

جامعة نيويورك أبوظبي



PSYCH-UH 2218: Language Science

Class 24: Language and thought

Prof. Jon Sprouse
Psychology

What is the relationship between language and thought?

We know that language is used to convey meanings to other humans.

For many humans, language also mediates some amount of our internal thoughts through an internal dialogue.

So we can ask the question: what exactly is the relationship between the cognitive ability we call **language** and the cognitive ability we call **thought**?

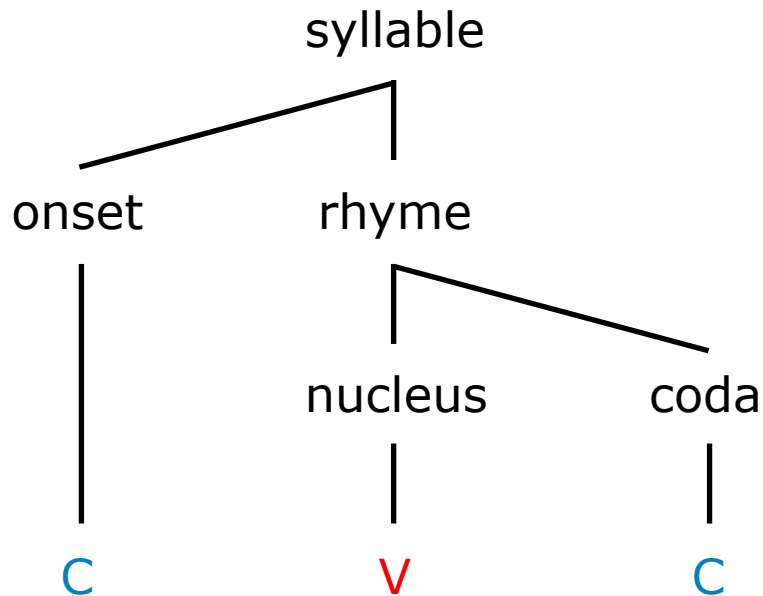


calamitiesofnature.com © 2011 Tony Piro

From a theoretical point of view, this is asking what the **scope** of our language theory is, and how it **interacts** with other cognitive mechanisms!

What is our theory of language?

We saw last time that our theory defines a language as a set of mental representations, and our theory of the cognitive system of language is a mathematical object, called a grammar, that defines that set intensionally:



/t/ → [t^h] / [syll ___]

/t/ → [r] / [V̇] ___ [V]

/t/ → [ʔ] / [V̇] ___ [ŋ]

derivational

-ful ADJ → N+ful

-ive ADJ → V+ive

-ize V → N+ize

-ship N → N+ship

/in/ → [il] / ___ l

/in/ → [ir] / ___ r

/in/ → [im] / ___ [bilabial]

/in/ → [inj] / ___ [velar]

What is our theory of thought?

This is **far** beyond the scope of this class, but we can sketch a plausible general shape of the theory:

A set of **primitive units** called **concepts**. These are the basic building blocks of complex thoughts.



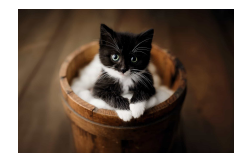
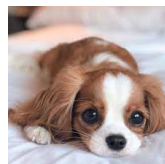
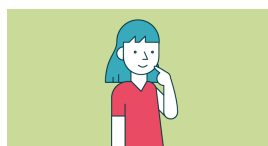
A **set of rules** for building more complex thoughts by **combining** primitive concepts:



Perhaps also a **set of rules** for establishing higher-order relationships among concepts:

Maybe something like logic?

$P \rightarrow Q$ "if P then Q"



Possible relationships

We can see some interesting similarities: both language and thought have primitive units, and both have rules for combining those units into larger and larger complex objects.

Extreme idea 1: Language **determines** thought

Cognitively, this would mean that the system of language is the system of thought or somehow constrains the system of thought such that the combinatorics of thought are identical to the combinatorics of language.

In-between: Language **influences** thought

Cognitively, this would mean that the two systems are separate, but there is an interaction. The term influence is **vague**. We will need to make it precise with cognitive mechanisms in order to explore it.

Extreme idea 2: Language and thought are completely **independent**.

Cognitively, this would mean that the two systems are separate, and there is no interaction.

“Linguistic Relativism”

These two ideas together are sometimes called **Linguistic Relativism**, because the idea is that thought (concepts, perception, etc) is relative to how one’s language works

Extreme idea 1: Language **determines** thought

Cognitively, this would mean that the system of language is the system of thought or somehow constrains the system of thought such that the combinatorics of thought are identical to the combinatorics of language.

In-between: Language **influences** thought

Cognitively, this would mean that the two systems are separate, but there is an interaction. The term influence is **vague**. We will need to make it precise with cognitive mechanisms in order to explore it.

Extreme idea 2: Language and thought are completely independent.

Cognitively, this would mean that the two systems are separate, and there is no interaction.

“Universalism”

This third idea is sometimes called **Universalism** because the idea is that all human thought is universal — we all conceptualize (and therefore experience) the world the same way.

Extreme idea 1: Language determines thought

Cognitively, this would mean that the system of language is the system of thought or somehow constrains the system of thought such that the combinatorics of thought are identical to the combinatorics of language.

In-between: Language influences thought

Cognitively, this would mean that the two systems are separate, but there is an interaction. The term influence is **vague**. We will need to make it precise with cognitive mechanisms in order to explore it.

Extreme idea 2: Language and thought are completely **independent**.

Cognitively, this would mean that the two systems are separate, and there is no interaction.

An outline for today

Extreme idea 1: Language **determines** thought

Nobody truly believes this any longer. But because it is the strongest form of the hypothesis it is worth exploring it a bit. This will help us to keep the theories clear when we evaluate claims. We can ask if the claim is about “determinism” or if it is about the weaker idea of “influence”.

In-between: Language **influences** thought

This is where all of the scientific debate is now. This is also the most complicated area of the hypothesis space because “influence” can mean any number of things. We will have to be precise in the mechanisms that we propose.

Extreme idea 2: Language and thought are completely **independent**.

Nobody truly believes this either. But it is important to keep this idea clear because all of the work is in teasing apart the parts of thought that are not influenced by language from the parts that are.

