

What Cannot Cross

On what grammar obliges a speaker to say, and what goes missing when it stops obliging

AICE Library · what-cannot-cross

11 chapters · single-language edition

Every number, name and year in this volume is traceable to a cited source.

01 • What Grammar Makes You Say

In June 1997, Kanchana Naksakul wrote in the Royal Institute Newsletter about the Thai dubbing of foreign films, in which a father calls himself **ผม** and calls his own child, not yet ten years old, **คุณ**.

Not one word of that dub was mistranslated. I is **ผม**, you is **คุณ**, every word is in the dictionary. What was wrong was the relationship. Kanchana wrote that it grates on the ear, because no father in Thailand speaks to his own child that way. That pair of pronouns announces two people who are distant and formal with one another. A father and his nine-year-old cannot stand in that relation.

The error did not begin in the dubbing booth. It began in the source. The English sentence handed over for translation never said how close these two people were, which of them stood higher, or whether the speaker was a man or a woman, because I and you oblige none of it. Thai obliges all three. So the translator had to supply what the source had never carried, and supplied it wrongly.

This is the spine of the book. **Translation does not lose words. It loses obligations.** Every language forces its speaker to specify certain things and lets others stay vague. The two lists have never matched between any pair of languages. What goes missing in the crossing is whatever sat on the forced list. On the far side it becomes something you may put in or leave out, and a thing you may leave out is, in the end, left out.

The obligation is at its plainest in Tariana, an Arawak language of the northwest Amazon. Alexandra Aikhenvald did the fieldwork herself and records in *A Grammar of Tariana* that almost every finite verb has to carry an evidential – the morpheme stating where the speaker got what he is saying. There are five to choose from, and they are fused with tense. Take it off and the sentence is ungrammatical at once. You cannot simply say "José played football." You have to commit, in the same breath, to how you know.

English has no such slot. "José played football" is complete, grammatical, and perfectly silent about where the information came from. Crossing from Tariana into English, the morpheme meaning that someone else told me has to turn into a phrase — apparently, it is said, I hear that, allegedly. Aikhenvald names this class of words as the standard way out. The trouble is not that the meaning distorts. The trouble is the status. On one side, a morpheme that cannot be left off; on the other, an adverb that can be cut at any moment. The next editor deletes one word, apparently, the sentence reads exactly as it did before, and what the source grammar obliged is gone with nobody noticing.

Turkish compresses the same thing into two morphemes. -DI is used of a past the speaker saw with his own eyes; -mİş of a past he inferred or heard from someone else. Two sentences differ at that one point and nowhere else, and Turkish children pick the distinction up early, on the foundational description by Aksu-Koç and Slobin of 1986. Cuzco Quechua has three: -mi for personal knowledge, -si for hearsay, -chá for conjecture. English needs a whole clause to say what these languages say in a syllable.

Thai obliges a different category, and obliges it harder than Thai speakers notice. English eat is one verb. Thai has four rungs. กิน is the bare one, **รับประทาน** the polite one, **ฉัน** is used of monks, **เสวย** of the King and high royalty. Picking the wrong rung is not a lapse of style. It becomes disrespect, or a formality so misplaced it is comic. The Thai sentence **ท่านเสวยแล้ว** states the subject's standing inside the verb itself. "He ate" cannot. It needs another clause bolted on.

On the second-person side the forms climb a four-rung ladder, and Kanchana sets all four out in the 1997 essay: **ใต้เท้า** for a respected superior, **ใต้ฝ่าพระบาท** for a Mom Chao, **ใต้ฝ่าละอองพระบาท** for Phra Ong Chao rank and above, **ใต้ฝ่าละอองธุลีพระบาท** for the King. The four are not interchangeable. Each pins the addressee's exact rank inside the word. You pins nothing, and cannot be made to without inserting a title beside it. **ครับ** and **ค่ะ** are the same story. Most polite Thai sentences cannot do without one, and it reports the speaker's own gender and the formality of

the utterance at the same time. Eric McCready built a compositional semantics for this system in 2014 and reached the same conclusion: Thai makes obligatory structure of what English leaves outside its grammar altogether.

"Untranslatable words" is therefore the wrong frame. A gloss is always available. **รับประทาน** can be explained in a single paragraph. The Office of the Royal Society explained it itself, on the radio programme **รู้ รัก ภาษา ไทย** broadcast on 30 October 2009: the word was originally two, **รับ** and **ประทาน**, used when a Prince or Princess of Phra Ong Chao or Mom Chao rank gives something to a commoner – and the programme instructs speakers to pause between the two syllables when using it in the **ราชาศัพท์** (royal register) sense, so that it does not blur into the single word that just means "eat." That is the whole explanation. Anyone may drop it in a footnote.

What cannot be done is to un-oblige the obligation. A gloss is an addition, and additions are always possible. An obligation is a slot that has to be filled, and a language without the slot cannot grow one. An English sentence about eating holds no space that forces rank to be named. You may put it in; nothing requires it; and nothing catches when it is left out. The distance between "may" and "must" is the whole of this book.

A line has to be drawn here, and drawn clearly. Nobody disputes the grammatical facts above. Whether grammar of this kind changes what a speaker remembers or thinks is a separate question, and it is unanswered. Tosun, Vaid and Geraci in 2013 reported that Turkish speakers remembered assertions marked as firsthand better, and were the more likely to misremember what they had been told as something they had seen. Ünal, Pinto, Bungler and Papafragou in 2016 ran it again with a task using no language at all, pictures in place of words, and found Turkish and English speakers equally prone to the same confusion. The two studies answer in opposite directions, and the research behind this book could not find a third that settles them.

There is a further study, "Lost in translation, apparently," in *Bilingualism: Language and Cognition*, testing this very crossing point directly on

Turkish-English bilinguals. This research confirmed its title and its journal and nothing else; the figures inside were not read. No number from it appears anywhere in this book. The line this book stands on is narrow, and narrow by design. It claims what a text loses in the crossing. It does not claim what a speaker's mind becomes.

There is one test for entry into this book. The case must carry a date, and its source must resolve to something real when opened — not an example that has been passed along. One chapter here exists to dismantle the most-cited evidence for the book's own thesis. The count of Eskimo words for snow swelled through repeated citation with nobody going back to count. Laura Martin traced that chain of citation in 1986; Geoffrey Pullum wrote it up further in 1989. The rule applies to the whole volume: a claim that flatters our own hypothesis is checked harder than one that contradicts it, not more gently.

This is the closing volume of a generation of books from this workshop. The previous generation wrote English first and rendered it into Thai afterwards. The Thai that came out was correct in every word and lifeless. The reason sits in the fourteen paragraphs I have just written. If the source is English, the sentence is obliged to specify only what English obliges, and the Thai walking along behind it can carry no more than that. The layers Thai does oblige — rank, distance, deference, the speaker's own gender — had nothing in the source demanding them, so the translator set a safe middle level in every sentence. The result was a book in which nobody stands anywhere.

This generation writes Thai first, deliberately, and the reason is not affection for the language. It is instrumentation. This workshop has a measure of Thai quality. It has no measure of English quality yet. So the side that can be measured is written first, and the other side follows a contract set down in writing. The next generation is to be beautiful in both languages at once. That sentence is a specification, not a wish, and its condition is already plain: the English-side instrument has to exist first.

Ünal, Pinto, Bungler & Papafragou 2016, <https://www.sciencedirect.com/science/article/abs/pii/S0749596X15001266> · "Lost in translation, apparently," *Bilingualism: Language and Cognition* (title and journal confirmed only; figures not checked), <https://www.cambridge.org/core/journals/bilingualism-language-and-cognition/article/abs/lost-in-translation-apparently-bilingual-language-processing-of-evidentiality-in-a-turkishenglish-translation-and-judgment-task/94BBB0151AB2214E72DFFE1BB71B9AF6> · Laura Martin 1986, <https://anthrosource.onlinelibrary.wiley.com/doi/abs/10.1525/aa.1986.88.2.02a00080> · Geoffrey Pullum 1989, <https://pullum.ppls.ed.ac.uk/EskimoHoax.pdf>

02 • That Young Woman

Sometime in the early-to-mid third century BCE, in Alexandria, a group of Jewish translators read the Hebrew word *almah* at Isaiah 7:14 and wrote it down in Greek as *parthenos*.

The two words do not contradict each other, and neither is a mistranslation of the other. What changed is the conclusion a reader was obliged to draw. The Hebrew obliged a conclusion about age. The Greek obliged something narrower. The gap between those two obligations is where two thousand years of doctrine came to sit.

The Hebrew of Isaiah 7:14, in the Masoretic Text, reads הַעֲלֻמָּה (*ha-almah*) — "that young woman," in the sense of a woman of childbearing age. Scholarly consensus (Sweeney 1996, Coogan 2007) holds that *almah* marks age and "has nothing to do with virginity." Hebrew already had a specific word for a virgin: *betulah*. The writer of Isaiah did not reach for it.

That does not mean Isaiah was saying she was not a virgin. It means the sentence does not touch the question at all. *almah* leaves it open. A Hebrew reader could conclude either way; the text compels neither.

parthenos is narrower. The Septuagint — the Hebrew Bible translated into Greek, made by Jews in Alexandria — chose a word meaning virgin. A young woman of childbearing age who has never married is an *almah* and a *parthenos* at the same time. So the Greek did not speak against the Hebrew. It closed a door the Hebrew had left open.

This is where the loss happens. A Hebrew reader was not obliged to conclude anything about virginity; a Greek reader could not avoid it. Nobody lied. No word went missing either. What went missing was the silence, and silence has nowhere to stand in a target language, because every word you pick up and put down obliges something. A translator chooses what to oblige. Obliging nothing at all is not one of the choices.

The scholarship does not agree here, and this book prints both sides. The Eerdmans Commentary on the Bible (Saldarini 2001, p. 1007) calls parthenos a mistranslation of almah, flatly. This book goes the other way. A mistranslation asserts what the source denies, and the Hebrew denies nothing at that point. So the real disagreement is over whether narrowing counts as an error, and that question is not closed.

The doctrinal weight came to rest on the Greek wording, not the Hebrew. Matthew 1:23 quotes the parthenos wording to present Jesus as born in fulfilment of the Isaiah prophecy, and Matthew 1:25 underlines that Joseph did not sleep with Mary before she gave birth. That is the textual basis of the Christian doctrine of a virgin birth.

The full Septuagint text read today does not come from the third century BCE. The oldest known Greek Bible manuscript before the Dead Sea Scrolls were found is Papyrus Rylands 458, a Pentateuch fragment, dated palaeographically to around the mid-second century BCE. The complete text survives to the present through three great uncial codices — Codex Vaticanus (4th century CE), Codex Sinaiticus (mid-4th century CE), and Codex Alexandrinus (5th century CE).

The account of who did the translating is later than the translation. The tradition descends from the Letter of Aristeas, which tells of seventy-two translators sent from Jerusalem, working in Alexandria under the Ptolemaic dynasty. But the Letter is itself a separate composition; most scholars date it to the second century BCE, later than the events it describes.

An Eastern Orthodox tradition holds that Simeon, one of the seventy translators, doubted the word parthenos and was told from above that he would not die until he saw it fulfilled. The story is used to explain Luke 2:26. This book's research file marks it explicitly as church tradition, not verifiable history, and this book prints it as that.

In 1952 the choice became a public argument again. The translators of the Revised Standard Version rendered almah as "young woman," and a controversy broke out among conservative American Christians. One

pastor publicly burned a copy of the RSV (Metzger 2001, p. 120). No new manuscript. No new evidence. One English word had changed.

Translations are still split today. The versions widely used among evangelicals, NIV and ESV, keep "virgin"; most other modern translations use "young woman." The dispute is not about the Hebrew — both sides read the same word. It is about which obligation English is going to carry.

