

Dar a luz una lengua: anatomía de las lenguas inventadas y de todo lo que las hace (por ahora) humanas

Jeqvmel kiñe zugun: Nentuke zugun tañi norkvnun ka kom ci zugu (fanten mew) celzuguñmafi egvn

Giving Birth to a Language: An Anatomy of Constructed Languages and Everything That Makes Them Human (For Now)

Nicolás Arellano

Universidad de Buenos Aires

Email

nicolas.a.arellano@gmail.com

ORCID<https://orcid.org/0000-0002-5197-5428>

RESUMEN. En este ensayo presentamos una introducción al fenómeno de las lenguas construidas o *conlangs*, desde su historia y clasificación hasta los desafíos que implica su creación. Partimos de describir los principales tipos (lenguas auxiliares, ingenierizadas y artísticas) y se presentan las decisiones fundamentales que todo *conlanger* debe tomar en los niveles fonológico, morfológico, sintáctico y léxico. A partir de la irrupción de las *genlangs*, lenguas generadas automáticamente por modelos de inteligencia artificial, explicamos los beneficios de tener en cuenta elementos sociales e históricos en la creación de lenguas, especialmente particularidades que muestran la variación según el hablante, irregularidades como huellas del cambio lingüístico y estrategias de evasión léxica como el *algospeak*.

Palabras clave: lenguas construidas, lenguas auxiliares, lenguas ingenierizadas, lenguas artísticas, *genlangs*

PICIWIRIKVNUN. Tvfaci pepilpepiltun wirin mew pepi konam ti zewma zugunvn zugu (*conlangs* ingvlezugun mew), kuyfitu kvpal ci ka ñi xvrvmel ta ka ñi kvzawgen tañi nentuel kvtu. Tvfa mew pegengeayin kiñe mufv wvneltu (kejuke zuguwigvn, kvzaw zuguwigvn ka az zuguwigvn) ka pu nentu kvzawfegelu tañi maypiyael ta fonoloqikugekelu zugu, morfoloqikugekelu zugu, sintatikugekelu zugu ka lesikugekelu zugu ta kimvrpuayin tati. Wefpvratulu ci zugun xipakelu celkvzaw tañi rakizuum mew, fey ta kvme zugu jemay tañi pu ce ñi kuyfi kimvn nvwgekelu ci zugu fey ta pegeluwayin tati, kizu tañi cemkvn pegekelu pu zugunfe ñi kakvnuwvn zugun mew, tañi xvrgelan egvn fey ta kakvnurpuy ñi zugun ka tañi cem nor rvpvtun tañi kuñil zugunegvn.

Rvftu nemvl: Zewma zugunvn, az zugun, zugun xipakelu celkvzaw tañi rakizuum mew, kakvnuwvn zugun¹.

¹La traducción al mapuzugun de título, resumen y palabras clave fue realizada por Pedro Cariman, integrante de *Kizu Iñciñ. Organización mapuche en la UNKo*.

1 | INTRODUCCIÓN

Imaginemos por un momento que te invitaron a un torneo poético (un *slam* de poesía que le dicen), un viernes a la noche. Subió al escenario alguien que no conocés, empieza a hablar y algo te llama la atención: no entendés nada. No es que las palabras estén en otro idioma que reconocés pero no hablás, como cuando quizás escuchás italiano o portugués, incluso alemán, inglés y quién te dice guaraní, y podés seguir el hilo aunque no sepas cada palabra. Es que no



Fuente: *Glotoida*. <https://glotoida.neocities.org/symbol>

reconocés absolutamente nada, ni el ritmo, ni los sonidos, ni la estructura de las frases. Después, cuando baja del escenario, te acercás y le preguntás de dónde es. Te mira con una sonrisa y te dice que es de acá, del barrio, que nació a diez cuadras. Entonces preguntás lo que obviamente hay que preguntar: qué idioma estaba hablando. Y te dice el nombre. Puede que lo reconozcas vagamente, como algo que leíste una vez o viste en una nota, pero no sabés bien qué es. Le preguntás si es una lengua de su familia, si sus padres la hablaban, que dónde se habla. Y te dice que no, que esa lengua la inventó alguien, quizá incluso alguien que nunca conoció y con quien nunca habló, o quizá ella misma, o tal vez es la lengua que hablan en esa serie que siempre te recomendaron pero que nunca encontraste el tiempo para empezar, que es una lengua que mucha gente estudia y que hasta quizás otros escriben poemas, manifiestos o llenan sus diarios íntimos con ella. En ese momento, lo más probable es que un primer juicio pase por tu cabeza. Al menos, es el que pasaría por la mía. No digo que sea un juicio necesariamente hostil, pero sí una especie de jerarquía implícita. Una lengua inventada huele, intuitivamente, a un artificio, a un ejercicio de imaginación sin dudas muy interesante pero que no puede equipararse a la lengua que hablo desde que nací, a la que hablaban mis abuelos, a la que aprendí en la escuela o a la que todavía no tengo el dinero para comenzar a estudiar. ¿De qué me serviría? ¿Por qué alguien decidió escribir y recitar ese poema con esa secuencia particular de sonidos que solo un grupo reducido de personas puede entender e incluso quizá nadie en el torneo de poesía?

