heard or said before?
- How can hearers understand utterances they've never heard before?

The generalization problem

- How can speakers produce utterances they've never heard or said before?
- How can hearers understand utterances they've never heard before?
- **Compositionality** as a solution that

The generalization problem

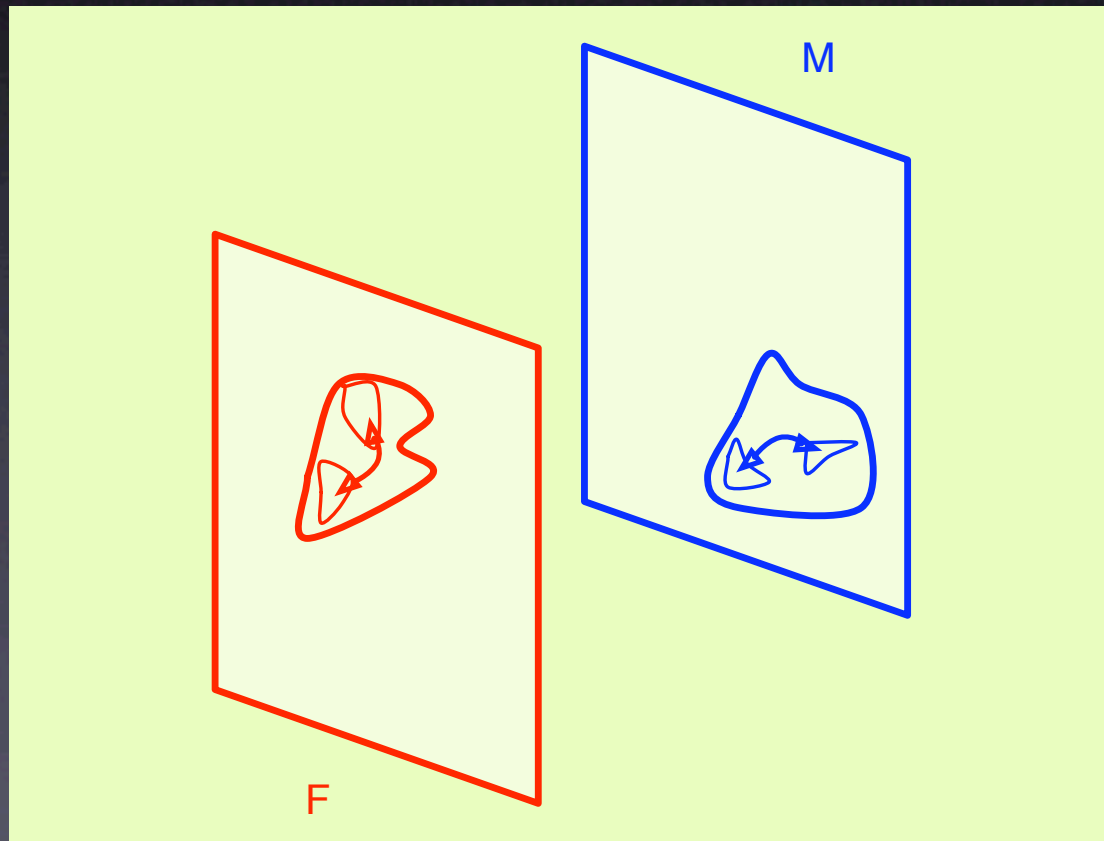
- How can speakers produce utterances they've never heard or said before?
- How can hearers understand utterances they've never heard before?
- **Compositionality** as a solution that
 - Emerged as human language evolved

The generalization problem

- How can speakers produce utterances they've never heard or said before?
- How can hearers understand utterances they've never heard before?
- **Compositionality** as a solution that
 - Emerged as human language evolved
 - Is rediscovered in each generation by language learners