What this entry claims is narrow. It does not claim anyone mistranslated. It does not claim the doctrine is right or wrong. It claims that one Hebrew word left a question open, one Greek word closed it, and that a target language has no empty space for an open question. The translators in Alexandria in the third century BCE were not choosing between right and wrong. They were choosing between two obligations, and they had to choose.

What could not cross: The silence of *almah* could not cross. Hebrew already had *betulah* for a virgin and did not use it, so the text obliged a conclusion about age and left status blank. Greek had to pick a word that obliged something, and once it was picked, that blank no longer existed.

Overtaken if: Someone shows the question of virginity was never open — either by showing that *almah* in Hebrew already obliged it, against Sweeney 1996 and Coogan 2007, or by showing that third-century BCE Greek had a word that obliged age alone and left status blank the way *almah* does, and that the translators passed it over. If such a word existed, *parthenos* was an option rather than an obligation, and this entry is wrong.

Sources: Isaiah 7:14 — https://en.wikipedia.org/wiki/Isaiah_7:14 · Letter of Aristeas — https://en.wikipedia.org/wiki/Letter_of_Aristeas · Papyrus Rylands 458 — https://en.wikipedia.org/wiki/Papyrus_Rylands_458 · Codex Vaticanus — https://en.wikipedia.org/wiki/Codex_Vaticanus · Codex Sinaiticus — https://en.wikipedia.org/wiki/Codex_Sinaiticus (these pages are the source of every attribution carried in this entry — Sweeney 1996, Coogan 2007, Saldarini 2001 p. 1007)

in the Eerdmans Commentary on the Bible, and Metzger 2001 p. 120; this book did not open those source volumes itself)

03 • Two Texts, One Country

On 6 February 1840 at Waitangi, Māori chiefs — rangatira — signed alongside William Hobson, Britain's representative, two documents that did not say the same thing. Both sides signed as though they did. The number who signed that day is recorded as a range, 43 to 46, not a single figure.

Article 1 of the English text states that the chiefs ceded to the Queen "all the rights and powers of sovereignty" over their land. Article 1 of the Māori text, known as te Tiriti, states that the chiefs ceded "te kawanatanga katoa" — the complete government. The whole of this chapter sits between those two lines.

Choosing the word kawanatanga was not carelessness. It was the word that actually existed that day. Kawanatanga is a transliteration of governance and governor, and Māori already knew the word kawana — from the translated Bible, and from the kawana of neighbouring New South Wales. Sovereignty, meanwhile, had no direct equivalent in Māori at the time. The translator was therefore not choosing between a right word and a wrong one. He chose the closest word the language had to offer.

Article 2 of the Māori text widens the gap further. There the chiefs were guaranteed "te tino rangatiratanga" over their lands, villages and treasures — the unqualified exercise of chieftainship. One text therefore says supreme authority has been handed over. The other says government has been handed over, then guarantees chiefly authority back in the very next article.

Read the English text to the end and what you have is sovereignty changing hands. Read the Māori text to the end and what you have is one governor, and chiefs who are still chiefs. Both readings are correct against the text each reader was reading.

Before the first signing there was real discussion. On the night of 5 February 1840, the missionary Henry Williams and the assembled chiefs discussed the Treaty's meaning through the night. The signing then went on for two further rounds: 10 February at Waimate, six chiefs; 12 February at Mangungu, in Hokianga, sixty-four more. Hobson signed in person on each occasion. The gap between the two texts survived the night of debate, and it survived two further signings.

New Zealand's own government agency puts this plainly. The official history site of the Ministry for Culture and Heritage states: "It is widely accepted that the use of the words 'kawanatanga' and 'tino rangatiratanga' ... contributed to later differences of view between the Crown and Māori over how much authority the chiefs would retain and how much the governor would have." That sentence accuses nobody. It identifies two words as the cause.

On 14 November 2014 the Waitangi Tribunal ruled on this in writing. The Tribunal is the standing New Zealand government body empowered to determine the Treaty's meaning, and the Stage 1 report of its Te Paparahi o Te Raki inquiry, in Northland, concluded: "The rangatira who signed te Tiriti o Waitangi in February 1840 did not cede their sovereignty to Britain. That is, they did not cede authority to make and enforce law over their people or their territories."

The Tribunal also set out how the British side had explained the Treaty. Hobson and his agents explained it as granting Britain "the power to control British subjects and thereby to protect Māori", and never explained the full sovereignty claim standing in the English text. The finding that follows is that the Crown "did not acquire that sovereignty through an informed cession by the rangatira who signed te Tiriti".

That ruling does not say anyone lied. The word carrying the weight is informed. What failed to happen was a cession of sovereignty made in knowledge of what was being ceded — not a deception. The record this chapter works from states what the British side explained, and what it did not explain. It does not state whether the omission was deliberate, and this chapter does not supply that answer.

The interval is the part to read slowly. Between the signings of February 1840 and the ruling of 14 November 2014 lie 174 years. Throughout that span both documents existed side by side, both legible, and each side read its own text exactly as written. The gap was never hidden. It stayed open in plain view.

This case is the strongest in the book for one reason: a standing arm of government wrote a finding about a translation gap down on paper, with a date on it. The other cases here have to infer the consequence from what survives. This one does not. The record states it.

The strongest counter-reading concerns the scope of the ruling itself. It is a Stage 1 report from a single district inquiry, Te Paparahi o Te Raki, and it speaks about the chiefs who signed in February 1840. The record this chapter works from does not say how the New Zealand government responded, and does not record any consequence for today's constitutional position – so this chapter does not speak for it. What the ruling establishes is one thing: the gap between the two texts is real, and it mattered enough for the state to come down and decide it.

What failed to cross on 6 February 1840 was not a word. Words could be found: sovereignty had *kawanatanga* to stand in for it. What failed to cross was the obligation the source word imposed on its reader's conclusion. Both sets of signatories read their own document, drew the correct conclusion, and drew different ones.

What could not cross: Sovereignty in the sense that 1840 English obliged a reader to conclude – ceded once and ceded for good, both the power to make law and the power to enforce it over people and territory. Māori had no word that obliged that conclusion. The closest was *kawanatanga*, and *kawanatanga* obliges only that a governor has arrived.

Overtaken if: A record from February 1840 shows that someone explained the English text's full sovereignty claim to the rangatira in Māori before they signed, or the Waitangi Tribunal issues a later finding reversing its conclusion of 14 November 2014. If the first

appears, this chapter changes from a case of a gap nobody closed into a case of a gap someone spoke aloud but nobody wrote into the text.

Sources:

- New Zealand Ministry for Culture and Heritage, "Differences between the texts" — <https://nzhistory.govt.nz/politics/treaty/read-the-Treaty/differences-between-the-texts>
- National Library of New Zealand, He Tohu, "Two languages, different meanings" — <https://natlib.govt.nz/he-tohu/korero/two-languages-different-meanings>
- "Kāwanatanga" — <https://en.wikipedia.org/wiki/K%C4%81wanatanga>
- Waitangi Tribunal, "Report on stage 1 of the Te Paparahi o Te Raki inquiry released", 14 November 2014 — <https://www.waitangitribunal.govt.nz/en/news/report-on-stage-1-of-the-te-paparahi-o-te-raki-inquiry-released>

04 • Five Kinds That Do Not Translate

In 645 Xuanzang came back to Chang'an after sixteen years travelling overland to India, and what he left behind as a translator was not a theory of how to translate but a list of five kinds of word that must not be translated. The list is called 五种不翻 — "five kinds that do not translate" — five categories of term to be transliterated rather than rendered by meaning.

Roughly two and a half centuries before that, another man did the same job by the opposite method. Kumārajīva (344–413), of Kuchean and Indian parentage, arrived in Chang'an in 401 and led the translation bureau there. Over about eleven years his team produced a run of sutras, among them the Lotus Sutra, finished in 406 in seven fascicles, which became the standard East Asian version from then on, and the Diamond Sutra, translated between 402 and 412.

Kumārajīva put meaning ahead of wording. One exchange with his disciple Sengrui is on record. Sengrui had rendered a line literally as 天見人, 人見天, "heaven sees man, and man sees heaven." Kumārajīva would not have it, and had it changed to 人天交接, 兩得相見, "man and heaven connect, and the two are able to see each other." Elsewhere he compared rendering Sanskrit incantations literally to 嚼飯與人, chewing up rice and feeding it to someone else: the food is pre-digested before it reaches the mouth, so none of the taste survives.

Xuanzang (602–664) worked the other way. He set out overland for India, stayed there sixteen years, returned to Chang'an in 645, and founded the Yogācāra school with his disciple Kuiji — the same school China calls Faxiang, Weishi and Ci'en — and the Daci'en Temple, established in 648, became its ancestral seat.

The difference between the two methods can be measured on a single line of the same sutra. In the Heart Sutra, Kumārajīva's version (Taishō T223) reads 非色異空, 非空異色, a doubled negation: "it is not that form

differs from emptiness, it is not that emptiness differs from form." Xuanzang's later version (T251) reads 色不異空, 空不異色 — the same meaning, differently built, saying directly that form is not different from emptiness instead of going around through the double negative. The chiastic core line, 色即是空, 空即是色, is identical in both versions and does not move at all.

Beyond that line, what separates the two men is terminological consistency. Xuanzang set a standing Chinese equivalent for each technical Sanskrit term and pinned it there, while Kumārajīva used a more fluid idiom and followed the meaning. The wording the source uses is that Xuanzang is "generally described as" combining literal with free method rather than being purely literal, and this book does not tighten that phrase beyond what the source wrote.

The list of five is not advice. It is a sorting rule, decidable one word at a time: this word falls under one of the categories, or it falls under none of them. The five, as recorded:

1. Secrets — dhāraṇī (陀羅尼), Sanskrit ritual speech and mantras.
2. One word, many meanings — Bhagavant (薄伽梵) carries six meanings at once: sovereignty, ablaze, solemnity, fame, auspiciousness and esteem. No single Chinese word covers all six.
3. Nothing of the kind in China — the jambu tree (閻浮樹) does not grow on Chinese soil, so Chinese has no name of its own for it.
4. Follow what is already there — terms earlier translators had transliterated until the transliteration stuck, such as Anouputi (阿耨菩提) for anuttara-samyak-saṃbodhi.
5. Preserve the weight of the word — keep Prajñā (般若) rather than translating it as 智慧, the Chinese word for wisdom, which sits lighter.

Only one of the five is about a missing word. Category three alone concerns a thing China does not have; the other four are about something else entirely. Category one says the word works by sound and

not by sense — translate it and the sound is gone. Category two says the source language lets one word hold six meanings at the same time, and the target language forces the choice of one. Category four says the term already has a history in the target language — render it afresh and the reader can no longer trace it back. Category five says the register of a word is part of what that word says, and lowering the register says something else.

A rule for what not to translate is stronger than any theory of translation, for two reasons. First, it can be wrong in a way anyone can see: take a word in category two, render it in one Chinese word carrying all six meanings, and category two falls on the spot. A theory holding that translation is a negotiation between source and target can never be knocked over that way, because nothing sticks out to knock. Second, it is operational. A translator stuck on a single word reads this list and knows what to do with that word on the next line.

What this book's dossier does not have is the text that records the five categories, or that text's date. The dossier gives all five categories and an example for each, but does not say who wrote them down, where, or when. What can be held is Xuanzang's own lifespan, 602 to 664, so the rule belongs to the seventh century as far as the evidence in hand certifies, and no further.

This book's contract follows Xuanzang's fourth category exactly, without having set out to. The requirement is that technical terms and proper names stay in their original form in both the Thai and the English — evidentiality, parthenos, kawanatanga, BLEU, and 五种不翻 itself. The reason is not laziness, and it is not taste. These terms already have a history where they came from. A reader who meets evidentiality on this page can type it out and walk from here to the source work; a reader handed a freshly coined Thai word instead cannot walk anywhere, holding a word nobody outside this page uses.