Esa intuición merece ser puesta a prueba. Si nos detenemos a pensar qué significa exactamente que una lengua sea “natural” y no “artificial”, las cosas se complican bastante rápido. Todas las lenguas tienen un origen, un momento en que ciertos sonidos (o ciertas señas en las lenguas de señas²) empezaron a significar ciertas cosas para ciertos grupos de personas, y ese origen no deja de ser un hallazgo humano, tan arbitrario y contingente como cualquier otro. Lo que llamamos lenguas naturales son el resultado de siglos o milenios de modificaciones acumuladas de a poco, de contactos entre comunidades, de olvidos colectivos, de creatividades individuales que nadie registró. Una lengua construida o artificial es, en ese sentido, una lengua con esa misma historia vuelta visible, gestionada por alguien con nombre y apellido en lugar de por generaciones de hablantes anónimos.

A estas lenguas construidas o artificiales (que se llaman *conlangs* en la bibliografía especializada y así también se las conoce en la aguerrida comunidad que las defiende e impulsa, del inglés *constructed languages*) las une el gesto deliberado de su creación, aunque las separe casi todo lo demás: sus propósitos, sus alcances, sus comunidades de uso y la pregunta de fondo que cada una intenta responder. Pero ¿por qué alguien se tomaría

² Antes de seguir, vale aclarar que cuando hablamos de lenguas naturales incluimos también a las lenguas de señas, que son lenguas completas con su propia gramática, su propia historia y sus propias comunidades de hablantes, y que no son versiones manuales de ninguna lengua oral. Que conste que en este humilde artículo no nos metemos con la posibilidad de que haya lenguas de señas construidas, aunque quizás recuerden varias escenas de los personajes de Duna (*Dune*) en que los miembros de la Casa Atreides se comunican por señas (ya veremos cómo el creador de esta lengua, David Peterson, es una figurita repetida del mundo de las lenguas inventadas).

el trabajo de construir una lengua desde cero y qué decisiones implica ese trabajo? ¿Hay cosas en común entre el esperanto, las lenguas élficas de Tolkien y el dothraki de Juego de Tronos (*Game of Thrones*)? ¿Qué puede aprender un estudiante de lingüística o un apasionado de las lenguas construyendo su propia lengua que no podría estudiando otras lenguas “naturales”? ¿Y qué dice sobre el lenguaje humano el hecho de que haya cosas que sean difíciles de replicar para las lenguas construidas y las inteligencias artificiales, por muy elaboradas que sean?

2 | DIME QUÉ LENGUA INVENTÁS Y TE DIRÉ QUIÉN ERES

A pesar de que la idea detrás de lo que significa *conlang* es bastante directa, hay algunos matices interesantes en los que concentrarse. Quizás dos de los ejemplos más famosos (o, por lo menos, con los que más suerte voy a tener en que los conozcas) son el esperanto, la lengua planificada más popular del mundo ([¡e incluso con hablantes nativos!](#)) y el dothraki, una de las [lenguas ficticias del universo de Juego de Tronos](#) (aunque no podemos culpar a la gramática del dothraki por el final de la serie). Para entender de qué estamos hablando cuando hablamos de lenguas inventadas, entonces, conviene empezar por ahí: por la pregunta de para qué se inventó cada una, porque esa respuesta, en muchos casos, determina lo demás.

La categoría más antigua y quizás más ambiciosa dentro de los posibles tipos de *conlangs* es la de las lenguas auxiliares, llamadas *auxlangs*, pensadas para resolver un problema de comunicación entre personas que no comparten lengua. Ya desde aquí partimos con una tensión, sobre todo si pensamos en que cualquier decisión sobre qué lengua adoptar como común beneficia a alguien y perjudica a otros. Si lo pensamos desde nuestra historia, sabemos que el español llegó desplazando lenguas que ya estaban y que las lenguas de migración (el italiano, el gallego, el yiddish, entre tantas otras opciones) nunca tuvieron el mismo estatus aunque millones las hablaran en sus casas. La utopía de la lengua auxiliar es exactamente lo contrario a esta idea: un idioma que no sea de nadie o que sea de todos por igual. El esperanto, desarrollado por el médico polaco L. L. Zamenhof y cuyo manifiesto fue publicado en 1887, es el intento más exitoso de materializar ese deseo.