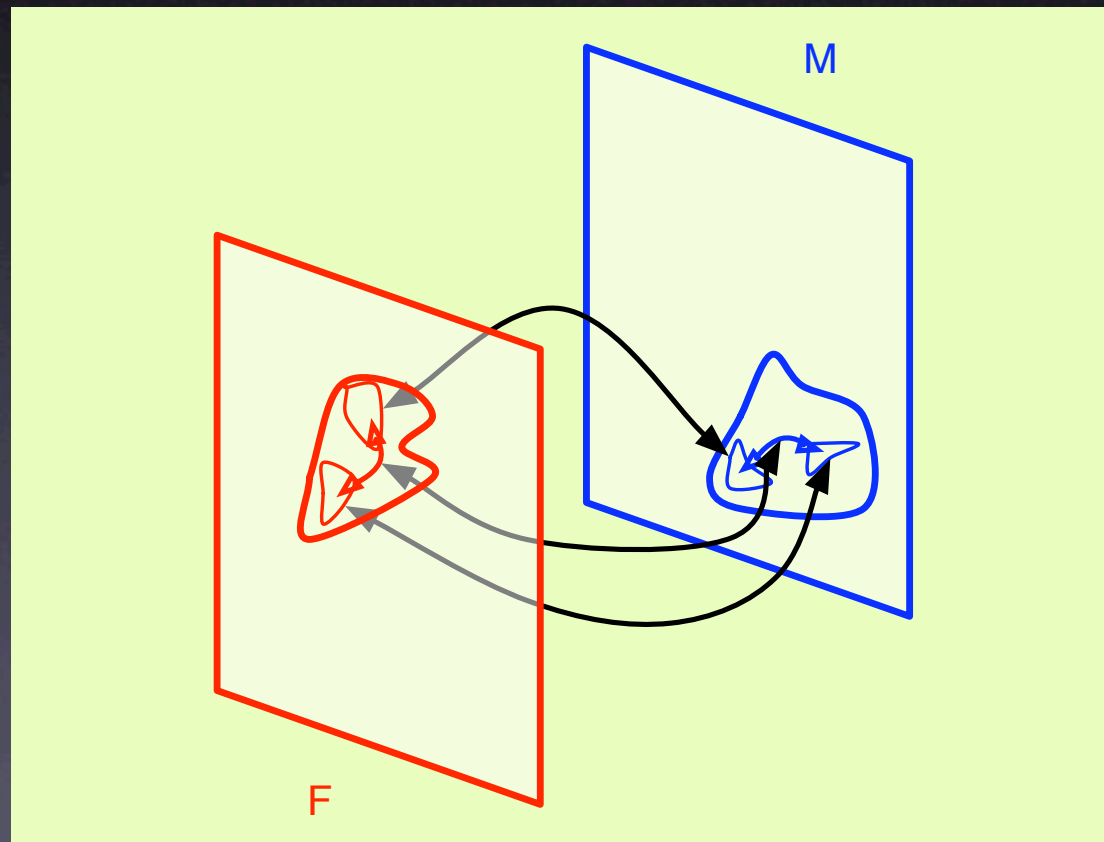
Composition 1

- Forms and Meanings have parts and relations between the parts.