Scientific strategies to keep in mind

1. When thinking about these theories, we need to be clear about whether we are claiming that language “**determines**” or “**influences**” thought. These are distinct theories.
2. We need to precisely describe the **linguistic properties**. Superficial linguistic work or vague terms can imply connections that do not exist.
3. We need to precisely describe **the cognitive mechanism** that would allow language to have that impact.

And most importantly, we need to keep in mind that any claim about restrictions on the thoughts that human can think are **extraordinary claims**. Our **null hypothesis** should be that **all humans have the same cognitive capacity, the same experience of being human, etc.** Because we all share the same genes, we are the same species. We need to be cautious about claims that tend to exoticize other humans, even if we intend it as a sign of respect or interest in their culture. It is a slippery slope to stereotyping and racism.

The Sapir-Whorf Hypothesis:
Language **determines** thought

The Sapir-Whorf Hypothesis



Edward Sapir
(1884-1939)



Benjamin Whorf
(1897-1941)

The idea of “linguistic determinism”, that language determines thought, was first, and most forcefully, advocated by two anthropologists: Edward Sapir and his student Benjamin Whorf.

Therefore it is often called the **Sapir-Whorf Hypothesis.**

“Human beings do not live in the objective world alone, nor alone in the world of social activity as ordinarily understood, but are very much at the mercy of the particular language which has become the medium of expression for their society.” - Sapir

“We dissect nature along lines laid down by our native languages. The categories and types that we isolate from the world of phenomena we do not find there because they stare every observer in the face; on the contrary, the world is presented in a kaleidoscopic flux of impressions which has to be organized by our minds - and this means largely by the linguistic systems in our minds.” - Whorf

A myth to set aside: “no word for X”

There are so many listicles and blog posts dedicated to words that exist in one language but not another. Google it. There are tons. I am reading the English internet, so these will be focused on English as an anchor point:

Here are words that exist in other languages, but not English:

<https://ihworld.com/news-blog/ih-blog/10-words-that-don-t-exist-in-english/>

Datsuzoku (Japanese): To break from a daily routine, to perceive things differently

Hygge (Danish): A strong feeling of cosiness.

And here is someone claiming that there are words that exist in English, but not certain other languages:

<https://www.tomedes.com/translator-hub/english-words-translated.php>

Awkward - no equivalent in Italian?

Bully - no equivalent in Portuguese?

This is not determinism

It is easy to see that these are not cases of determinism — English speakers can still think about the concepts that are associated with these words.

You can see this because we can define them easily.

You can also see this because **we can learn the words!**

To learn a word, you must be able to conceptualize the meaning without the use of the word — otherwise you'd only be able to learn the phonetic form. You'd literally fail to learn the meaning, because you cannot think it. So, for any word that is ever learned, you can conclude that the learner was able to think of the meaning without using that word.

As a concrete example, the German word **schadenfreude** has no equivalent in English. But all English speakers are able to learn that word. So they must be able to conceive of the meaning without the word.

schadenfreude: enjoyment from observing someone's misfortune

The start of the Sapir-Whorf hypothesis: Hopi and time

Hopi is Native American language (Uto-Aztecan) spoken by the Hopi people who live in the southwest US:

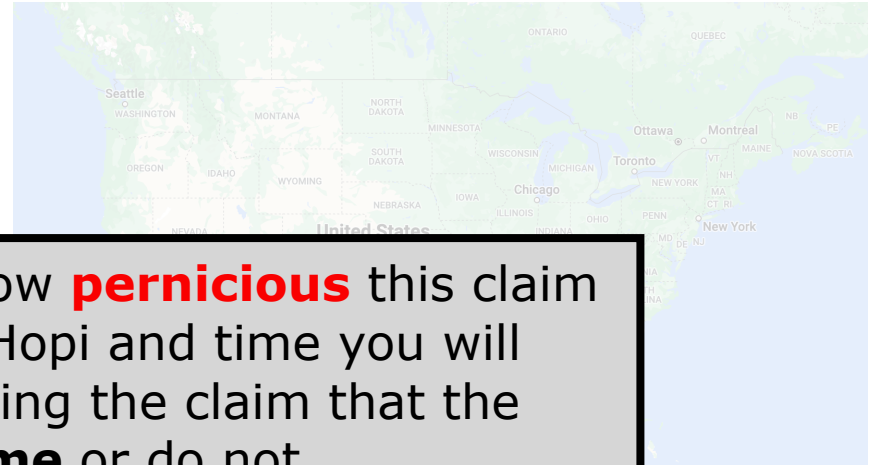


Benjamin Whorf
(1897-1941)

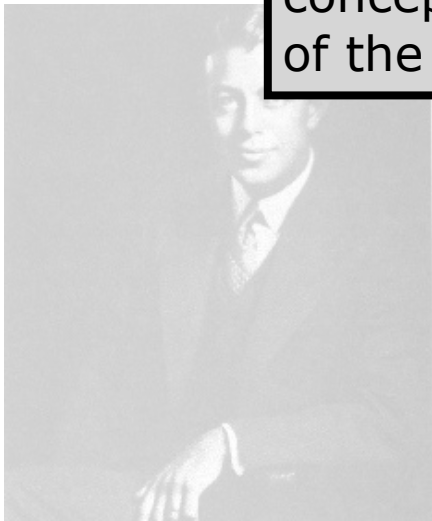
“I find it gratuitous to assume that a Hopi who knows only the Hopi language and the cultural ideas of his own society has the same notions, often supposed to be intuitions, of time and space as we have, and that are generally assumed to be universal. In particular he has no notion or intuition of time as a smooth flowing continuum in which everything in the universe proceeds at an equal rate, out of a future into a present and into a past After a long and careful analysis, the Hopi language is seen to contain no words, grammatical forms, construction or expressions that refer directly to what we call 'time', or to past, present or future ...”

The start of the Sapir-Whorf hypothesis: Hopi and time

Hopi is Native American language (Uto-Aztecan) spoken by the Hopi people who live in the southwest US:



It is difficult to convey just how **pernicious** this claim is. To this day, if you google Hopi and time you will find reputable sources repeating the claim that the Hopi **have no concept of time** or do not conceptualize time the same as other cultures because of the way their language works.