The seventh century answered one of this book's questions already. The answer was not that everything translates, and not that nothing does. It was to sit down and write out what cannot cross, then lift those things

out and set them aside in their original form. Whoever does that takes one measured loss in front of him in order to keep it from spreading through the whole book.

What could not cross: The sound of a dhāraṇī, and the six meanings Bhagavānt holds at once. Sanskrit lets a syllable work as sound without having to mean anything, and lets one word hold several meanings simultaneously. Chinese obliges the translator to pick one meaning, and obliges him to write it out in characters that always mean something.

Overtaken if: Someone renders Bhagavānt in a single Chinese word that a Chinese reader takes in all six meanings without a footnote – category two falls. And if someone points to the text that records 五種不翻 and that text is dated outside Xuanzang's lifetime, this entry's attachment of the rule to the seventh century falls with it.

Sources:

- Wikipedia, "Chinese translation theory" – https://en.wikipedia.org/wiki/Chinese_translation_theory
- Wikipedia, "Kumārajīva" – <https://en.wikipedia.org/wiki/Kum%C4%81raj%C4%Bva>
- Wikipedia, "Lotus Sutra" – https://en.wikipedia.org/wiki/Lotus_Sutra
- Xue, "Study on Xuanzang's Translation of Heart Sutra" – <https://www.semanticscholar.org/paper/Study-on-Xuanzang's-Translation-of-Heart-Sutra-%E2%80%94of-Xue/04b07be9bdd8bcfe889969b5dae027ba8404c9b5>

05 • The Snow That Never Was

In 1986 Laura Martin traced the number everyone was citing — how many words for snow does Eskimo have — back to its origin, and found that the origin had mentioned four roots. Not fifty. Not two hundred. And the origin made no claim whatever about how its speakers thought.

The origin is Franz Boas, 1911, in the introduction to *The Handbook of North American Indians*. He mentioned four roots: *aput*, snow on the ground; *qana*, falling snow; *piqsirpoq*, drifting snow; and *qimuqsuq*, a snow drift. His point was purely technical. One language builds these meanings from separate roots; another builds them from a single root with modifiers attached. He did not write that the people who spoke that language saw snow differently from anyone else.

Whorf enlarged the number in 1940. The essay "Science and Linguistics," in *Technology Review*, raised the implied count to at least seven, with no footnote and no source, and asserted in addition that English covers all those distinctions with one word, *snow*. Pullum later established that the second assertion is false. The version later generations have cited sits in the collected volume Carroll edited in 1956, page 216.

What Martin documented was not a liar. It was a number swelling one relay at a time. Roger Brown in 1958 wrote three, misreading Whorf's own figure. Carol Eastman in 1975 wrote "many words" six lines from a passage she herself quoted that said three. Every relay cited the relay before it. No relay went back and counted.

Geoffrey Pullum collected the specimens that had escaped into the wild in a 1989 paper, "The Great Eskimo Vocabulary Hoax," published in *Natural Language & Linguistic Theory*:

- Nine — a trivia encyclopedia, 1984
- About four dozen — *New York Times*, 9 February 1988
- Fifty — Lanford Wilson's play *The Fifth of July*, 1978
- One hundred — *New York Times* editorial, 9 February 1984

- Two hundred — a television weather forecast in Cleveland

The two figures from the same newspaper do not agree, and this book prints both. One hundred in 1984 and about four dozen in 1988, four years apart, matching in masthead and in calendar date. Between those two days nobody went and counted. A number that was never measured can move in any direction.

The closest thing to a measurement in this affair is one Pullum took from C. W. Schultz-Lorentzen's 1927 Dictionary of the West Greenlandic Eskimo Language. It holds two roots bearing directly on snow: qanik, snow in the air or a snowflake, and aput, snow on the ground. Two. Not seven, not two hundred.

The fault heavier than the number is that no experiment was ever underneath this claim, not once. Nobody measured whether speakers of one language distinguish snow more finely than speakers of another. Nobody set a criterion for what counts as one word. What the claim had was a citation chain that lengthened every time it was retold. Both Martin and Pullum offer it as a case study in the field's own standards of evidence, not as a refutation of any relation between language and thought.

The line most often attributed to Sapir is not Sapir's position. In the 1929 paper "The Status of Linguistics as a Science," published in *Language* volume 5, pages 207–214, he wrote at page 209 that the real world "is to a large extent unconsciously built up on the language habits of the group." Elsewhere in his work he argued against the proposition that language determines thought, explicitly and more than once. The weak-versus-strong pairing that travels with his name is not his phrase and not Whorf's. A later generation fitted that frame back onto them.

Whorf's own language is genuinely stronger. In the 1940 essay, in the version Carroll collected in 1956, page 213, he wrote that the world reaches us in a "kaleidoscopic flux" that thought must organise, "and this means largely by the linguistic systems in our minds." Whorf was not a trained professional linguist. He was a fire-prevention engineer at the Hartford Fire Insurance Company who studied linguistics as an avocation

under Sapir at Yale. That biography settles nothing on its own. What settles it is the evidence base of his flagship claim.

The flagship claim was that Hopi grammar makes no reference to time in the way English does, and that Hopi speakers therefore hold no comparable concept of time. Its evidence base is thin. The source states that it reportedly rests on interviews with a single Hopi speaker, Ernest Naquayouma, in New York, not in the territory where the language is spoken. The source says reportedly, not confirmed, and this book keeps the word.

In 1983 Ekkehart Malotki published *Hopi Time: A Linguistic Analysis of the Temporal Concepts in the Hopi Language* with Mouton, running to some six hundred pages. The book documents that Hopi has tense, has words for units of time such as days, weeks and months, and has spatial metaphors for time. Malotki stated his own aim: he wrote it to refute Whorf. That intent is a fact about the author, not an argument against the document itself. What does the work is the grammatical description, and the linguistics that followed treats it as having succeeded.

This is not a replication coming back empty. It is a description of a language wrong at the level of fact. An empirical claim that was a load-bearing pillar of the theory fell because someone went into the field specifically to check it. Terry Regier and Yang Xu, in a 2017 review in *WIREs Cognitive Science*, wrote that the hypothesis occupies a "troubled and contested" position and carries an "inconsistent record of replication." Both write from inside the research tradition that supports it.

The bounds of this verdict have to be stated plainly. The snow story does not refute the claim that language affects thought. It refutes the most-repeated piece of evidence ever used to prove that claim. The two are not the same thing. Effects that survive scrutiny do exist, and they appear in the next chapter of this book, but they are small, they depend on the task the subject is given, and they are measurable only under narrow conditions. Under the snow there is nothing at all.

The rule this chapter lays down for the whole book has one clause. A claim that flatters this book's thesis gets checked harder than a claim that cuts against it, not more softly. The snow story is the best opening this book could have had — short, memorable, favourable to the thesis in every particular. It has no evidentiary content whatever, and the book prints that instead.

What could not cross: What failed to cross in this chapter is not a word between two languages but the condition that made a number checkable. Boas distinguished a root from the forms built on it. That distinction is the whole of what gave his "four" any meaning, and it was the first thing to go when the claim was retold in the English ordinary writers write. Once the counting rule stops travelling with the number, the number can be anything.

Overtured if: A document published before 1940 gives a snow-word count above four for some named Eskimo language, states its counting rule — roots, or forms built on roots — and the same rule applied to English still leaves a real gap. Or field work shows that the tense and the time-unit vocabulary Malotki recorded in 1983 do not match the Hopi people actually speak.

Sources: Laura Martin, "Eskimo Words for Snow': A Case Study in the Genesis and Decay of an Anthropological Example," *American Anthropologist* 88(2):418–423, 1986 — <https://anthrosource.onlinelibrary.wiley.com/doi/abs/10.1525/aa.1986.88.2.02a00080> · Geoffrey Pullum, "The Great Eskimo Vocabulary Hoax," *Natural Language & Linguistic Theory* 7:275–281, 1989 — <https://pullum.ppls.ed.ac.uk/EskimoHoax.pdf> · Terry Regier & Yang Xu, *WIREs Cognitive Science* 8:e1440, 2017 — <https://terry-regier.github.io/lclab/papers/whorf-wires-preprint-v2.pdf> · Stanford Encyclopedia of Philosophy, "Whorfianism" — <https://plato.stanford.edu/entries/linguistics/whorfianism.html> · Ekkehart Malotki, *Hopi Time: A Linguistic Analysis of the Temporal Concepts in the Hopi Language*, Mouton, 1983 — <https://archive.org/details/hopitimelinguist0000malo> · "Hopi time controversy," Wikipedia, secondary source — <https://en.wikipedia.org/>

wiki/Hopi_time_controversy · "Ekkehart Malotki," Wikipedia, background — https://en.wikipedia.org/wiki/Ekkehart_Malotki · Boas 1911, Whorf 1940 and the 1956 edition, Brown 1958, Eastman 1975 and Schultz-Lorentzen 1927 are cited through Martin 1986 and Pullum 1989; Sapir 1929 page 209 and Whorf 1956 page 213 are cited through Regier & Xu 2017. This book did not read those primary documents directly.

06 • The Two Blues of Russian

In 2007, at MIT, twenty-one Russian speakers judged pairs of blue that straddled the line between goluboy and siniy faster than pairs sitting on the same side of it: 1,164 milliseconds against 1,288.

A difference of 124 milliseconds. That is the true size of the effect this book will accept, and it is far smaller than what has been claimed of it. So this chapter prints what replicated and what did not on the same pages. Print them apart and the reader gets the number without its boundary.

In 1969 the colour question had already been answered once, and answered the other way. Brent Berlin and Paul Kay published *Basic Color Terms: Their Universality and Evolution*, surveyed 98 languages, and proposed that every language draws its basic colour terms from a universal set of eleven categories, admitting them to the lexicon in a broadly predictable order. The proposal was aimed straight at the strongest reading of Whorf: the belief that each language carves the spectrum as it pleases.

What has survived to today is not the eleven-stage sequence exactly. The World Color Survey that followed — 110 languages, and the dataset cited by both Winawer et al. 2007 and Kay & Regier 2007 — pushed the fixed order into a statistical tendency. The surviving conclusion is weaker, but it is not empty: some force presses wholly unrelated languages to name colours in the same direction.

And the two sides do not really contradict each other. In their own 2007 reanalysis, Kay and Regier hold two things at once — colour boundaries are not set arbitrarily, and colour boundaries correlate with perceptual differences between speakers of different languages. The field's principal researchers call the "universalist versus relativist" frame a false dichotomy.

Russian has no neutral word for blue. A speaker must choose between goluboy for light blue and siniy for dark blue, every time, not only when

precision is wanted. English carries no such obligation. Winawer, Witthoft, Frank, Wu, Wade and Boroditsky took that difference into the laboratory and published it in PNAS 104(19):7780–7785 in 2007.

The participants were 26 native Russian speakers and 24 native English speakers, recruited in Boston and tested at MIT; after exclusion for missing more than 25% of trials, 21 remained on each side (3 English, 5 Russian excluded). The task was to look at colour squares and say, as fast as possible, which of the two options matched the target. It ran in three forms – no interference, a digit string held in mind, and a spatial grid to remember.

The result is much narrower than the retelling. Russian speakers were indeed faster when a pair straddled the goluboy/siniy line, but only for pairs close enough to be hard to tell apart, and only when there was no digit string to hold. The category $\times$ interference interaction came out at $F(2,40) = 5.3$, $p = 0.009$. English speakers showed no category advantage under any condition: $F(1,20) = 0.150$, $p = 0.703$.