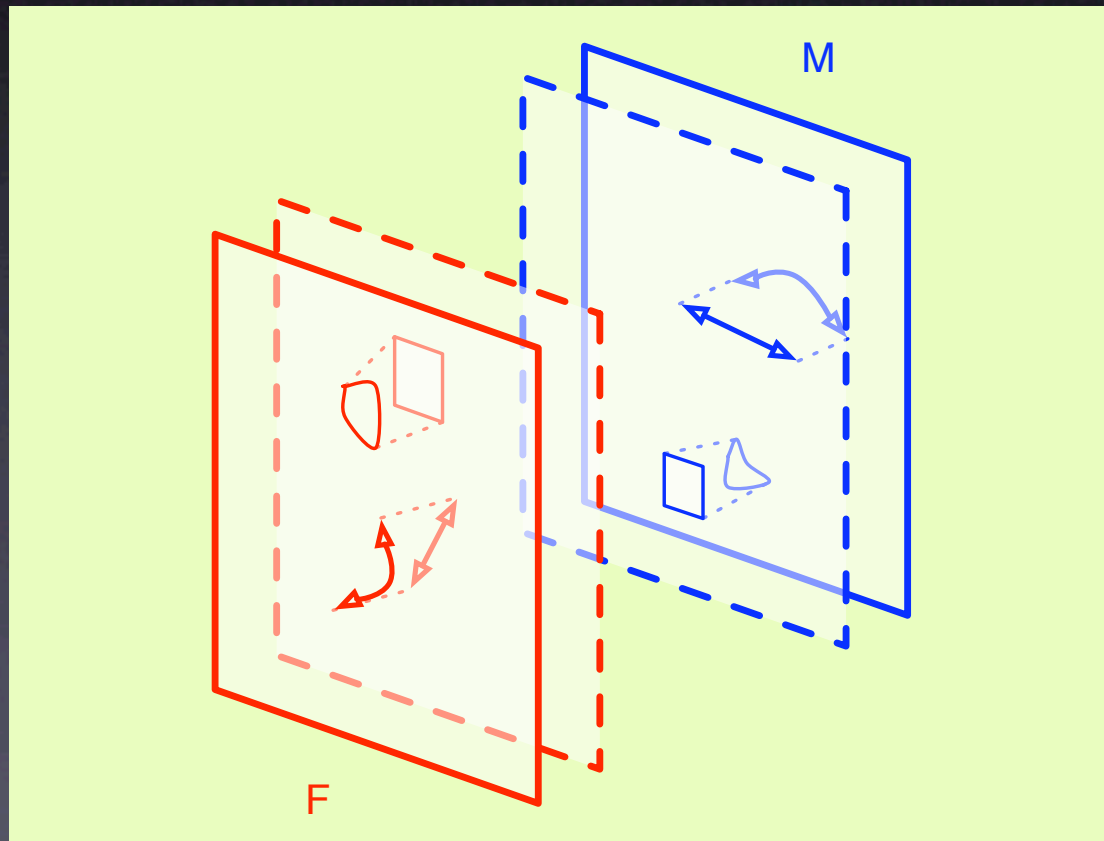
Composition 1

- Composition associates Form and Meaning parts and relations.