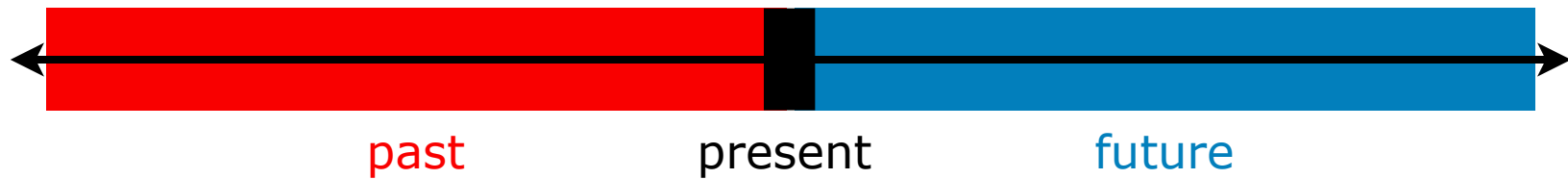
Benjamin Whorf
(1897-1941)

who knows only the Hopi language and the cultural ideas of his own society has the same notions, often supposed to be intuitions, of time and space as we have, and that are generally assumed to be universal. In particular he has no notion or intuition of time as a smooth flowing continuum in which everything in the universe proceeds at an equal rate, out of a future into a present and into a past After a long and careful analysis, the Hopi language is seen to contain no words, grammatical forms, construction or expressions that refer directly to what we call 'time', or to past, present or future ..."

To analyze Hopi, we need to understand how language convey time

There are at least three ways that languages convey time information:

- 1. Tense:** a property of verbs that tells us when in time the event described by the verb took place:



Some languages, like English, make a distinction between all three

English	past	present	future
cook	cook-ed	cook	will cook
jump	jump-ed	jump	will jump
speak	spoke	speak	will speak

To analyze Hopi, we need to understand how language convey time

But some languages, like Finnish, make a two-way distinction (past/non-past):

Finnish	past	present	future
I speak	puhu-i-n	puhun	puhun
You speak	puhu-i-t	puhut	puhut
He speaks	puhu-i	puhuu	puhuu
We speak	puhu-i-mme	puhumme	puhumme

Some languages, like Mandarin, don't make any tense distinctions:

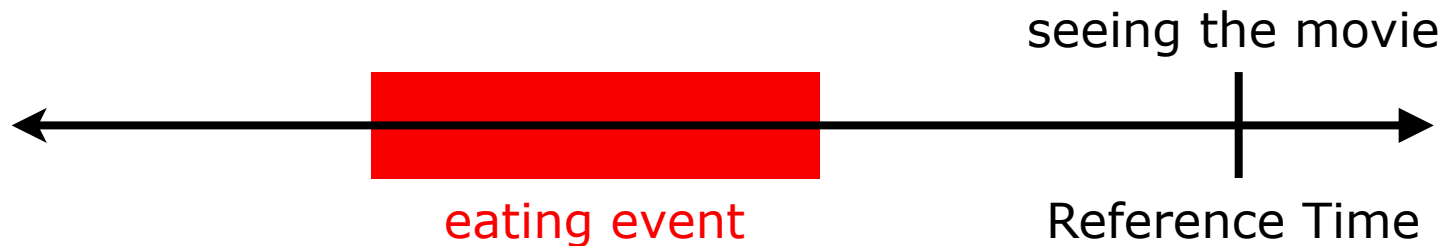
Mandarin	past	present	future
to eat	chi	chi	chi
to drink	he	he	he
to watch	kan	kan	kan

To analyze Hopi, we need to understand how language convey time

There are at least three ways that languages convey time information:

- 2. Aspect:** a property of verbs that tells us about how the event extends over time (completed, ongoing, etc). This is usually in reference to a specific time point (not necessarily the present) called reference time.

I **had eaten** when I saw the movie.



I **was eating** when I saw the movie.



To analyze Hopi, we need to understand how language convey time

There are at least three ways that languages convey time information:

3. Temporal words: These are lexical items that convey something about time independently of the verb.

In English, these tend to be adverbs, but we have to allow that there could be variability across languages in the exact syntactic category of these items. What is crucial is that they convey information about time:

yesterday, tomorrow, early, later, finally, meanwhile, again, then, etc.

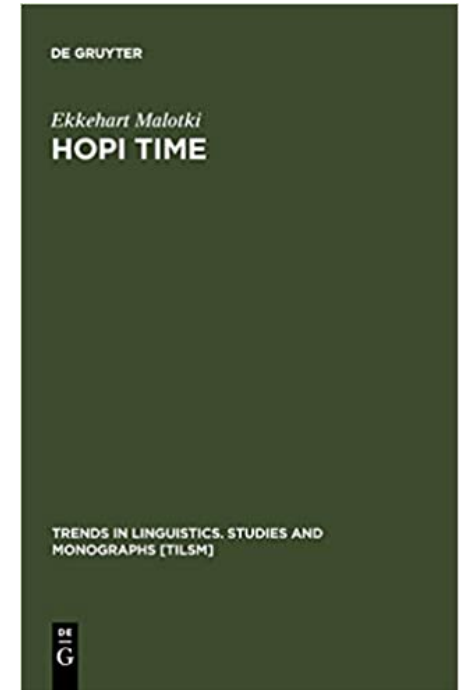
Whorf's claim was based sloppy linguistics

The quote we saw, plus many others in Whorf's (posthumous) writings, include claims that Hopi does not have **tense** or **temporal words**.

In 1983, a German linguist named Ekkehart Malotki published a **677 page book based on 4 years of fieldwork with the Hopi** dedicated exclusively to constructions in their language related to time, and dedicated to systematically refuting every claim that Whorf made about Hopi and time. It is truly an overwhelming piece of work:

The chapter titles alone tell us how bad this going to be:

1. The device of spatio-temporal metaphor (>200 pages)
2. Units of time (~200 pages)
- 3-6: Horizon-based sun time, Stellar orientation, The ceremonial calendar, Timekeeping devices
7. Pluralization and quantification of time expressions
8. Temporal particles (e.g., 'early', 'late', 'meanwhile'...) >70 pages
- 9.4. The Hopi tense system



It is clear that Hopi has words to describe time

This is just a screenshot of the table of contents, showing just the first half of chapter 8. There are lots of words describing time.