The sharpest number is how much of that advantage the digit string ate. On the Russian side, 173 milliseconds; on the English side, 14 [t(40) = 2.142, $p = 0.038$]; and in a parallel comparison, 189 against 15 milliseconds [t(40) = 2.17, $p = 0.036$]. Within the Russian group alone, the advantage on near colours was significant with no interference (1,164 against 1,288 milliseconds; $t(20) = 2.59$, $p = 0.0176$) and under spatial interference (1,162 against 1,270 milliseconds; $t(20) = 2.18$, $p = 0.041$). With a digit string to hold, it was not significant.

The authors themselves concluded something narrower than their readers passed on. The effect is small, it depends on the task – only hard discriminations, only under speeded conditions that block thinking in words – and it is generated on the spot. Language is recruited during the judgment and can be switched off by occupying the verbal channel. This is not a visual system permanently retuned.

Supporting work exists, but it does not stand as far off as it should. Winawer et al. cite three studies using the same triad paradigm –

Witthoft et al. 2003 (conference proceedings), Drivonikou et al. 2007 (PNAS 104:1097–1102), and Gilbert, Regier, Kay and Ivry 2006 (PNAS 103:489–494), whose own title reports the effect in the right visual field and not the left. The two laboratory lines, Kay/Regier/Ivry and Boroditsky/Winawer, overlap. The corroboration is real, but it is not corroboration from outside.

Here the other half has to be printed. Several other studies in this same line did not replicate – and the ones that did not replicate are the ones retold most often.

Grammatical gender does not change how people describe objects. That is the verdict after two replication attempts. The original claim comes from Boroditsky, Schmidt and Phillips 2003, "Sex, Syntax, and Semantics," in Gentner and Goldin-Meadow's *Language in Mind* (MIT Press). In German, key is masculine, and German speakers described keys in English as hard, heavy, jagged. In Spanish the same word is feminine, and Spanish speakers described them as beautiful, elegant, tiny. Its companion study was Phillips and Boroditsky 2003, on similarity judgments between pictures of people and pictures of gendered objects. At the time they were replicated, neither had passed peer review in a journal: one sat in a book, the other in conference proceedings.

The first time anyone tried to replicate it, nothing came out. Mickan, Schiefke and Stefanowitsch 2014, "Key is a llave is a Schlüssel: A failure to replicate an experiment from Boroditsky et al. 2003," in the *Yearbook of the German Cognitive Linguistics Association*, ran two separate attempts – a word association task and a primed lexical decision task – with German and Spanish speakers. Neither recovered the original effect. The authors were left with two options: the original was an artefact of undocumented procedural detail, or it was a statistical fluke.

The second attempt was preregistered and better powered, and still nothing came out. Elpers, Jensen and Holmes 2022, "Does grammatical gender affect object concepts? Registered replication of Phillips and Boroditsky (2003)," in the *Journal of Memory and Language* 127:104357, collected a total $N = 375$. Once the appropriate sources of error variance

were modelled, Spanish and German speakers' similarity judgments showed no gender-congruency effect at all. The original pattern surfaced only in a separate control condition, where English speakers were explicitly trained on artificial categories built to mimic gender. That is an effect of training, not of a native grammar.

Mandarin's vertical time fell the same way. Boroditsky 2001, "Does Language Shape Thought? Mandarin and English Speakers' Conceptions of Time," in *Cognitive Psychology* 43(1):1-22, ran three experiments – the first with 26 English-speaking and 20 Mandarin-speaking students, the second with 25 bilingual students, the third with 70 native English speakers – and reported that Mandarin-English bilinguals think about time vertically more than English monolinguals do. The reason offered was that Mandarin has additional vertical spatial metaphors for time.

Chen 2007, in *Cognition*, titled the paper "Do Chinese and English speakers think about time differently? Failure of replicating Boroditsky (2001)," and found no reaction-time difference of the predicted kind. Worse for the original was the corpus analysis in the same paper: Chinese speakers use horizontal time metaphors more often than vertical ones, which pulls the floor out from under the explanatory mechanism. Several further replication attempts are named in this literature – January & Kako 2007, Tse & Altarriba 2008, Chen & O'Seaghdha 2013 – but this book's dossier did not retrieve and check the texts itself, so they are printed as the citing literature reports them, not as independently confirmed.

And colour-memory effects in general, weighed across the whole set, are not reliable. Wright, Davies and Franklin 2015, "Whorfian effects on colour memory are not reliable," in the *Quarterly Journal of Experimental Psychology* 68:745-758 (DOI 10.1080/17470218.2014.966123), gathered cross-linguistic and developmental data and used Bayesian model comparison to weigh "Whorfian effect present" against "Whorfian effect absent" directly. The evidence consistently favoured absent. This is the paper that makes the present chapter most uncomfortable, because it

does not knock down the opposing camp's work. It knocks down the line siniy and goluboy stand on.

The tidiest explanation for statistics scattered like this comes from Regier and Xu 2017, in WIREs Cognitive Science 8:e1440. What decides whether an effect appears is not language but uncertainty in the non-linguistic representation. Tasks answerable with confidence — easy, simultaneous, no memory load — show almost nothing; uncertain tasks — delayed, difficult — give larger effects. The two of them offer this as an explanation of a mixed replication record, not as a new experimental finding, and this book prints it that way.

So the size of the real effect can be stated only this far: small, task-dependent, and visible only under time pressure, or when something else has taken the verbal channel. Outside those conditions it disappears; and weighed across the whole set the way Wright, Davies and Franklin weighed it, it does not stand. Anyone claiming that speakers of different languages live in different worlds will find no number in this chapter to borrow. A hundred and twenty-four milliseconds is not another world.

But a translator works in exactly those conditions: one sentence at a time, under time, at the boundary where two options are nearly the same — precisely where this effect shows up, on every count. When a Russian sentence obliges a choice between goluboy and siniy and the English sentence can take only blue, what goes missing is not the shade. What goes missing is the necessity of choosing.

There is no aggregate number for this field to print. This book's dossier searched for a meta-analysis quantifying an average effect size for linguistic relativity as a whole and did not find one. No such number appears on this page, and none should appear anywhere without a source.

What could not cross: Russian obliges its speaker to place every shade of blue on one side or the other of the goluboy/siniy line before the sentence can be finished. English does not oblige it, and cannot receive it

– "light blue" restores the shade but not the necessity of choosing, and what is not necessary a translator may always drop.

Overtuned if: a laboratory unconnected to both the Boroditsky/Winawer and the Kay/Regier lines preregisters and repeats the Winawer et al. 2007 triad task at higher power. Recover the category × interference interaction and this page grows firmer. Fail to recover it and this page joins grammatical gender and vertical time on the same pile. Either result overturns this chapter.

Sources: Winawer, Witthoft, Frank, Wu, Wade & Boroditsky 2007, PNAS 104(19):7780–7785 – <http://lera.ucsd.edu/papers/pnas-2007.pdf> ·

Berlin & Kay 1969, Basic Color Terms – <https://press.uchicago.edu/ucp/books/book/distributed/B/bo3643648.html> ·

Kay & Regier 2007, Cognition 102:289–298 – <https://lclab.berkeley.edu/papers/berinmo-cognition-in-press.pdf> ·

Mickan, Schiefke & Stefanowitsch 2014 – <https://www.degruyterbrill.com/document/doi/10.1515/gcla-2014-0004/html?lang=en> ·

Elpers, Jensen & Holmes 2022, Journal of Memory and Language 127:104357 – https://www.reed.edu/psychology/thought-lab/assets/publications/Elpers,%20Jensen,%20_%20Holmes%202022%20-%20JML.pdf ·

Phillips & Boroditsky, "Effects of Grammatical Gender on Object Description" – <https://escholarship.org/uc/item/3j37f380> ·

Boroditsky 2001, Cognitive Psychology 43(1):1–22 – <http://lera.ucsd.edu/papers/mandarin.pdf> ·

Chen 2007, Cognition – <https://pubmed.ncbi.nlm.nih.gov/17070793/> ·

Wright, Davies & Franklin 2015, QJEP 68:745–758 – <https://www.tandfonline.com/doi/abs/10.1080/17470218.2014.966123> ·

Regier & Xu 2017, WIREs Cognitive Science 8:e1440 — <https://terry-regier.github.io/lclab/papers/whorf-wires-preprint-v2.pdf>

07 • How Do You Know

In 2004, Alexandra Aikhenvald published *Evidentiality* with Oxford University Press, setting out what the grammar of many languages obliges in every single sentence. A speaker cannot simply say what happened. They must also say how they came to know it — saw it, were told it, worked it out from traces. English can say all three. English never obliges any of them. The distance between can say and must say is the whole subject of this book, and this chapter is where that distance is easiest to see.

The name for it is evidentiality — grammar that makes the source of a piece of knowledge something that has to be specified, not something that may be added. Aikhenvald sorts these systems into two kinds. The first marks only that evidence exists, without saying what kind. The second names the kind outright: seen with the eyes, heard from someone, worked out by inference. What separates the two is not finer vocabulary. It is the number of slots a speaker has to fill before the sentence can end.

Turkish does this with two past-tense morphemes. "-DI" is used for a past the speaker witnessed; "-mIş" for a past that was reported or inferred. Reaching for the wrong one is not a lapse of manners. It is a false statement about where the knowledge came from. Aksu-Koç and Slobin wrote a psychological account of this pair in 1986 and recorded that Turkish children acquire the distinction early.

Cuzco Quechua divides the ground into three slots: "-mi" for what is known directly, "-si" for what was heard from someone else, "-chá" for what is conjectured. Three is not the family's ceiling. Martina Faller's dissertation on Cuzco Quechua evidentials, and the SIL documentation alongside it, record a five-choice system in South Conchucos Quechua and a six-choice one in Sihuas Quechua.

Tariana, of the Arawak family, in northwest Amazonia, takes the obligation to its limit. Aikhenvald did the fieldwork on this language herself, and recorded in *A Grammar of Tariana*, with Cambridge University Press, that five evidentials must appear on essentially every finite verb, fused with tense so that the two cannot be pulled apart. Leave one out and the sentence is ungrammatical — not curt, wrong. A Tariana speaker therefore cannot say a bare sentence like "José played football." The answer to how do you know has to be tied into that same verb.

Take one claim and mark it three ways, and what the grammar carries for free becomes visible.

- **Witnessed** — the speaker was at the pitch. Turkish takes "-DI", Cuzco Quechua takes "-mi", Tariana selects the one of its five that fits seeing. The sentence is not a word longer.
- **Reported** — someone came and told them. Turkish takes "-mİş", Cuzco Quechua takes "-si", Tariana selects the one for a report passed on. The sentence is still not a word longer.
- **Inferred** — the speaker was not there, saw only the traces and concluded it. Turkish takes the same "-mİş", because the Turkish system has two slots and does not split heard-from-someone off from worked-out. Cuzco Quechua takes "-chá", the conjecture slot specifically. Tariana still has to select one of its own, and not selecting is not speaking. The sentence is still not a word longer.

English has to find extra words to carry the same load: I saw for the witnessed, they say or it is said for the reported, apparently, allegedly or it seems that for the inferred. These words carry the meaning in full. Not one of them is a mistranslation. What differs is their standing. In Turkish, Quechua and Tariana those markers are the grammar. In English they are words that may be put in or left out.

This is where the whole book stands. When an obligation crosses into a language that does not oblige it, it does not vanish on arrival. It becomes optional — and optional is what can be dropped. A careful translator will put apparently in every time. A hurried one will put it in sometimes. A machine that learns from human translations learns exactly that: in or out, either passes. Not one step in that chain is wrong, and not one step

can stop it, because the only thing that ever kept the information in every sentence was that it was compulsory.

What is lost, then, is not meaning but a guarantee. A reader of Tariana knows in advance that every sentence passing under their eye states its source, because a sentence that does not state it could never have been written in the first place. A reader of the English translation can draw no such conclusion. A sentence silent about its source may mean the speaker saw it, or may mean the marker fell off somewhere along the way. The page does not say which.