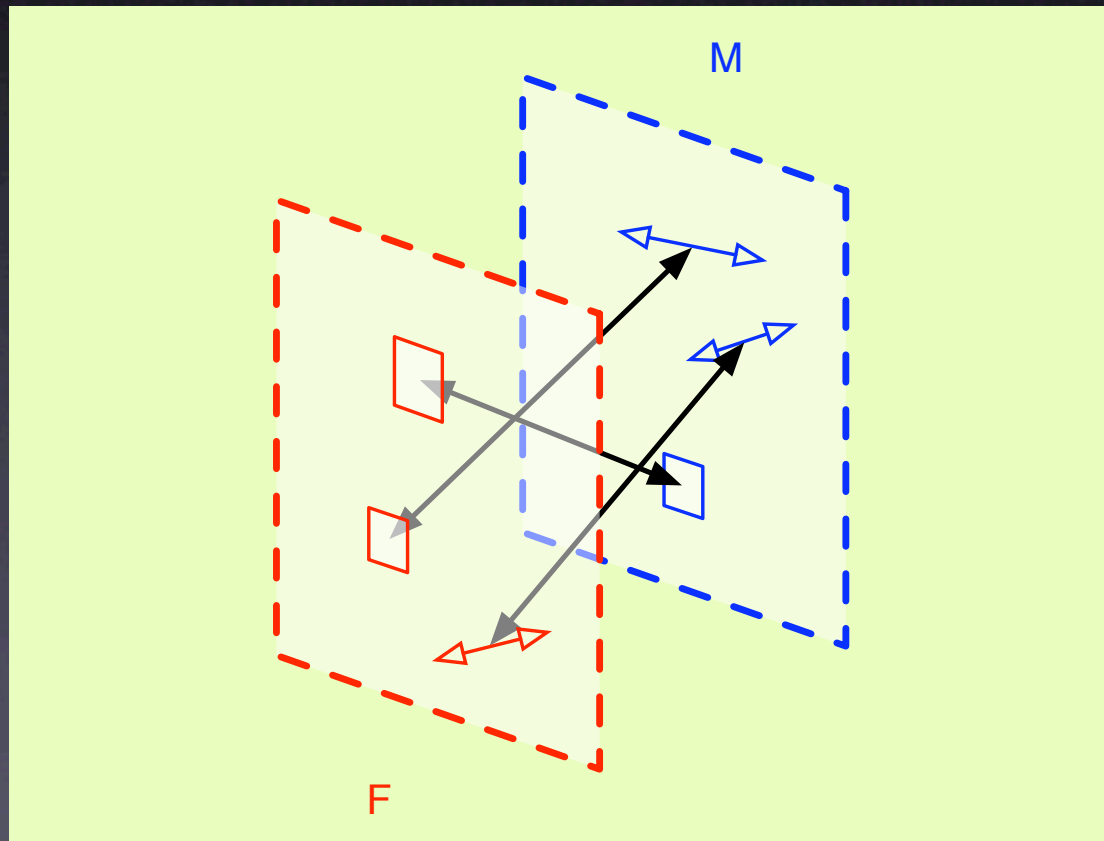
Composition 2

- Linguistic generalizations abstract over parts and relations.



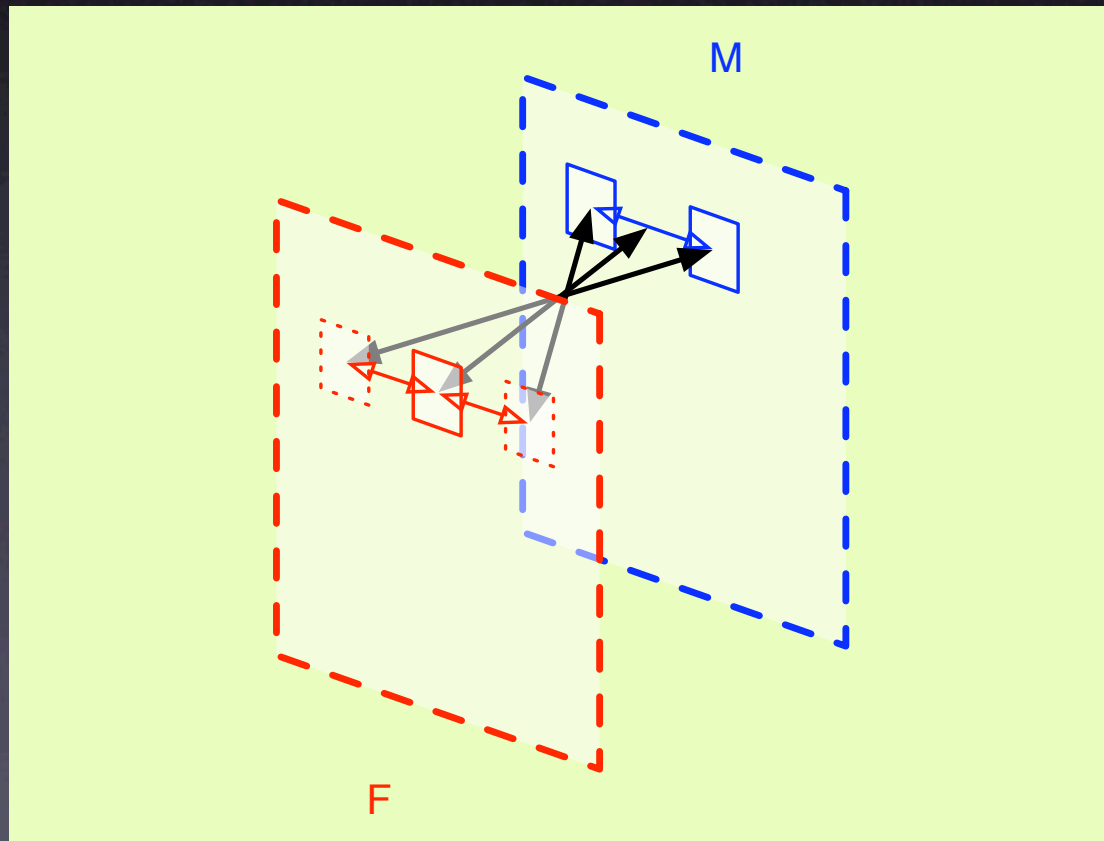
Composition 2

- Linguistic knowledge is about associations between abstracted parts and relations.