8	<i>Temporal particles</i>	531
8.0	Introduction	531
8.1	<i>aapiy</i> 'prematurely'	532
8.2	<i>angwu</i> 'beforehand/aforetime'	533
8.3	<i>ason</i> 'later/subsequently'	535
8.4	<i>haak</i> 'temporarily'	543
8.5	<i>iits</i> 'early'	547
8.6	<i>mòoti</i> 'at first'	550
8.7	<i>naat</i> 'still'	556
8.7.1	Simple <i>naat</i>	556
8.7.2	Negated <i>naat</i>	557
8.7.3	<i>naat pu</i> 'just'	559
8.7.4	<i>naat piw</i> 'again'	562
8.7.5	<i>naat suus</i> 'for the first time'	563
8.7.6	Variant <i>naato</i> and derivatives	563
8.8	<i>nawis'ew</i> 'finally'	564
8.9	<i>nen</i> 'and then'	565
8.10	<i>nit</i> 'and after that'	566
8.11	<i>nungwu</i> 'meanwhile'	568
8.12	<i>paapu</i> 'progressively'	570
8.13	<i>panis</i> 'constantly'	572
8.14	<i>pay</i> 'right now/already'	575
8.15	<i>pi'ép</i> 'repetitively'	578
8.16	<i>piw</i> 'again'	579



It is also clear that Hopi has aspect

Hopi actually has a fairly complex aspect system, with a number of different distinctions about events that are completed (perfective) and ongoing (imperfective):

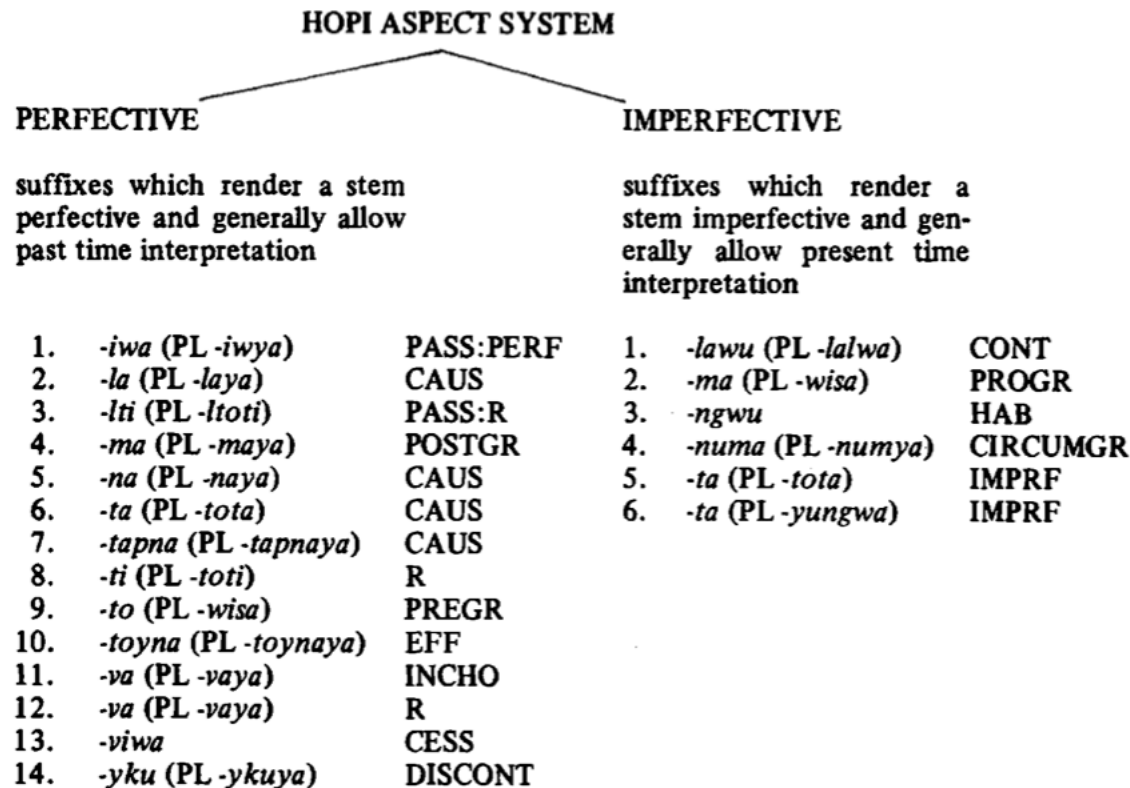
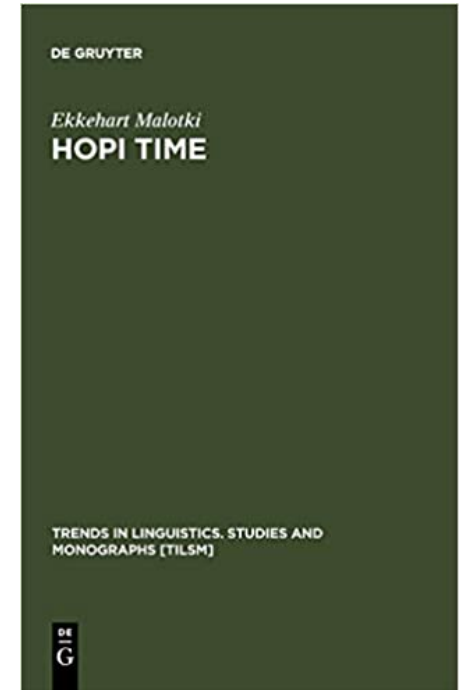


Figure 9



Whether Hopi has tense or not is unclear

There is a suffix **-ni** that Malotki calls future tense. If this is true, it would give Hopi a two-way distinction (non-future/future) similar to the way Finnish has a two-way distinction (past/non-past):

Hopi	past	present	future
to eat	nöösa	nöösa	nösni

But this could also be **mood**. Mood is a third piece of information carried by verbs. It tells us whether an event is **real** (indicative) or **imaginary** (subjunctive) as in things that are wished for, hoped for, hypothesized, etc. Mood is not about time, but it often correlates with the future, because the future is not real yet!

indicative:

subjunctive:

English: When I **was** happy.