The next question is whether the obligation changes memory, and the answer has not settled. In 2013, Tosun, Vaid and Geraci published in the *Journal of Memory and Language*, volume 69, issue 2, pages 121–134, reporting that Turkish speakers — both monolinguals and Turkish-English bilinguals — remembered statements marked as firsthand significantly better than statements marked as non-firsthand, both for recognising them and for getting their source right, and were more likely to misremember non-firsthand information as firsthand. The pattern was reported as weaker, or absent, in English speakers.

In 2016, Ünal, Pinto, Bungler and Papafragou published in the *Journal of Memory and Language*, volume 87, pages 157–176, and designed the language out of the task entirely. Participants viewed photographs, some depicting an event directly, some only cueing it for inference, and source memory was measured without any spoken report. The result: speakers of both languages misremembered inferred events as seen events at much the same rate. The mechanism may not have been linguistic at all. The two studies answer in opposite directions, differing in the response format asked of participants, and no third study has arrived to settle it. This book prints both and says the question is open.

The scope of this chapter has to be stated plainly. Nobody disputes the grammatical facts: Turkish, Quechua and Tariana do oblige the source, and English does not. The cognitive consequence is unresolved. Nor was any memory experiment found on Quechua or Tariana speakers comparable to the Turkish work — not one. There is a further study that

lands exactly on this crossing point, titled "Lost in translation, apparently: Bilingual language processing of evidentiality in a Turkish–English Translation and judgment task," in *Bilingualism: Language and Cognition*; only its title and journal could be confirmed. The numbers inside it have not been read. This chapter therefore prints no figure from it.

The verdict of this chapter needs no memory experiment at all. It rests on the grammar alone. One language obliges every sentence to answer how do you know. Another does not. After the crossing the answer can still be said – but nothing obliges anyone to say it any more.

What could not cross: The compulsory standing of the evidential. Turkish, Quechua and Tariana oblige every sentence to say how the speaker came by the claim. English can say the same thing in full – I saw, they say, apparently – but no piece of its grammar obliges it. Nothing was mistranslated. The information simply stopped being compulsory, and compulsory was the only reason it was ever in every sentence.

Overtaken if: Someone measures a Turkish–English or Quechua–English parallel corpus and finds that source-language evidentials come through as English wording in nearly every sentence. If the drop rate really is low, the verdict of this chapter narrows to an observation about grammar, not an observation about what goes missing in the crossing.

Sources:

- Alexandra Aikhenvald, *Evidentiality*, Oxford University Press, 2004 – the standard cross-linguistic typology for this; the record cited is Aikhenvald's chapter on the typology of evidentiality in *The Grammar of Knowledge* – <https://academic.oup.com/book/11481/chapter/160189133>
- Alexandra Aikhenvald, *A Grammar of Tariana*, from Northwest Amazonia, Cambridge University Press – <https://www.cambridge.org/core/books/grammar-of-tariana-from-northwest-amazonia/FC9135A672C8EB1F643905ABD87C456C>
- Aksu-Koç and Slobin, "A psychological account of the development and use of evidentials in Turkish," 1986 – <https://wals.info/refdb/record/Aksu-Koc-and-Slobin-1986>

- Martina Faller, dissertation on Cuzco Quechua evidentials, full text — <https://personalpages.manchester.ac.uk/staff/martina.t.faller/documents/thesis-a4.pdf>
- Tosun, Vaid and Geraci, "Does obligatory linguistic marking of source of evidence affect source memory? A Turkish/English investigation," *Journal of Memory and Language* 69(2):121-134, 2013 — <https://www.sciencedirect.com/science/article/abs/pii/S0749596X13000235>
- Ünal, Pinto, Bungler and Papafragou, "Monitoring sources of event memories: A cross-linguistic investigation," *Journal of Memory and Language* 87:157-176, 2016 — <https://www.sciencedirect.com/science/article/abs/pii/S0749596X15001266>
- "Lost in translation, apparently: Bilingual language processing of evidentiality in a Turkish-English Translation and judgment task," *Bilingualism: Language and Cognition* — title, venue and topic confirmed only; sample size and effect size have not been read, so this chapter prints no figure from it — <https://www.cambridge.org/core/journals/bilingualism-language-and-cognition/article/abs/lost-in-translation-apparently-bilingual-language-processing-of-evidentiality-in-a-turkishenglish-translation-and-judgment-task/94BBB0151AB2214E72DFFE1BB71B9AF6>

08 • The Dust Beneath the Royal Feet

In 1997 the Royal Institute academician Kanchana Naksakul wrote in the Royal Institute Newsletter that carrying Thai's **ฉัน** and **คุณ** — "I" and "you" — into English takes almost no thought: I and you, and that is the end of it. Going the other way, the translator has to establish first who is speaking, how old they are, their gender, who they are speaking to, how close the two are, whether the setting is formal, whether the speaker is being humble or is angry, and what social class they come from.

That list is what English never has to answer. Thai has to, and answers it from the first word. The most formal part of the system has a name of its own — **ราชาศัพท์**, the royal register, the vocabulary reserved for the king, the nobility and the monks — but the same machinery runs down as far as a sentence a father says to his child. The rank of the two people talking is not floating somewhere in the context. It is embedded in the verb and in the pronoun.

The verb meaning "eat" has four forms, and taking the wrong one is not a lapse of style. **เสวย** is used of the king and of high royalty. The verb **ฉัน** is used of Buddhist monks — a different axis from rank, since it marks a religious role and not a class. **รับประทาน** is the ordinary polite register. **กิน** is the bare word. Use the wrong rung and it lands as disrespect, or as formality absurd enough to be funny.

รับประทาน is itself the fossil of a sentence. The Office of the Royal Society explained on its radio programme "**รู้ รัก ภาษาไทย**" on 30 October 2009 that as royal-register vocabulary it is two words — **รับ**, "to receive," and **ประทาน**, "to give" — used when a prince or princess of **พระองค์เจ้า** or **หม่อมเจ้า** rank gives something to a commoner. Run together as a single word, with no pause, the two syllables have eroded into a polite verb meaning simply "eat." The Office goes as far as instructing speakers to pause between the syllables when they mean the older sense.

English has one word, ate, and no rungs to choose between. A translator carrying **เสวย** across has to put in words that were not there — "the King ate" — so what Thai has already said inside the verb, English says again outside it. The information does not vanish in the crossing. It changes category: from something grammar obliges to something context can supply, and what is not obliged can always be cut.

The second person has four forms that cannot be exchanged, ordered by the rank of the person addressed. **ใต้เท้า**, "beneath your feet," for a respected superior. **ใต้ฝ่าพระบาท**, "beneath the royal sole," for a **หม่อมเจ้า**. **ใต้ฝ่าละอองพระบาท** adds the dust rising off that sole, and is used from **พระองค์เจ้า** rank upward. **ใต้ฝ่าละอองธุลีพระบาท** adds a mote within that dust, and is reserved for the king. The speaker is not choosing how polite to be. He is announcing the rank of the person he addresses loudly enough for the room to hear it. You cannot do that.

The first person climbs in the opposite direction. **ข้าพระพุทธเจ้า** means "servant of the Lord Buddha," a religious self-abasement carried over into address of the king. **เกล้ากระหม่อม** and **หม่อมฉัน** put up the crown of the speaker's own head in place of the speaker, and are used by lower nobility addressing higher. Kanchana's account is that Thai treats the head as the highest part of the body, so the speaker offers his head rather than the whole of himself. And **ผม**, the word used every day, means literally "hair" — the same logic, fallen to the ordinary register.

Kanchana gives one case where it breaks. In foreign films dubbed into Thai you sometimes hear a father call himself **ผม** and address his own child, under ten years old, as **คุณ**. Every word is correct by the dictionary, and it grates badly, because no Thai father speaks to his child that way. The error is not lexical but relational — that pair of pronouns carries a formal distance built into it, which a father and child do not have.

The sentence-final particles **ครับ**, spoken by men, and **ค่ะ**, spoken by women, close nearly every polite Thai utterance, and mark two things at once: the speaker's own gender and the formality of the sentence. English has no word class that does this work. Eric McCready set out a

semantics for it in "A Semantics for Honorifics with Reference to Thai" (PACLIC 28, 2014) — pronouns, verb forms and particles jointly fix a politeness level, and that level must then be compatible with a formality level the discourse context has already set.

In 2020 the President of the Office of the Royal Society, ศ.เกียรติคุณ นพ.สุรพล อิศรางกูร ศิล, said of a word his own institution had coined: "ตอนนี้มีใครใช้คำว่าคณิตกรณ์ ก็เรียกว่าคอมพิวเตอร์กันหมด" — does anyone use the word *khanitkorn* now? Everyone just says computer. A language that can oblige rank inside its verbs could not oblige its own vocabulary.

The instrument that should have been able to oblige it exists, and has a birthday. The body that became *ราชบัณฑิตยสภา* was founded in 1933; formal terminology-coining began in 1942, per its Secretary-General, *กนกวลี ชูชัยยะ*, in 2019. The method is stated in three tiers, per *สุรพล*'s interview of 6 July 2020 — find a native Thai word first; failing that, build one from Pali-Sanskrit roots; transliterate only as a last resort. And it has worked. *โทรทัศน์*, "television," is *โทร* ("far") plus *ทัศน์* ("to see").

The words that lost have names and dates too. The Society's science and computing committees coined *คณิตกรณ์* from Pali-Sanskrit roots, roughly "calculating instrument," alongside the transliteration *คอมพิวเตอร์*. *สุรพล* gave the outcome himself in 2020: "*ปรากฏว่ามีเพียงคำทับศัพท์คอมพิวเตอร์เพียงคำเดียวที่รอด*" — in the event, only the transliterated word "computer" survived. *จดหมายอิเล็กทรอนิกส์*, "electronic letter," coined for email, holds on in official paperwork and some teachers' documents; Thai mouths say *อีเมล* — this from *รศ.อรพัช บวรรักษา* (Assoc. Prof. Orapat Bavornraksa) in the same set of interviews.

The older mechanism belonged to one man rather than a committee. Rama VI (Vajiravudh, reigned 1910–1925, died 26 November 1925) translated three plays of Shakespeare's into Thai — *เวนิสวานิช*, *The Merchant of Venice*, in 1916; *ตามใจท่าน*, *As You Like It*, in 1921; *โรมิโอและจูเลียต*, *Romeo and Juliet*, in 1922. He set out his own method in the preface to the last of them, signed "*รามวชิราวุธ ประ*" and dated 1 June 1922: "*ในการประพันธ์คำกลอนภาษาไทยในเรื่องละครนี้ ข้าพเจ้าตั้งใจให้ใกล้เคียงสำนวนและ*

โวหารเดิมของเชกส์เปียร์ มากกว่าคำนึงถึงความไพเราะทางกวีภาษาไทย" — in composing the Thai verse for this play, I intended to keep close to Shakespeare's own idiom and style, rather than prioritise the mellifluousness of Thai poetic convention. That is a source-fidelity policy, declared 73 years before Lawrence Venuti coined the term foreignization in 1995.

His coining was not confined to the plays. The Surnames Act of พ.ศ. 2456, 1913, required every Siamese subject to take a hereditary family name for the first time, and many of the compound Pali-Sanskrit surnames that followed he devised or bestowed himself. Another was ลูกเสือ, "tiger cub," the name he gave the Boy Scouts of Siam in 1911, built on เสือป่า, the Wild Tiger Corps he founded the same year. This is a different mechanism from the terminology office, which was not born until 1933, after both reigns.

The word "first" here has already been contested. อาชญาสิทธิ์ ศรีสุวรรณ wrote in The101.world on 31 January 2023 that Rama VI's preface claimed priority only for translating the play as line-by-line stage dialogue, not for bringing the story into Thai at all. ขุนสุนทรภักษ์ (Thanom Keyanon), a schoolteacher, published "โรมยูเลียดคำฉันท" between August and October 1921, a full year ahead of the King's version. Whether the King knew of it is undocumented; the article's author states that he could not determine it.