Composition 3

- Form parts (like verbs) can be about Meaning relations and can have associated Form slots.



Limits to compositionality

Limits to compositionality

- Chunking and idioms

Limits to compositionality

- Chunking and idioms
 - *weapons of mass destruction*

Limits to compositionality

- Chunking and idioms
 - *weapons of mass destruction*
 - *the Pledge of Allegiance*

Limits to compositionality

- Chunking and idioms
 - *weapons of mass destruction*
 - *the Pledge of Allegiance*
 - *Give me a break.*

Limits to compositionality

- Chunking and idioms
 - *weapons of mass destruction*
 - *the Pledge of Allegiance*
 - *Give me a break.*
 - *throw in the towel*

Units of composition

Units of composition

- Segmentation and aggregation

Units of composition

- Segmentation and aggregation
- What units should do

Units of composition

- Segmentation and aggregation
- What units should do
 - Each potential unit is a category with F and M parts.

Units of composition

- Segmentation and aggregation
- What units should do
 - Each potential unit is a category with F and M parts.
 - A “good” unit is one that supports compositionality in the comprehension and production of utterances.

Units of composition

- Segmentation and aggregation
- What units should do
 - Each potential unit is a category with F and M parts.
 - A “good” unit is one that supports compositionality in the comprehension and production of utterances.
 - For a good unit, both the F and M parts as well as their association should be relatively coherent, bounded, stable.

Units of composition

- Segmentation and aggregation
- What units should do
 - Each potential unit is a category with F and M parts.
 - A “good” unit is one that supports compositionality in the comprehension and production of utterances.
 - For a good unit, both the F and M parts as well as their association should be relatively coherent, bounded, stable.
- Dimensions of “unithood”

Units of composition

- Segmentation and aggregation
- What units should do
 - Each potential unit is a category with F and M parts.
 - A “good” unit is one that supports compositionality in the comprehension and production of utterances.
 - For a good unit, both the F and M parts as well as their association should be relatively coherent, bounded, stable.
- Dimensions of “unithood”
 - Phonetic

Units of composition

- Segmentation and aggregation
- What units should do
 - Each potential unit is a category with F and M parts.
 - A “good” unit is one that supports compositionality in the comprehension and production of utterances.
 - For a good unit, both the F and M parts as well as their association should be relatively coherent, bounded, stable.
- Dimensions of “unithood”
 - Phonetic
 - Syntactic/morphological

Units of composition

- Segmentation and aggregation
- What units should do
 - Each potential unit is a category with F and M parts.
 - A “good” unit is one that supports compositionality in the comprehension and production of utterances.
 - For a good unit, both the F and M parts as well as their association should be relatively coherent, bounded, stable.
- Dimensions of “unithood”
 - Phonetic
 - Syntactic/morphological
 - Semantic

Units of composition

- Segmentation and aggregation
- What units should do
 - Each potential unit is a category with F and M parts.
 - A “good” unit is one that supports compositionality in the comprehension and production of utterances.
 - For a good unit, both the F and M parts as well as their association should be relatively coherent, bounded, stable.
- Dimensions of “unithood”
 - Phonetic
 - Syntactic/morphological
 - Semantic
 - F–M association

Differences between languages

Differences between languages

- Languages differ in the complexity of “words”, in the possibilities for “bound morphology” (the analytic–synthetic dimension).

Differences between languages

- Languages differ in the complexity of “words”, in the possibilities for “bound morphology” (the analytic–synthetic dimension).
- Languages differ in terms of how much they make use of primitive, abstract units to express particular meanings. Analytic languages may have many more possible “words” and may have many fewer morphemes than synthetic languages.