If I **were** happy...

Persian: **mi**-nevisam

be-nevisam

"sees"

Spanish: habla**a**

hable

"speaks"

Is Tense required to convey time?

Earlier we saw that Mandarin has no tense distinctions on the verb. Shouldn't Whorf claim that Mandarin speakers have no conception of time?

Mandarin	past	present	future
to eat	chi	chi	chi

No, of course not. Mandarin has various temporal words to convey time, as well as **aspect** markers that can be used to infer temporal information because they refer to different reference times:

Wǒ shuāiduàn-**le** tuǐ

I break-**le** leg

(and it's still in a cast)

Wǒ shuāiduàn-**guo** tuǐ

I break-**guo** leg

(and it has healed since)

It is telling that Whorf did not make this claim about Mandarin. Mandarin was well understood by linguists and anthropologists (and all Mandarin speakers!!!) at the time. For Hopi, he was one of the only white people to have had experience with the language, and the Hopi community is very small (~5,000).

The current debate:
Language **influences** thought

Weakening the Sapir-Whorf Hypothesis

The Hopi controversy was fairly catastrophic for the first, strongest form of the Sapir-Whorf Hypothesis. But a much weaker form of the hypothesis has remained in the field:

Weak form: Language **influences** thought

Here is one such claim:

Speakers of languages with **grammatical gender** are more likely to conceptualize common nouns, like “key”, as having a **(human) gender** that corresponds with the noun’s grammatical gender.

Remember our lessons:

We need to precisely describe the **linguistic properties**. Superficial work or vague terms can imply connections that do not exist.

We need to precisely describe **the cognitive mechanism** that would allow language to have that impact.

What is grammatical gender?

The term “grammatical gender” is a bit misleading. At a linguistic level, we should probably call it **agreement class**.

Agreement is when one word in a sentence takes a specific form because of the properties of another word in the sentence.

Class just means a set of nouns that behave similar. For agreement, it means causing the same agreement on other words. Here is an example from Spanish:

The **class** of the noun determines agreement with the **article** and with the **adjective**.

class 1

La bicicleta es cara.

The bike is expensive.

La revista es cara. (magazine)

La foto es cara. (photo)

La llave es cara. (key)

class 2

El carro es caro.

The car is expensive.

El libro es caro. (book)

El pastel es caro. (dessert)

El mapa es caro. (map)

How many classes do languages have?

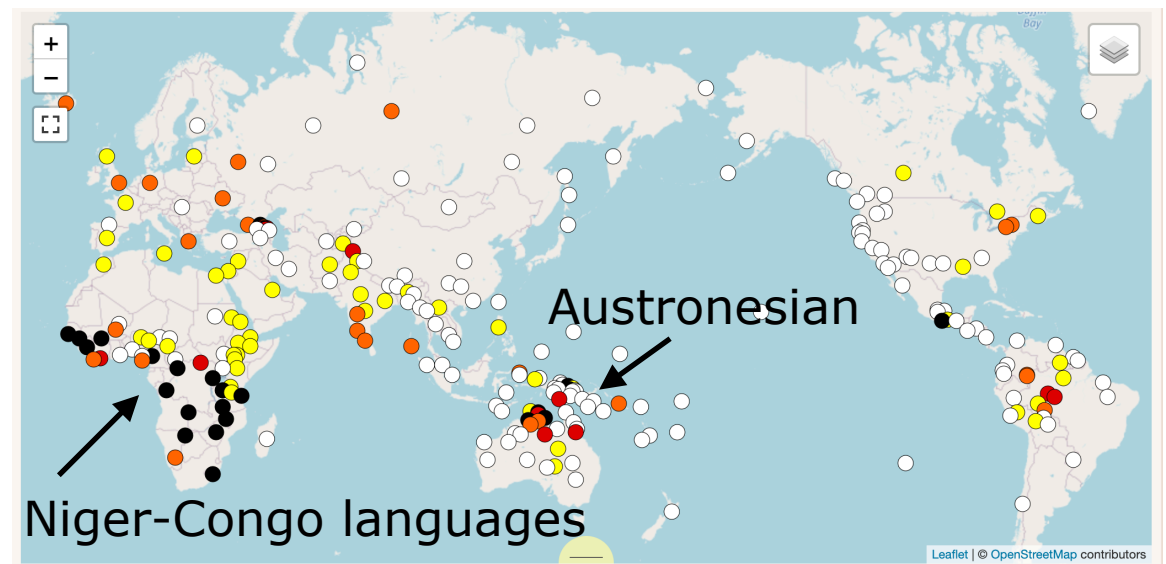
The World Atlas of Language Structures is a database and website put together by linguists to collate what we know about cross-linguistic variation of many different phenomena. There are several chapters on grammatical gender.

<https://wals.info/chapter/30>



Currently, WALS has a sample of 257 languages that are characterized based on whether they have grammatical gender (i.e., agreement classes) and how many classes they have. WALS also provides a map so we can see the distribution:

Value	Representation
○ None	145
● Two	50
● Three	26
● Four	12
● Five or more	24
Total:	257



Three independent concepts of gender

I am about to begin to discuss the link between agreement classes (grammatical gender) and concepts about human gender

Human gender is a complex topic of study. To respect this, in language science, we try to distinguish three concepts:

Grammatical “gender”: These are **agreement classes**. It is a linguistic property.

Conceptual gender: This is a **semantic** (conceptual/semantic) property. It is the property that is used by a perceiver to **classify** a referent.

Biosocial gender: This is **gender identity** as experienced by a human (about themselves).

These three concepts are **independent**, meaning they can freely vary. Crucially, this means that the values that each can take are independent of the others. For example, the fact that “grammatical gender” in a language may only take two values does not imply that biosocial gender only takes two values.

So what defines the classes?

There appears to be five ways that languages define agreement classes:

1. **Semantics of animacy and other semantic categories**

In this type of language, all nouns are categorized based on how animate they are. Exactly what this means varies by language. Roughly speaking, animacy corresponds with the semantic concepts of agency, mobility, power-over-the-environment, etc. Humans (and deities) are always the most animate.