On the Pali and Sanskrit share of the modern Thai lexicon, this book prints no number. The research for this chapter found no verifiable measurement. The figures in circulation — "30–40%," or "more than half" — could not be traced back to anyone who did the counting or to a method. No number does not mean there is nothing there to measure. It means nobody has measured it. What has actually been counted is narrower: Søren Egerod counted 181 Teochew loanwords in Siamese, in 1959.

The highest word in this chapter is itself borrowed. Rikker Dockum read roughly a hundred Sukhothai inscriptions, 13th to 15th century, and found Sanskrit arriving in at least two waves. The first came through Khmer,

carrying Khmer's own sound changes in with it — which is why Thai has บ and ด at all; the second was taken directly. The result is doublets still in use, บิดา beside บิดุ, and the word บาท, from pāda, "foot." The บาท in ใต้ฝ่าละอองธุลีพระบาท is therefore a loan that arrived by way of Khmer, and no committee ever voted it in.

That is the whole difference between the two halves of this chapter. Rank could take up residence inside the verb because nobody had to approve it. คณิศรกรณ had a resolution behind it, a stated method behind it, a state office behind it, and it lost. Grammar obliges. Policy can only ask.

The strongest argument against this verdict is that transliteration is a form of absorption too, and Thai has been absorbing from Khmer, Pali, Sanskrit, Chinese and Portuguese for centuries. Nor do dead coinages always die of English: พุทธิธรรม lost to วัฒนธรรม, for "culture," before the digital era and before the terminology office existed. So the claim made here is narrow. The Office stated its own order of preference plainly — Thai first, Pali-Sanskrit second, transliteration last — and its own President said in 2020 that last place was the place that survived.

What could not cross: The rank of speaker and hearer, which Thai binds into its verbs and its pronouns so tightly that it cannot go unsaid, cannot cross into English, because English has nowhere in its grammar to put it. The translator has to add words instead, and words that can be added can always be cut.

Overtaken if: Someone produces a lexical count — the counter named, the year given, the method stated — that actually measures the Pali-Sanskrit share of the modern Thai lexicon; or usage data showing that คณิศรกรณ or จดหมายอิเล็กทรอนิกส์ is still current in ordinary speech and not only in official paperwork.

Sources: กาญจนา นาคสกุล (Kanchana Naksakul), "คำสรรพนามในภาษาไทยสะท้อนวัฒนธรรมไทย", จดหมายข่าวราชบัณฑิตยสถาน vol. 7 no. 73 (June 1997)

— <http://legacy.orst.go.th/?>

knowledges=%E0%B8%84%E0%B8%B3%E0%B8%AA%E0%B8%A3%E0%B8%A3%E0

· สำนักงานราชบัณฑิตยสภา, "รับประทาน", radio programme "รู้ รัก ภาษาไทย", broadcast 30 October 2009 — <http://legacy.orst.go.th/?knowledges=%E0%B8%A3%E0%B8%B1%E0%B8%9A%E0%B8%9B%E0%B8%A3%E0%B9%93%E0%B9%90%E0%B8%95%E0%B8%B8%E0%B8%A5%E0%B8%B2%E0%B8%84%E0%B8%A1%E0%B9%92%E0%B9%95%E0%B9%95%E0%B9%92> · Eric McCready, "A Semantics for Honorifics with Reference to Thai", PACLIC 28 (2014) — <https://aclanthology.org/Y14-1058.pdf> · "'ราชบัณฑิต' ยุคใหม่ กับ 'ศัพท์บัญญัติ' ดิสรบั้ง", Bangkok Biznews, 6 July 2020 (interviews with สุรพล อิศโรกรศีล and อรพัช บวรรักษา) — <https://www.bangkokbiznews.com/social/888035> · "'ระบบศัพท์บัญญัติ' สำนักงานราชบัณฑิตยสภา ปรับปรุงล่าสุด พร้อมใช้งานแล้ววันนี้", NECTEC, 30 April 2019 (quoting กนกวลี ชูชัยยะ) — <https://www.nectec.or.th/news/news-pr-news/royin-coined.html> · ศัพท์บัญญัติ ๔๐ สาขาวิชา, Office of the Royal Society — <https://coined-word.orst.go.th/> · Kanlaya Coulsting, "Shakespeare in Thai", Shakespeare Birthplace Trust, 1 November 2017 — <https://www.shakespeare.org.uk/explore-shakespeare/blogs/shakespeare-thai/> · "To be or not to be in Thailand", Bangkok Post Special Report (2016) — <https://www.bangkokpost.com/thailand/special-reports/1241218/to-be-or-not-to-be-in-thailand> · อาชญาสิทธิ์ ศรีสุวรรณ, "มีผู้แปล Romeo and Juliet มาเป็นภาษาไทยก่อนในหลวงรัชกาลที่ 6", The101.world, 31 January 2023 — <https://www.the101.world/romeo-and-juliet-thai/> · "Vajiravudh", Wikipedia — <https://en.wikipedia.org/wiki/Vajiravudh> · "Thai name", Wikipedia — https://en.wikipedia.org/wiki/Thai_name · "กำเนิด 'เสือป่า' ก่อนมี 'ลูกเสือ'! ร.๖ ทรงตอบโต้คำวิจารณ์...", MGR Online — <https://mgronline.com/onlinection/detail/9630000067423> · Søren Egerod, "Swatow loan-words in Siamese", Acta Orientalia 23 (1959), 137-156 — <https://journals.uio.no/actaorientalia/article/view/5303> · Rikker Dockum, "Sanskrit loanword adaptation in Old Thai: epigraphic evidence from the Sukhothai corpus", 16th World Sanskrit Conference — https://www.academia.edu/28516211/Sanskrit_loanword_adaptation_in_Old_Thai_epigraphic_evidence_from_the_Sukhothai_corpus · "List of loanwords in Thai", Wikipedia — https://en.wikipedia.org/wiki/List_of_loanwords_in_Thai

09 • Human Parity

In March 2018 Microsoft Research announced that its machine had reached human parity translating Chinese news into English, and within the same year two teams working independently showed that the word parity was a product of the evaluation method, not a property of the translation.

The claim was not weak, and it deserves its full force before it is broken. Hany Hassan and colleagues published arXiv:1803.05567, "Achieving Human Parity on Automatic Chinese to English News Translation," reporting that their neural system's output on the WMT 2017 Chinese-to-English news test set was statistically indistinguishable from the work of professional human translators. The evidence came in two layers. The first was automatic metrics. The second was a formal protocol in which bilingual judges scored the output. Nothing offhand about it, and both layers said the same thing.

Samuel Läubli, Rico Sennrich and Martin Volk took the same translations back to human judges, changing one variable — the raters read whole documents instead of isolated sentences served in shuffled order. The paper, "Has Machine Translation Achieved Human Parity? A Case for Document-level Evaluation," appeared at EMNLP 2018, and its result reads: "human raters assessing adequacy and fluency show a stronger preference for human over machine translation when evaluating documents as compared to isolated sentences." Same system, same translations, human judges in both conditions. The only thing that changed was the length of what was put in front of them.

The errors that surface over a longer read have names of their own. Dropped pronouns. Coherence that breaks between sentences. One term called two different things in the same story. None of the three is visible in an isolated sentence, because an isolated sentence gives them nowhere to appear. A sentence cut out of a paragraph has nothing for a

pronoun to point back to, and no second occurrence for a term to contradict its first.

Antonio Toral, Sheila Castilho, Ke Hu and Andy Way checked the same claim by another route. Their paper, "Attaining the Unattainable? Reassessing Claims of Human Parity in Neural Machine Translation," at WMT 2018, found that the parity result depended on how the evaluation had been designed, and named two choices that flattered the machine directly – the source text used for testing was itself translated material rather than text originally authored in that language, and the raters were shown no inter-sentential context.

The sentence-sized window is sixteen years older than this episode. Kishore Papineni, Salim Roukos, Todd Ward and Wei-Jing Zhu proposed BLEU at ACL 2002 as a modified n-gram precision score against reference translations, with a brevity penalty to stop a system scoring by writing short, and its own authors called it an "understudy to skilled human judges." An understudy, not the lead. The word was in the founding paper from 2002, and the understudy then held the stage for roughly two decades.

BLEU cannot report what breaks across sentences, because it never looks at anything longer than a sentence. This is not a flaw in the formula. It is the size of the window. Chris Callison-Burch, Miles Osborne and Philipp Koehn wrote at EACL 2006 that a rise in BLEU is "neither necessary nor sufficient" for a real improvement in quality, with counterexamples in both directions: systems BLEU ranked higher and humans ranked lower, and systems where it ran the other way. Twelve years before the 2018 announcement, the warning was already in the field's own proceedings.

Human judges fail the same way if given the same window. Markus Freitag, George Foster, David Grangier, Viresh Ratnakar, Qijun Tan and Wolfgang Macherey published in TACL 9 in 2021 the largest MQM study yet conducted, having professional translators read with full document context and re-score the WMT2020 systems in two language pairs. The outcome was "a substantially different ranking of evaluated systems from the one established by the WMT crowd workers." The fault, then, is not

whether the judge is a machine or a person. It is what the judge was shown.

The line runs back to 1966. ALPAC was established in April 1964, chaired by John R. Pierce of Bell Telephone Laboratories with six other members. Its report, *Language and Machines — Computers in Translation and Linguistics*, was issued in November 1966. It defined machine translation very narrowly, as translation "by algorithm from machine-readable source text to useful target text, without recourse to human translation or editing," and concluded under that definition that "there has been no machine translation of general scientific text, and none is in immediate prospect." Funding was cut, and MT research in America stopped for roughly twenty years.

The two cases run the same mechanism with the sign reversed. In 1966 a narrow definition reduced an entire field's output to zero. In 2018 a narrow protocol raised one set of translations to human parity. Every generation's evaluation regime measures what it happens to be able to measure, and what it can measure then becomes the meaning of the word good. One detail from the ALPAC report itself is worth noting: one of the three things it recommended funding was translation evaluation methods.

The field has fixed this, and it fixed it by widening the window rather than by waiting for the models to improve. Maja Popović proposed chrF at WMT 2015, scoring on character n-grams instead. Ricardo Rei and colleagues proposed COMET at EMNLP 2020, learning directly from human judgments rather than counting matching n-grams. The WMT24 organizers themselves name the two top performing automatic metrics as MetricX-23-XL and CometKiwi-DA-XL — BLEU is no longer the yardstick by which the field's best systems are ranked. WMT25 went further, to 32 language pairs and 60 systems, adding a new Literary domain built from documents of roughly 5,000 words, for the reason printed in the report's own title — "Time to Stop Evaluating on Easy Test Sets."

The Thai-side number cannot be confirmed, and this book will not guess it. NLLB-200 reports a 44% average BLEU improvement over the prior state of the art across more than 40,000 translation directions on FLORES-200, which spans 204 language and script pairs, and Thai (tha_Thai) is among them. But a BLEU or chrF score specific to Thai could not be traced to a primary source in this research pass. As for the WMT25 system tables, all 32 language pairs of them, inspection turned up no English-Thai pair among them. The wider window has not yet reached the language this book is written in.

What the sentence-sized window cannot see is the set of obligations that arise only once a text runs longer than one sentence: the pronoun that has to point back to a person named three sentences ago, the term that has to be called the same thing every time until the end, and the register that says who is speaking to whom. The last of those has evidence of its own. Luis Cihlar and colleagues examined the automatic translation of Korean honorifics and reported at the IAAI workshop at LREC 2026 that models shift politeness level according to whether the addressee is stated explicitly or has to be inferred. The systems match a cue at the surface of the text; they do not model the relationship. A single sentence has no addressee. A whole text does.