Morphology: inflectional

- Nouns: number, class, case
- Verbs: tense, aspect, mood; subject and object agreement
 - *you see me*
kinawilo (K'iche')
 - unaniona* (Swahili)
- Why *you see me* is considered three words and *unaniona* one word

Morphology: derivational

Morphology: derivational

- *darken*, *watery*, *sleepiness*, *funny* (?)

Morphology: derivational

- *darken*, *watery*, *sleepiness*, *funny* (?)
- Stems and roots: *darkening*

Morphology: derivational

- *darken*, *watery*, *sleepiness*, *funny* (?)
- Stems and roots: *darkening*
- *arguu* 'to see' (Oromo)
argamuu 'to be seen'
agarsiisuu 'to cause to see; to show'
argamsiisuu 'to cause to be seen; to bring about'

Morphology: derivational

- *darken*, *watery*, *sleepiness*, *funny* (?)
- Stems and roots: *darkening*
- *arguu* 'to see' (Oromo)
argamuu 'to be seen'
agarsiisuu 'to cause to see; to show'
argamsiisuu 'to cause to be seen; to bring about'
- *llank'ay* 'to work' (Quechua)
llank'arpariy 'to work completely'

Morphology: derivational

- *darken*, *watery*, *sleepiness*, *funny* (?)
- Stems and roots: *darkening*
- *arguu* ‘to see’ (Oromo)
argamuu ‘to be seen’
agarsiisuu ‘to cause to see; to show’
argamsiisuu ‘to cause to be seen; to bring about’
- *llank’ay* ‘to work’ (Quechua)
llank’arpariy ‘to work completely’
- *yəmmittikkəbabbərulliñ* (Amharic)
kkəbabbər = *kbr* + *C1C1əC2aC2C2əC3*
‘[you(pl.)] who respect each other for me’

Morphology: incorporation

- *mepitetatsihetixakítetuirietsitiximekaitikaku*
‘however, they already wanted to start ordering us to
bring them the dishes’ (Huichol, Iturrioz & Gómez López)

An very analytic language: Nuxalk vs. English words

An very analytic language: Nuxalk vs. English words

- *smnmnmuuc* 'mute'

An very analytic language: Nuxalk vs. English words

- *smnmnmuuc* 'mute'
- *nusʔaqulik^wsta* 'prison'

An very analytic language: Nuxalk vs. English words

- *smnmnmuuc* 'mute'
- *nusʔaḡulik^wsta* 'prison'
- *snuṭnusltʔ* 'twins'

An very analytic language: Nuxalk vs. English words

- *smnmnmuuc* 'mute'
- *nusʔaqwsta* 'prison'
- *snutnuslt* 'twins'
- *kitnuus* 'eight'

An very analytic language: Nuxalk vs. English words

- *smnmnmuuc* 'mute'
- *nusʔaq
 - u
lik^wsta* 'prison'
- *snutnuslt* 'twins'
- *kitnuus* 'eight'
- *ʔittutcuut* 'naked'

An very analytic language: Nuxalk vs. English words

- *smnmnmuuc* 'mute'
- *nusʔaqulik^wsta* 'prison'
- *snułnusłt* 'twins'
- *kıtnuus* 'eight'
- *ʔitłutcuut* 'naked'
- *nutixcnalsiix^w* 'uvula'

An very analytic language: Nuxalk vs. English words

- *smnmnmuuc* 'mute'
- *nusʔaqulik^wsta* 'prison'
- *snutnuslt* 'twins'
- *kitnuus* 'eight'
- *ʔittucuut* 'naked'
- *nutixcnalsiix^w* 'uvula'
- *squpuc* 'beard'

Lexicalization in Nuxalk

Lexicalization in Nuxalk

- *ciiqnka* ‘be busy digging roots’

Lexicalization in Nuxalk

- *ciiqnka* ‘be busy digging roots’
- *numawaluuti* ‘be all alone in one’s boat’

Lexicalization in Nuxalk

- *ciiqnka* ‘be busy digging roots’
- *numawaluuti* ‘be all alone in one’s boat’
- *nuciq[˜]waq[˜]ws* ‘break sbd’s glasses’