Fulfulde, a set of Niger-Congo languages spoken in over 20 countries, has 20+ classes depending on the specific language (commonly called dialect)!

There is a category for humans, one for large animals, and then a bunch for different semantic categories of things.

Crucially, human gender is not part of the system. All humans are in the same class!

So what defines the classes?

There appears to be five ways that languages define agreement classes:

2. Semantics of conceptual gender (only humans)

In this type of language, only human nouns (man, woman, boy, girl, etc) are placed into agreement classes, very obviously based on values of conceptual (human) gender.

Two classes:	Kala Lagaw Ya (Australia)	Masculine: all male humans Feminine: all other nouns
	Diyari (Australia)	Feminine: all female humans Masculine: all other nouns
Three classes:	Kannada (India)	Feminine: all female humans Masculine: all male humans Neuter: all other nouns

So what defines the classes?

There appears to be five ways that languages define agreement classes:

3. Semantics of conceptual gender for human nouns and other semantic categories

In this type of language, human nouns (man, woman, boy, girl, etc) are placed into agreement classes, very obviously based on two values of conceptual (human) gender. Other objects can be placed in these categories, or in other categories that are defined some other way!

Bininj Gun-Wok
(Australian)

Masculine: male humans, animate entities, rain, compass points, honey...

Feminine: female humans, female animates, sun

Vegetable: plants, some body parts, songs, fire, food, boats/cars, drink, water...

Neuter: animal parts, weather, sea, time, language, some landscape features...

So what defines the classes?

There appears to be five ways that languages define agreement classes:

4. **Semantics of conceptual gender for human nouns, and phonology for non-human nouns.**

In this type of language, human nouns (man, woman, boy, girl, etc) are placed into agreement classes, roughly based on two values of conceptual (human) gender (with some strange nouns here or there). Non-human nouns are placed into these categories as well, but based on their [phonological properties](#).

Qafar (Ethiopia and Djibouti): accented vowel are feminine (catò - help), and all others are masculine. (gilàl - winter, tàmu - taste)

Spanish nouns ending in -o are almost always masculine, and nouns ending in -a are almost always feminine. There are some exceptions. Nouns with other endings show sub-patterns.

French is sometimes claimed to be arbitrary, but it has many phonological components. For example, of 938 nouns ending in /ɛ̃/, 99% are masculine.

So what defines the classes?

There appears to be five ways that languages define agreement classes:

5. Semantics of conceptual gender for human nouns, and morphology for non-human nouns.

In this type of language, human nouns (man, woman, boy, girl, etc) are placed into agreement classes, roughly based on two values of conceptual (human) gender (with some strange nouns here or there). Non-human nouns are placed into these categories as well, but based on their [morphological properties](#).

Russian:

Masculine

Feminine

Neuter

Each row has similar semantics, so this shows that it is not semantic!

žurnal (magazine)

gazeta (newspaper)

pis'mo (letter)

lokot' (elbow)

kost (bone)

koleno (knee)

klarnet (clarinet)

skripka (violin)

fortep'jano (piano)

kostjum (suit)

jupka (skirt)

plat'e (dress)

So what defines the classes?

There appears to be five ways that languages define agreement classes:

6. Semantics of conceptual gender for human nouns, and morphology for non-human nouns.

In this type of language, human nouns (man, woman, boy, girl, etc) are placed into agreement classes, roughly based on two values of conceptual (human) gender (with some strange nouns here or there). Non-human nouns are placed into these categories as well, but based on their morphological properties.

Russian: Gender is based on morphological variations (based on something called "case")

	Masculine	Feminine		Neuter
	I	II	III	IV
NOMINATIVE	žurnal	gazeta	kost'	pis'mo
ACCUSATIVE	žurnal	gazetu	kost'	pis'mo
GENITIVE	žurnala	gazety	kosti	pis'ma
DATIVE	žurnalu	gazete	kosti	pis'mu
INSTRUMENTAL	žurnalom	gazetoj	kost'ju	pis'mom
LOCATIVE	žurnale	gazete	kosti	pis'me
gloss	'magazine'	'newspaper'	'bone'	'letter'

6 types of languages, and we can see why traditional grammarians called it “gender”

1. **No** agreement classes
2. Semantics of **animacy** and other semantic categories.
3. Semantics of **conceptual gender** for humans nouns only
4. Semantics of **conceptual gender** for human nouns and other semantic categories
5. Semantics of **conceptual gender** for human nouns, and phonology for non-human nouns.
6. Semantics of **conceptual gender** for human nouns, and morphology for non-human nouns.

But we need to notice that this claim is only about languages in type 5 and 6 where common nouns are placed into gender-based classes based on their formal linguistic properties (phonology and morphology).

When you read about these, they are sometimes grouped this way

1. No agreement classes

2. Semantics of animacy and other semantic categories.

3. Semantics of conceptual gender for human nouns only

4. Semantics of conceptual gender for human nouns and other semantic categories

5. Semantics of conceptual gender for human nouns, and phonology for non-human nouns.

6. Semantics of conceptual gender for human nouns, and morphology for non-human nouns.

The idea is that languages can have no gender, they can have semantic-only gender, or they can have semantic plus formal gender. I just want you to know this in case you start reading the literature!

What is the cognitive claim?

Earlier, I wrote it like this:

Speakers of languages with **grammatical gender** are more likely to conceptualize common nouns, like “key”, as having a **(human) gender** that corresponds with its grammatical gender.