What could not cross: The pronoun that has to point across a sentence boundary, the term that has to stay fixed for the length of a story, and the register that specifies who is speaking to whom — all three are obligations of a whole text, not of a single sentence. An instrument with a one-sentence window therefore reports that nothing is missing, because it never looks at the place the missing thing was.

Overtaken if: Take Hassan et al. 2018's translations back to human judges under both conditions, shuffled isolated sentences and full documents. If the raters do not prefer the human translation more strongly when reading full documents, or if Läubli, Sennrich and Volk 2018's result fails to replicate in other language pairs, this verdict fails.

Sources: Hassan et al., "Achieving Human Parity on Automatic Chinese to English News Translation," arXiv:1803.05567 (2018) — <https://arxiv.org/abs/1803.05567> · Läubli, Sennrich, Volk, "Has Machine Translation Achieved Human Parity? A Case for Document-level Evaluation," EMNLP 2018 — <https://aclanthology.org/D18-1512/> · Toral, Castilho, Hu, Way, "Attaining the Unattainable? Reassessing Claims of Human Parity in Neural Machine Translation," WMT 2018 — <https://aclanthology.org/W18-6312/> · Papineni, Roukos, Ward, Zhu, "BLEU: a Method for Automatic Evaluation of Machine Translation," ACL 2002 — <https://aclanthology.org/P02-1040/> · Callison-Burch, Osborne, Koehn, "Re-evaluating the Role of Bleu in Machine Translation Research," EACL 2006 — <https://aclanthology.org/E06-1032/> · Popović, "chrF: character n-gram F-score for automatic MT evaluation," WMT 2015 — <https://aclanthology.org/W15-3049/> · Rei et al., "COMET: A Neural Framework for MT Evaluation," EMNLP 2020 — <https://aclanthology.org/2020.emnlp-main.213/> · Freitag, Foster, Grangier, Ratnakar, Tan, Macherey, "Experts, Errors, and Context: A Large-Scale Study of Human Evaluation for Machine Translation," TACL 9, 2021 — <https://aclanthology.org/2021.tacl-1.87/> · Pierce, Carroll, et al., Language and Machines — Computers in Translation and Linguistics (ALPAC Report), National Academy of Sciences, 1966 — <http://www.mt-archive.info/ALPAC-1966.pdf> · Kocmi et al., "Findings of the WMT24 General Machine Translation Shared Task: The LLM Era Is Here but MT Is Not Solved Yet," WMT 2024 — <https://aclanthology.org/2024.wmt-1.1/> · Kocmi et al., "Findings of the WMT25 General Machine Translation Shared Task: Time to Stop Evaluating on Easy Test Sets," WMT 2025 — <https://aclanthology.org/2025.wmt-1.22/> · Kocmi et al., "Preliminary Ranking of WMT25 General Machine Translation Systems," arXiv:2508.14909 (2025) — <https://arxiv.org/abs/2508.14909> · NLLB Team et al., "No Language Left Behind: Scaling Human-Centered Machine Translation," arXiv:2207.04672 (2022) — <https://arxiv.org/abs/2207.04672> · Cihlar, Bui, Park, Mager, Bisang, von der Wense, "Investigating the Automatic Translation of Korean Honorifics," LREC 2026 IAAI Workshop — <https://aclanthology.org/2026.iaai-1.3/>

10 • The Vegetarian

On 2 July 2017, Charse Yun, a professor at Ewha Womans University, published a critique of the translation that had won the Man Booker International Prize the year before. His charge was not that it had been mistranslated. It was that it had been over-translated.

The book had a life before it crossed. Han Kang wrote *The Vegetarian* (Chaesikjuuija) in three parts between 2003 and 2005; it appeared as a novel in Korea in 2007, and its central section won the Yi Sang Literary Award in 2005. Deborah Smith began learning Korean in 2009, founded the press Tilted Axis in 2015, and translated this book as her first book-length translation. The English edition came out in January 2015 from Portobello Books. The following year the two of them shared the Man Booker International Prize, in the first year the prize credited author and translator equally.

Yun's criticism came in two layers. In the first he rated the translation a high achievement. In the second he argued that it had drifted from Han Kang's spare, restrained manner by adding intensifiers, superlatives and emphatic word choices — not once or twice, but on close to every other page. One senior translator he consulted took the view that Smith's work was better called an adaptation than a translation. The first piece ran in Korea Exposé on 2 July 2017 under the title "You Say Melon, I Say Lemon: Deborah Smith's Flawed Yet Remarkable Translation of 'The Vegetarian'"; he then extended the argument in the Los Angeles Times on 22 September 2017.

The two sets of evidence in this entry do not carry equal weight. Yun's wording reached this book through secondhand citation, not through the Los Angeles Times text itself, so his argument is reported here in summary rather than quoted. The words of Smith and Han Kang below were checked against their sources one word at a time.

Smith answered for herself in the Los Angeles Review of Books on 11 January 2018. Her position sits in a single sentence: "there is no such thing as a truly literal translation... there can be no such thing as a translation that is not 'creative.'" She separated the charges into two piles. The first was outright mistranslation, and for that she took responsibility. The second was stylistic divergence, which she defended as one legitimate reading of Han Kang's prose.

Han Kang addressed it herself in an interview the Booker Prize Foundation published on 28 July 2023: "I do not believe that the translator has deliberately undermined the original, nor do I believe that she has created a new work that is completely different from the original." She pointed out that the dispute had mixed two things together. The first was genuine mistranslation: Smith made 67 corrections in 2017, and they went into the 2018 UK and US reprints. The second was the argument about style — and the first was being used as evidence for the second.

The example she gave herself is the word *cheohyeong*, which means a wife's older sister, though its two syllables read literally as wife plus older brother — a trap for anyone who did not grow up in the language. Some critics read that slip as evidence that Smith had changed a character's sex, when it was purely linguistic. She also disclosed that around half a page of the Korean text is missing from the English edition. That was an editorial cut, not a failure of translation, because the translator had translated every sentence. The two went on working together; the manuscript of *The White Book* (2017) was compared sentence by sentence.

The dispute is not an argument about who did the better job. It is an argument about what a translation owes first. The same question had ended a friendship once already, fifty years earlier.

In 1964 Vladimir Nabokov published his translation of Pushkin's Eugene Onegin with the Bollingen Foundation: four volumes of unrhymed, near-literal iambic prose, plus more than 930 pages of commentary. He had declared the position long before, in "The Art of Translation" in *The New Republic* on 4 August 1941. That essay opens: "Three grades of evil can be

discerned in the queer world of verbal transmigration." The worst grade is prettifying a text to suit the taste of the audience it is being carried to.

Edmund Wilson reviewed the translation in *The New York Review of Books* on 15 July 1965, under the title "The Strange Case of Pushkin and Nabokov". The two had been friends since 1940, when Wilson was *The New Republic's* critic and helped bring Nabokov into American literary life. His verdict on the translation: "It has produced a bald and awkward language which has nothing in common with Pushkin or with the usual writing of Nabokov."

The heavier charge was about interpretation. Wilson read Onegin's killing of Lensky in the duel as malice consistent with the character, where Nabokov read it as out of character. Wilson wrote: "There are no out-of-character actions in Evgeni Onegin. Nabokov has simply not seen the point." He still closed the review with praise, calling Nabokov a live wire strung between the reader and the Russian past.

Nabokov replied in a letter in the issue of 26 August 1965. He opened by affirming the friendship, then itemised seven errors in Wilson's Russian grammar and usage, from confusing two senses of the colloquial *netu* to getting the sound of the letter "ë" wrong. He closed: "Its mixture of pompous aplomb and peevish ignorance is certainly not conducive to a sensible discussion of Pushkin's language and mine." Wilson answered in the same issue, conceding two of the seven: "I must acknowledge two errors... I felt that it sounded more damaging than I had meant it to be, and this has given him a chance to score." The rest he contested point by point.

The exchange ran on for months in the NYRB letters pages, with other Slavists joining, David Magarshack and Ernest J. Simmons among them, and Nabokov returning with a piece titled "Translation" on 20 January 1966. A twenty-five-year friendship did not come back. Alex Beam, in *The Feud* (2016), treats this as the rupture point.

Under the insults there was a real disagreement about method. Wilson wanted a verse translation that read fluently as English, to the point of

offering a competing stanza of his own in the review. Nabokov held that any deviation from literal sense for rhyme or euphony betrayed the poem. He had said so in 1941 and said it again in the prefatory poem "On Translating Eugene Onegin" (1955): a faithful crib is more honest than a beautiful lie.

The frame used for a conflict of this kind comes from Lawrence Venuti, in *The Translator's Invisibility: A History of Translation* (Routledge, 1995; revised edition 2008). A translation fluent and idiomatic enough in the target language that readers believe they are reading an original, he calls domestication. A translation that deliberately keeps the friction, the unfamiliar syntax and the traces of the source language, he calls foreignization. His proposal does not stop at description. The fluency Anglophone publishing has always rewarded is not a neutral taste; it is an ideology, and it forces the translator to disappear. In his view foreignization is therefore the more ethical choice, because the reader stays aware throughout of reading something that has crossed.

The frame drew hard pushback, chiefly from Anthony Pym, whose argument is reported here in summary because his exact wording was not checked in this research pass. It has three layers. First, foreignization has no workable method. Second, Venuti does not show that the democratic cultural exchange he claims actually follows from it. Third, the domestication and foreignization axis is too crude to describe what translators do; real work sits on a spectrum of partial adaptation. Maria Tymoczko's work on postcolonial translation ethics is cited as extending the argument past that binary. Venuti's own answer is that fluency is not the enemy, that translations should be readable, and that what he attacks is the squeezing of readable down to one register of standard English, whatever register the source was written in.

Put the two cases on this frame and the picture sharpens. Smith chose the domestication side: the book must work as literature in English first. Nabokov chose the foreignization side to its limit: the reader should see through the English back to the Russian sense, and the English may be deformed to allow it. Each side's loss was predictable from that choice

alone. One loses the register and restraint of the original; the other loses the poetry. Both sides had even been named in advance. Wilson called the second side's result "bald and awkward" in 1965; Nabokov had called the first side's method the highest grade of evil in 1941, before either book existed.

One part of the dispute can be settled, and it is far smaller than what the two sides argued over. The 67 mistranslations in *The Vegetarian* were wrong under either standard: whatever obligation a translator holds, *cheohyeong* still means a wife's older sister. The seven errors of Russian are the same kind of thing. Wilson conceded two of them himself; the rest are still argued. What cannot be settled is the much larger part, the question of what a translation owes first.

This book can support exactly one verdict. These two translators worked to different obligations, and each one's loss was predictable from the obligation chosen, before a first word was translated. The verdict this book cannot support is that either of them was wrong, because no measurement adjudicates between two obligations. There is only the check of whether a translator did what they said they were doing, and whether an error is an error under both standards. So this book says plainly that it cannot rule, rather than pretending that it has.

What could not cross: Korean kinship terms oblige a speaker to specify side and seniority in a single word: *cheohyeong* says both that the relative is on the wife's side and that she is the elder. English obliges neither, and cannot put them back without adding words. The other thing that could not cross is the verdict itself — no measure weighs an obligation to the reader on the far side against an obligation to the text on the near one.

Overtaken if: Someone builds a checkable criterion that ranks domestication against foreignization without first assuming what a translation is for; or if someone reads Charse Yun's text in the *Los Angeles Times* of 22 September 2017 directly and finds his charge differs from what is reported here.