Lexicalization in Nuxalk

- *ciiqnka* ‘be busy digging roots’
- *numawaluuti* ‘be all alone in one’s boat’
- *nuciq[˜]waq[˜]ws* ‘break sbd’s glasses’
- *tayamkiiχ^{wm}* ‘throw away one’s hat’

Lexicalization in Nuxalk

- *ciiqnka* ‘be busy digging roots’
- *numawaluuti* ‘be all alone in one’s boat’
- *nuciq[˜]waq[˜]ws* ‘break sbd’s glasses’
- *tayamkiiχ^{wm}* ‘throw away one’s hat’
- *nusq^wtaʔ* ‘have a fishbone stuck in one’s throat’

Lexicalization in Nuxalk

- *ciiqnka* ‘be busy digging roots’
- *numawaluuti* ‘be all alone in one’s boat’
- *nuciq[˜]waq[˜]ws* ‘break sbd’s glasses’
- *tayamkiiχ^{wm}* ‘throw away one’s hat’
- *nusq^wtatt* ‘have a fishbone stuck in one’s throat’
- *plxlyslayx* ‘be no longer able to speak’

Lexicalization in Nuxalk

- *ciiqnka* ‘be busy digging roots’
- *numawaluuti* ‘be all alone in one’s boat’
- *nuciq[˜]waq[˜]ws* ‘break sbd’s glasses’
- *tayamkiiχ^{wm}* ‘throw away one’s hat’
- *nusq^wtatt* ‘have a fishbone stuck in one’s throat’
- *plxlyslayx* ‘be no longer able to speak’
- *spuxamtimut* ‘change oneself into eagle down’

Lexicalization in Nuxalk

- *ciiqnka* ‘be busy digging roots’
- *numawaluuti* ‘be all alone in one’s boat’
- *nuciq[˜]waq[˜]ws* ‘break sbd’s glasses’
- *tayamkiiχ^{wm}* ‘throw away one’s hat’
- *nusq^wtatt* ‘have a fishbone stuck in one’s throat’
- *plxlyslayx* ‘be no longer able to speak’
- *spuxamtimut* ‘change oneself into eagle down’
- *?itcayanm* ‘be the time of year to pick berries’

Lexicalization in Nuxalk

- *ciiqnka* ‘be busy digging roots’
- *numawaluuti* ‘be all alone in one’s boat’
- *nuciq[˜]waq[˜]ws* ‘break sbd’s glasses’
- *tayamkiiχ^{wm}* ‘throw away one’s hat’
- *nusq^wtatt* ‘have a fishbone stuck in one’s throat’
- *plxlχslayx* ‘be no longer able to speak’
- *spuxamtimut* ‘change oneself into eagle down’
- *ʔitcayanm* ‘be the time of year to pick berries’
- *knixalstn* ‘eat sbd else’s food’

Abstract morphemes in Nuxalk

- *sti-* ‘asymmetric’
 - *stispułankiiq^w* ‘have a lump on one side of the head’
 - *stik^łaałqii* ‘have one leg shorter than the other’
- *-ak-*, *-aak-* ‘hand, work’
 - *cpakm* ‘wipe sth with one’s hand’
 - *numawaakmicut* ‘be all alone at one’s work’
 - *stip^łnakt* ‘tree branch’
- *-uc-*, *-uuc-* ‘mouth, opening, speaking, eating, water’
 - *cpu^ucm* ‘wipe one’s mouth’
 - *atnap^uc* ‘know what sbd says’
 - *txulał^łsu^uc* ‘tide comes in’

Terms for kinds of units

Terms for kinds of units

- Wordform

Terms for kinds of units

- **Wordform**
 - *weakening, bæmmittikkəbabbərulliñ, plxlɣslayxc*

Terms for kinds of units

- Wordform
 - *weakening, bæmmittikkəbabbərulliñ, plxlɣslayxc*
- Morpheme

Terms for kinds of units

- **Wordform**
 - *weakening, bæmmittikkəbabbərulliñ, plxlɣslayxc*
- **Morpheme**
 - *weak-en-ing, bæ-mmi-tti-kbr+C1C1əC2aC2C2əC3-u-ll-iñ*