But now we can see this is not quite correct:

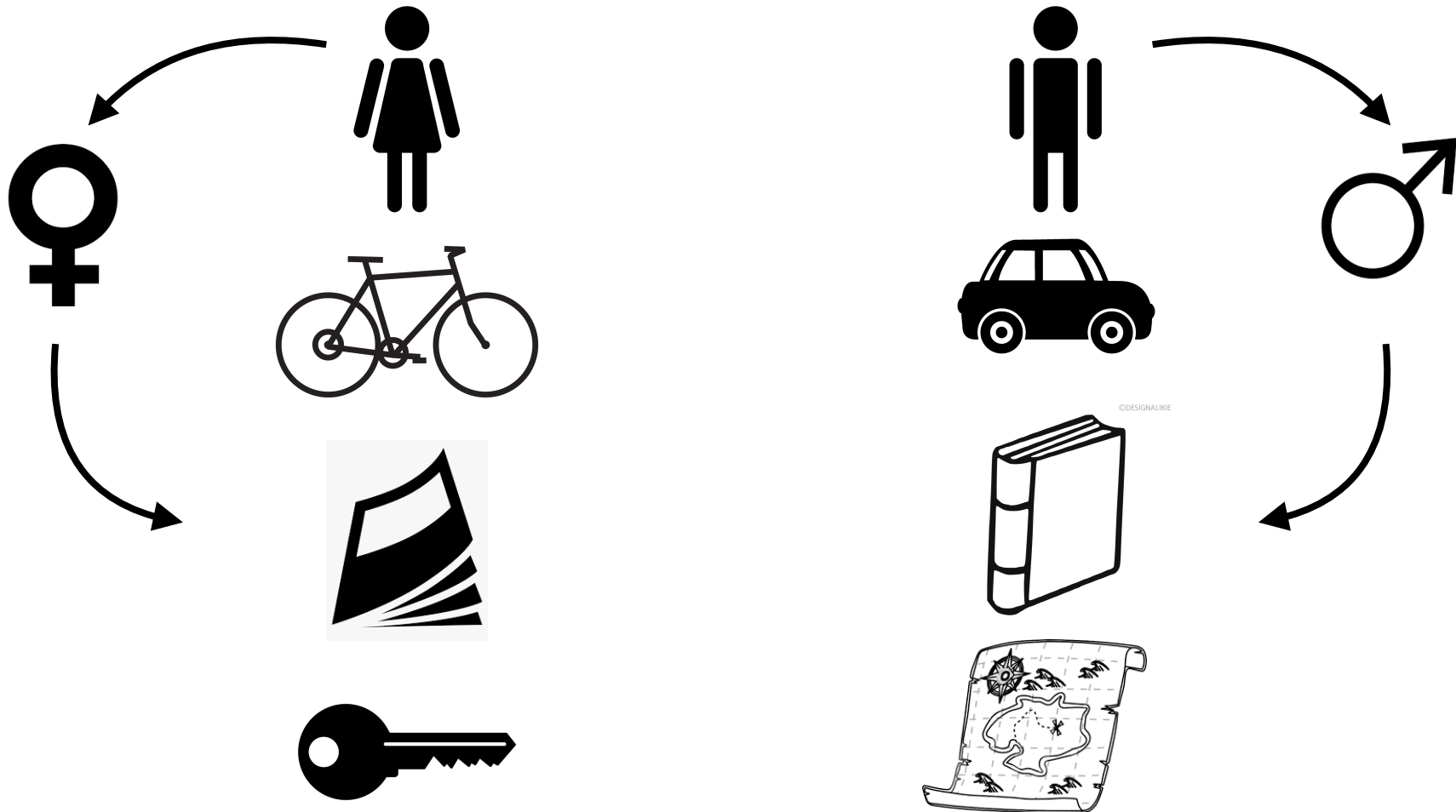
Speakers of languages with **agreement classes based on both human gender for humans and phonology or morphology for common nouns** are more likely to conceptualize common nouns, like “key”, as having the **(human) gender of the human nouns that are also in the class.**

Notice this claim can't apply to human gender only systems, because common nouns are classed with them. And it can't apply to animacy systems, because they don't involve human gender at all.

What is the cognitive mechanism for this?

Option 1: the semantics “rubs-off”

This is the claim that the **concept** of “key” is changed by the fact that its phonology puts it in the same class as humans with a certain gender.

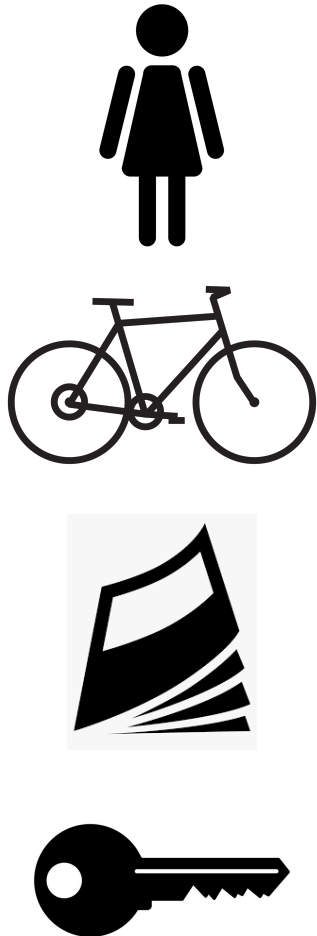


What is the cognitive mechanism for this?

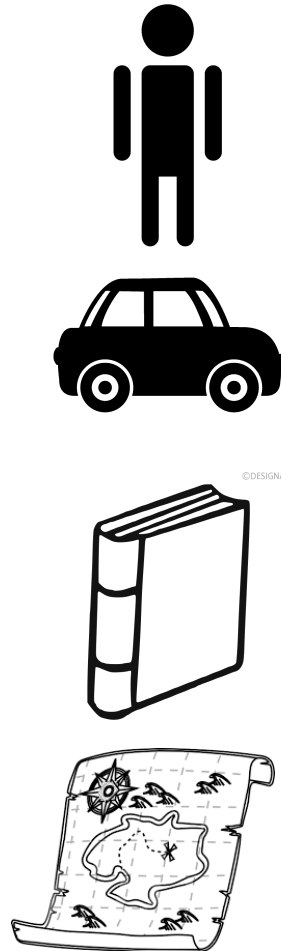
Option 2: It is the **labels** that is driving this!

This is the claim that humans, particularly adults, know that we call the noun classes "masculine" and "feminine". So we create an **association**.

"feminine"



"masculine"



But this does not change the fundamental **concept** — "key" doesn't get a conceptual gender. We just "know" that the idea of "gender" can be applied to these words.

Statistical associations are real. We have a statistical association between **ice cream** and **hot days**! But that doesn't mean that the concept of ice cream is altered by the concept of a hot day, or that the concept of a hot day is altered by ice cream.