Sources:

- Han Kang, interviewed by the Booker Prize Foundation, "I never imagined The Vegetarian would find so many readers," 28 July 2023 — <https://thebookerprizes.com/the-booker-library/features/han-kang-interview-i-never-imagined-the-vegetarian-would-find-so-many>
- Deborah Smith, "What We Talk About When We Talk About Translation," Los Angeles Review of Books, 11 January 2018 — <https://lareviewofbooks.org/article/what-we-talk-about-when-we-talk-about-translation/>
- "Deborah Smith (translator)," Wikipedia, citing Charse Yun via the Los Angeles Times, 22 September 2017 — [https://en.wikipedia.org/wiki/Deborah_Smith_\(translator\)](https://en.wikipedia.org/wiki/Deborah_Smith_(translator)) (Yun's wording reached this entry through secondhand citation only)
- "Yi Sang Literary Award," Wikipedia — https://en.wikipedia.org/wiki/Yi_Sang_Literary_Award
- Edmund Wilson, "The Strange Case of Pushkin and Nabokov," The New York Review of Books, 15 July 1965 — <https://www.nybooks.com/articles/1965/07/15/the-strange-case-of-pushkin-and-nabokov/>
- "Letters: The Strange Case of Nabokov and Wilson," NYRB, 26 August 1965 — <https://www.nybooks.com/articles/1965/08/26/letters-the-strange-case-of-nabokov-and-wilson/>
- Vladimir Nabokov, "The Art of Translation," The New Republic, 4 August 1941 — <https://newrepublic.com/article/113310/vladimir-nabokov-art-translation>
- "The Translator's Invisibility," Wikipedia — https://en.wikipedia.org/wiki/The_Translator%27s_Invisibility
- "Domestication and foreignization," Wikipedia — https://en.wikipedia.org/wiki/Domestication_and_foreignization
- Anthony Pym, Method in Translation History (St. Jerome, 1998) — <https://www.taylorfrancis.com/books/mono/10.4324/9781315760049/method-translation-history-anthony-pym>

11 • The Facing Page

In 1911 James Loeb laid down a single rule of page layout: the Greek or Latin original on the left-hand page, the English translation on the right-hand page facing it. The first twenty volumes appeared in 1912, published simultaneously by Heinemann in Britain and Macmillan in America. The format has a name of its own: en-face.

Three more publishers picked up the same format over the century that followed. The I Tatti Renaissance Library began in 2001, Latin on the left, English on the right. The Clay Sanskrit Library released its first twelve titles in February 2005, setting Sanskrit in Roman transliteration on the left-hand page, and closed the series at 56 volumes in 2009.

The Murty Classical Library of India began in 2015 and stated its reasoning on its own web page — "the page spread is designed so text and translation reflect each other rather than allowing one to decide the other's fate." The companion faces for all six Indic scripts were commissioned from John Hudson and Fiona Ross at Tiro Typeworks. One studio, one brief.

These four publishers did not invent the same format four separate times. Clay states that it was modelled on Loeb, and Harvard University Press says that every facing-page series it has built since follows Loeb's model. This is inheritance, not four independent discoveries. What the evidence does establish is something else: en-face was carried onto scripts from entirely different worlds — Greek, Latin, Sanskrit in transliteration, seven Indic scripts — and it worked every time. Along the way the trim size changed, the typefaces changed, the cover colours changed. What nobody changed was the facing page. The thing that survives while everything around it is altered is the thing carrying the load.

Thai makes a page like that more expensive, because Thai puts no space between words. The W3C's Thai Layout Requirements of 30 April 2024,

edited by Richard Ishida, states it plainly: "Spaces separate phrases, rather than words." The consequence is that UAX #14, Unicode's line-breaking algorithm, hands Thai, Lao, Khmer and Burmese off to a separate dictionary-based word segmenter. ICU does this work with its DictionaryBasedBreakIterator. Ordinary page-layout software does not do it for you, so Thai copy has to pass through a segmenter first.

The ambiguity is a property of the language itself, not a defect in the software. Wirote Aroonmanakun of Chulalongkorn University measured it in 2007: he asked twelve Thai speakers to mark, on the same page of text, where each sentence ended. Of 59 candidate positions, only 15 drew agreement from 10 to 12 of them, and 24 drew agreement from only 1 to 3. His conclusion was that "the notion of Thai sentence is fuzzy" — for native readers, not merely for machines. Where the dictionary breaks the text wrongly, the compositor fixes it by hand, with ZWSP (U+200B) and WJ (U+2060).

The second problem is vertical. Thai sits on the same baseline as Latin; Ishida writes that "Thai uses the so-called 'alphabetic' baseline, which is the same as for Latin and many other scripts." The difficulty is therefore not aligning the two languages' lines but the space above them. A composite Thai vowel can involve as many as four glyphs surrounding a base consonant on three sides, a tone mark stacks above that again, and it rides higher still when a vowel sign already sits beneath it. Ishida's summary is short: "Thai tends to add more interline spacing than Latin text does."

This is the point at which this book has to stop and say it does not know. Two Thai line-height ratios circulate on the web: one says at least 1.55, the other says 1.8 or more. Neither traces back to a foundry, a named type designer, or an attributable study; neither appears in the W3C or Ishida documentation; and they contradict each other by a wide margin. One of the two sources announces itself as "the definitive English-language resource for Thai graphic design" while carrying the marks of AI-generated SEO content. So this book prints neither figure. A book

about what cannot cross does not get to fake a number about setting two scripts together.

Two things survive. The first is a direction that has a source. Noto splits its fonts into two families: the standard one, with the generous line height that continuous reading requires, and the UI one, compressed for screens. The generous family is the default. The second is the one figure with a name attached to it. Ishii Genta, DTP Director at Human Science Corporation, reports that InDesign's auto-leading default differs by the application's own language — 175% in the Japanese version against 120% in the English one. That is Japanese, not Thai. For Thai, nobody has yet written anything at the level of the page: the "Page & book layout" section of that same W3C document contains one word, "tbd."

The volume you are holding is not set en-face. It is set in alternating paragraphs — one Thai paragraph, one English paragraph, in a single column. I chose it because our production line could already do it, not because it is better. The cost is that a reader cannot read either language as continuous prose. Every time a paragraph ends, the other language gets in the way.

I cannot tell you how large that cost is. No study comparing alternating paragraphs against the other two layouts was found. The one comparison that does exist set interlinear against facing-page — 38 Korean university students read both — and found comprehension at roughly the same level, which means en-face does not trade comprehension away for elegance. As for the read-so-many-percent-faster figures circulating in advertising copy, they trace back to no study at all.

This is the closing volume of one generation of books from this workshop. The previous generation wrote English first and rendered Thai from it, and the result was Thai that was correct in every word and lifeless. This generation reversed the direction on purpose: Thai first, English as the render. The reason is not love of the language, it is instrumentation. This workshop has an instrument for measuring the quality of Thai and does not yet have one for English, so the language we

cannot yet measure was put on the render side. That explanation expires the moment the second instrument exists.

The next generation is to be beautiful in both languages at once. That sentence carries no weight as a wish. A specification differs from a wish in saying what will be built, with what numbers, and who can check that it was done.

1. **Layout.** Use en-face, Thai always verso, English always recto, throughout the volume; open any spread and you can check it. Sourced, from the four publishers above. Rejecting alternating paragraphs is a judgment call, not an experimental result.

2. **The typeface pair.** Thai and Latin must come from one brief. Either use a family whose foundry has published its own pairing rationale, such as Cadson Demak's Prompt, or commission a new pair from a single studio on the Murty model. The failing side is already documented: in Angsana New the Latin x-height falls below the height of the Thai **น**, and Thai users have been setting Latin at 12 pt against Thai at 16 pt by hand for 23 years.

3. **Point size.** Set Thai roughly 10 to 20% larger than English, then correct by eye on the actual pair, because the base height of Thai letters falls between Latin's x-height and cap-height. That range is a judgment call, not a figure anyone looked up.

4. **Line height.** Let the Thai column set the shared grid and the English column absorb the surplus. The direction is sourced, from both Ishida and Noto. The figure of 150 to 160% is a judgment call, not a figure anyone looked up, and must be tested against copy carrying stacked tone marks before it is locked.

5. **Thai line-breaking.** All Thai copy must pass through a dictionary-based segmenter, ICU or equivalent, before it reaches the layout program, with wrong breaks corrected by hand using ZWSP and WJ. Justify by stretching the space between phrases, not between words. Machine-checkable.

6. Language tagging. Every paragraph must carry the Lang value of its own language, per PDF19 and H58, to satisfy WCAG 3.1.2. Never one tag across a spread. The file must conform to PDF/UA. Machine-checkable.

Items 1, 5 and 6 can be checked by machine today. Items 2, 3 and 4 cannot: a sample spread has to be printed and put in front of Thai readers first.

This book does not fully do what it argues for. It argues about what grammar obliges, and then sets its two languages so that they interrupt each other on every page. The six items above are a debt this volume incurred and cannot pay itself. The next volume pays it, and will know whether it has paid in full, because somebody wrote down in advance what was owed – which items have numbers, and which do not.

What could not cross: Line height could not cross, because Thai orthography obliges a tone mark to stack above a vowel sign that is already stacked on a consonant – vertical space Latin never had to ask for. Nor could the line break, because Thai does not write word boundaries into the text. What Latin gets free from its spaces, Thai has to buy with a dictionary.

Overtured if: Someone sets the same bilingual text both ways – en-face and alternating paragraphs – measures comprehension and reading speed with Thai readers, and finds no difference; item 1 falls immediately. And if a foundry or a standards body publishes a Thai-to-Latin line-height ratio with a name and a method attached to it, items 3 and 4 stop being judgment calls and become figures that can be looked up.

Sources: Loeb Classical Library, publisher history <https://www.loebclassics.com/page/history> · I Tatti Renaissance Library <https://itatti.harvard.edu/publications/book-series/i-tatti-renaissance-library> · Clay Sanskrit Library https://en.wikipedia.org/wiki/Clay_Sanskrit_Library · Murty Classical Library of India, design and typography <https://www.murtylibrary.com/about/design-and->

typography · W3C, Thai Layout Requirements, Group Draft Note, 30 April 2024 <https://www.w3.org/TR/2024/DNOTE-thai-lreq-20240430/> · Richard Ishida, Thai Orthography Notes <https://r12a.github.io/scripts/thai/th.html> · UAX #14, Unicode Line Breaking Algorithm <https://unicode.org/reports/tr14/tr14-51.html> · ICU, Boundary Analysis <https://unicode-org.github.io/icu/userguide/boundaryanalysis/> · Wirote Aroonmanakun, "Thoughts on Word and Sentence Segmentation in Thai," SNLP 2007 <http://pioneer.netserv.chula.ac.th/~awiwrote/ling/snlp2007-wiwrote.pdf> · Microsoft typography-issues #1116 on Angsana New, January 2024 <https://github.com/MicrosoftDocs/typography-issues/issues/1116> · Cadson Demak, Prompt <https://github.com/cadsondemak/prompt> · Type Together, Beginner's Guide to Multiscript Typography <https://www.type-together.com/multiscript-typography-guide> · Noto usage guidance <https://notofonts.github.io/noto-docs/website/use/> · Ishii Genta, Human Science Corporation <https://www.science.co.jp/en/nmt/blog/27655/> · the interlinear-versus-parallel comparison with 38 Korean students https://www.researchgate.net/publication/337704420_The_Effects_of_Dual-language_L2_Texts_on_English_Reading_Comprehension_and_Text-type_Preferences · WCAG 2.0 Understanding SC 3.1.2 Language of Parts <https://www.w3.org/TR/UNDERSTANDING-WCAG20/meaning-other-lang-id.html> · WCAG technique PDF19 <https://www.w3.org/TR/2016/NOTE-WCAG20-TECHS-20161007/PDF19> · WCAG technique H58 <https://www.w3.org/TR/2016/NOTE-WCAG20-TECHS-20161007/H58> · the two sources whose figures this book does not accept, named so they can be checked <https://www.samui-infotech.com/the-importance-of-thai-typography-in-web-design/> and <https://thaigraph.com/faq/>