Terms for kinds of units

- **Wordform**
 - *weakening, bæmmittikkəbabbərulliñ, plxlɣslayxc*
- **Morpheme**
 - *weak-en-ing, bæ-mmi-tti-kbr+C1C1əC2aC2C2əC3-u-ll-iñ*
- **Lexeme**

Terms for kinds of units

- **Wordform**
 - *weakening, bæmmittikkəbabbərulliñ, plxlɣslayxc*
- **Morpheme**
 - *weak-en-ing, bæ-mmi-tti-kbr+C1C1əC2aC2C2əC3-u-ll-iñ*
- **Lexeme**
 - *WEAK: weak, weakly, weaker, weakest, weaken?*

Terms for kinds of units

- **Wordform**
 - *weakening*, *bəmmittikkəbabbərulliñ*, *plxlɣslayxc*
- **Morpheme**
 - *weak-en-ing*, *bə-mmi-tti-kbr+C1C1əC2aC2C2əC3-u-ll-iñ*
- **Lexeme**
 - *WEAK*: *weak, weakly, weaker, weakest, weaken?*
 - *KBR* 'respect' (Amharic)

Terms for kinds of units

- **Wordform**
 - *weakening*, *bəmmittikkəbabbərulliñ*, *plxlɣslayxc*
- **Morpheme**
 - *weak-en-ing*, *bə-mmi-tti-kbr+C1C1əC2aC2C2əC3-u-ll-iñ*
- **Lexeme**
 - *WEAK*: *weak*, *weakly*, *weaker*, *weakest*, *weaken?*
 - *KBR* 'respect' (Amharic)
- **Lemma**: canonical form of lexeme

Terms for kinds of units

- **Wordform**
 - *weakening*, *bəmmittikkəbabbərulliñ*, *plxlɣslayxc*
- **Morpheme**
 - *weak-en-ing*, *bə-mmi-tti-kbr+C1C1əC2aC2C2əC3-u-ll-iñ*
- **Lexeme**
 - *WEAK*: *weak, weakly, weaker, weakest, weaken?*
 - *KBR* 'respect' (Amharic)
- **Lemma**: canonical form of lexeme
 - *kəbbərə* 'he respected' (Amharic)

Terms for kinds of units

- **Wordform**
 - *weakening*, *bəmmittikkəbabbərulliñ*, *plxlɣslayxc*
- **Morpheme**
 - *weak-en-ing*, *bə-mmi-tti-kbr+C1C1əC2aC2C2əC3-u-ll-iñ*
- **Lexeme**
 - *WEAK*: *weak, weakly, weaker, weakest, weaken?*
 - *KBR* 'respect' (Amharic)
- **Lemma**: canonical form of lexeme
 - *kəbbərə* 'he respected' (Amharic)
- **Lexical item**

Terms for kinds of units

- **Wordform**
 - *weakening*, *bəmmittikkəbabbərulliñ*, *plxlɣslayxc*
- **Morpheme**
 - *weak-en-ing*, *bə-mmi-tti-kbr+C1C1əC2aC2C2əC3-u-ll-iñ*
- **Lexeme**
 - *WEAK*: *weak, weakly, weaker, weakest, weaken?*
 - *KBR* 'respect' (Amharic)
- **Lemma**: canonical form of lexeme
 - *kəbbərə* 'he respected' (Amharic)
- **Lexical item**
 - *WEAK*

Terms for kinds of units

- **Wordform**
 - *weakening, bəmmittikkəbabbərulliñ, plxlɣslayxc*
- **Morpheme**
 - *weak-en-ing, bə-mmi-tti-kbr+C1C1əC2aC2C2əC3-u-ll-iñ*
- **Lexeme**
 - *WEAK: weak, weakly, weaker, weakest, weaken?*
 - *KBR 'respect' (Amharic)*
- **Lemma**: canonical form of lexeme
 - *kəbbərə 'he respected' (Amharic)*
- **Lexical item**
 - *WEAK*
 - *THROW IN THE TOWEL*