How do we test these?

There are a number of tasks that people have attempted. Here are two:

Properties: Participants are shown words or images of common nouns, and asked to list adjectives to describe the noun. Those adjectives are then rated (by other participants) as masculine or feminine.

Voice change: Participants are shown an object, and asked to create a simulated voice for the object. Other participants then rate the voice as masculine or feminine.

These two tasks differ in how explicit the idea of conceptual gender is in the task. For adjective generation, there is no explicit use of gender. But for voices, there is (under the assumption that imaginary individuals who can speak will be given a conceptual gender).

This implicit/explicit difference has been shown to matter for the results...

What do we find?

A recent meta-study looked at 43 articles covering nearly 6,000 participants.

Psychonomic Bulletin & Review (2019) 26:1767–1786
<https://doi.org/10.3758/s13423-019-01652-3>

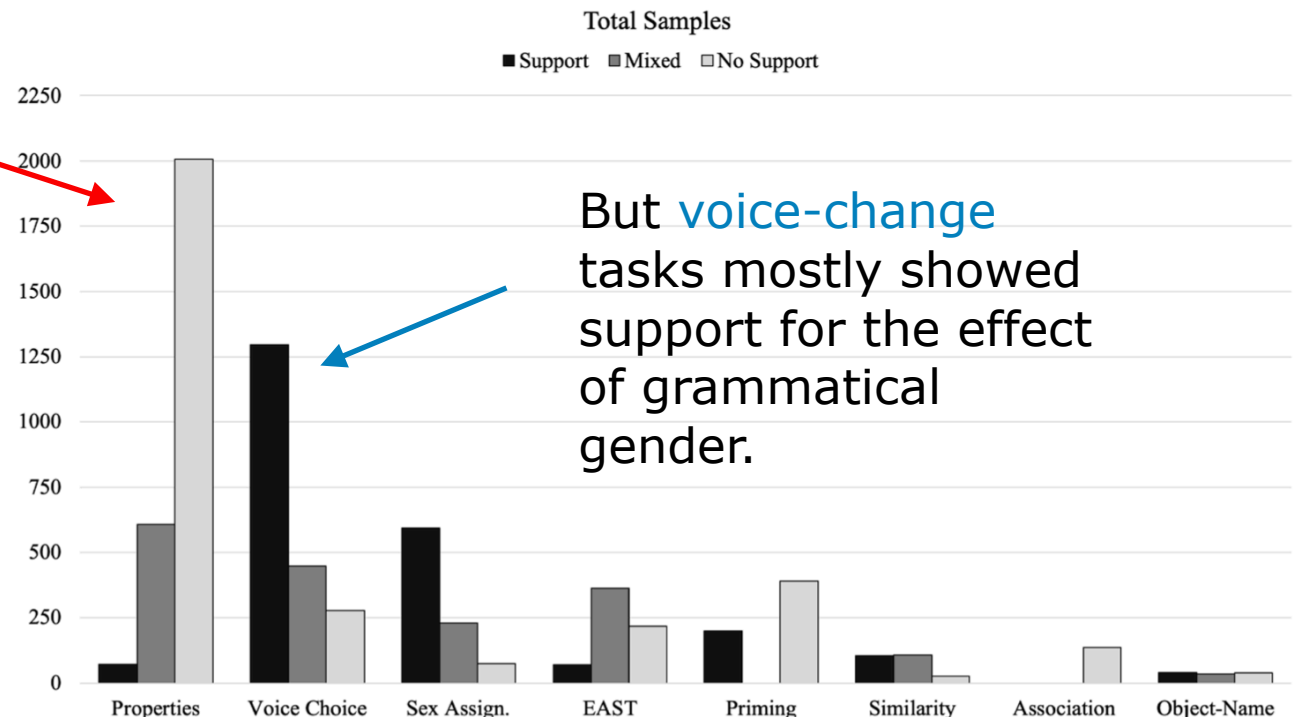
THEORETICAL REVIEW



Grammatical gender and linguistic relativity: A systematic review

Steven Samuel¹ · Geoff Cole¹ · Madeline J. Eacott¹

They found that **properties** tasks mostly showed no support for the effect of grammatical gender.



But **voice-change** tasks mostly showed support for the effect of grammatical gender.

Fig. 2 Classifications of support, mixed support, and no support by task type. The total number of samples is shown on the y-axis.

What is the cognitive mechanism for this?

Option 1: the semantics “rubs-off”

This is the claim that the **concept** of “key” is changed by the fact that its phonology puts it in the same class as humans with a certain gender.

Option 2: It is the **labels** that is driving this!

This is the claim that humans, particularly adults, know that we call the noun classes “masculine” and “feminine”. So we create an **association**.

The authors of the meta-study argue that option 2 is the most consistent with the idea that only explicit gender tasks show an effect. The idea is that explicit gender tasks engage conscious knowledge of the labels (and their associations). Implicit tasks do not engage that knowledge. They argue that option 1 should predict affects in both tasks.

What can we conclude?

This topic is still debated. But here's where we seem to stand:

Grammatical gender is probably misnamed. Adding precision about the phenomenon (agreement classes) and types of systems (semantic, phonological, morphological) can help clarify what we are testing.

Grammatical gender is very complicated. The claims in this area are for only a portion of the grammatical gender systems, and relatively few languages have been tested.

There are at least two mechanisms that could lead to an interaction between language and thought — the stronger is that the concept changes, the weaker is that we experience statistical associations, perhaps with the labels.

Current studies suggest it is the weaker claim - no change in the concepts, just an association, possibly conscious.

So what do we think about these?

We can see some interesting similarities: both language and thought have primitive units, and both have rules for combining those units into larger and larger complex objects.

Extreme idea 1: Language **determines** thought

Cognitively, this would mean that the system of language is the system of thought or somehow constrains the system of thought such that the combinatorics of thought are identical to the combinatorics of language.

In-between: Language **influences** thought

Cognitively, this would mean that the two systems are separate, but there is an interaction. The term influence is **vague**. We will need to make it precise with cognitive mechanisms in order to explore it.

Extreme idea 2: Language and thought are completely **independent**.

Cognitively, this would mean that the two systems are separate, and there is no interaction.