Zamenhof construyó una lengua con una gramática completamente regular, sin casi frases idiomáticas ni verbos irregulares, con un sistema de afijos que permite formar palabras nuevas a partir de raíces conocidas, tomando su vocabulario principalmente de las lenguas romances con aportes germánicos y eslavos. Hoy, casi ciento cuarenta años después, el esperanto tiene comunidades activas en decenas de países e incluso en Internet (<https://esperanto.masto.host/>), [su propio curso en Duolingo](#), de donde saqué la frase que aparece en (1) y la Imagen 2, y algo que pocas lenguas construidas pueden decir que tienen: hablantes que lo aprendieron de sus padres desde la infancia, para quienes fue la primera lengua antes que cualquier otra (incluyendo al músico Kim J. Henriksen, que cuenta con [un canal de Youtube](#) donde sube sus composiciones en esta lengua).

(1) *Li volas kreskigi plantojn en la gârdeno*

‘Él quiere cultivar plantas en el jardín’

Las *engelang*s (que más o menos se podrían traducir como lenguas ingenierizadas desde el inglés *engineered languages*) parten de una motivación diferente: no comunicar “mejor” sino pensar diferente. O, al menos, poner a prueba algo sobre la relación entre la estructura de una lengua y la manera en que sus hablantes perciben el mundo. Esa búsqueda tiene nombre propio en lingüística: la hipótesis Sapir-Whorf, que en su versión más fuerte

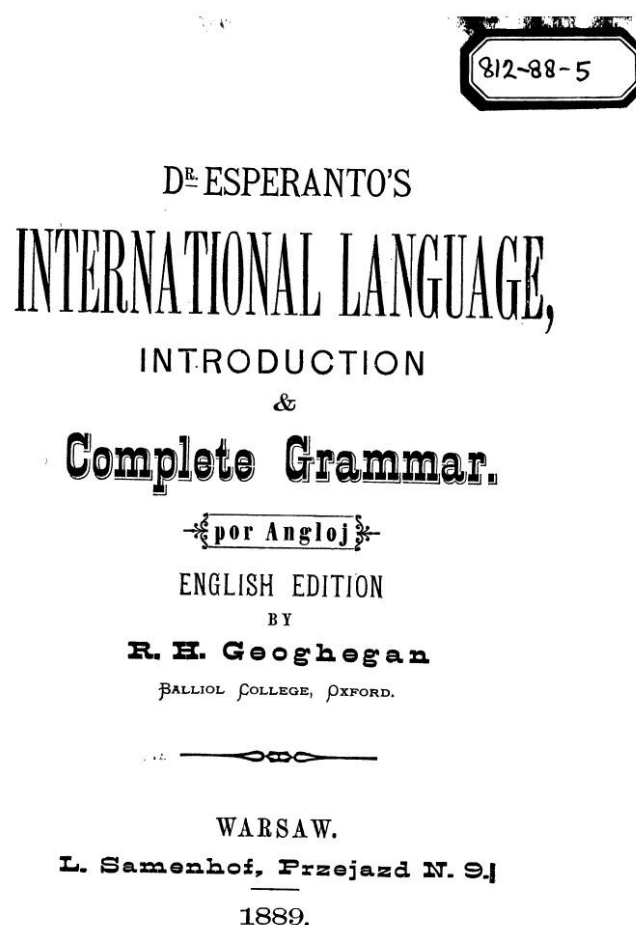


IMAGEN 1 Portada en inglés de *Dr. Esperanto's International Language*, traducción de 1889 de *Unua Libro*, original publicado por Zamenhof en 1887.

sostiene que la lengua que hablamos determina o condiciona activamente cómo pensamos. Es una hipótesis que genera debate desde hace décadas y que la evidencia empírica ha matizado bastante, pero que sigue siendo lo suficientemente provocadora como para inspirar proyectos enteros. El [loglan](#), creado en 1955 por el lingüista James Cooke Brown, fue diseñado explícitamente para ponerla a prueba. La idea era construir una lengua basada en lógica de predicados, que es la manera en que la filosofía y las matemáticas formalizan los enunciados: en lugar de decir “Juan corre”, un sistema lógico diría algo como “existe un individuo que se llama Juan y del cual es verdad el predicado correr” (recuerdos de guerra para muchos estudiantes de semántica formal o lógica e incluso pensamiento científico). Traducido a una lengua, eso significa que cada oración tiene una estructura unívoca, sin ambigüedades posibles. La pregunta de fondo seguía siendo la de Sapir-Whorf: si le damos a alguien una lengua así de precisa, ¿cambia la manera en que piensa? Quizás si ves la peli *La llegada* (*Arrival*), la de Amy Adams, podrías tener una noción un poco más práctica y artística sobre este interrogante.

Las *artlangs*, lenguas construidas artísticas, son probablemente las más visibles en la cultura popular y las que mejor ilustran hasta dónde puede llegar la deliberación en el diseño de una lengua. Es casi como si cada generación tuviera la suya: los frikis de la Generación X [tienen al klingon](#), los milenials crecieron con las [lenguas élficas de Tolkien](#) revividas en el cine, y la generación siguiente recibió [el alto valyrio](#) y [el na'vi](#)

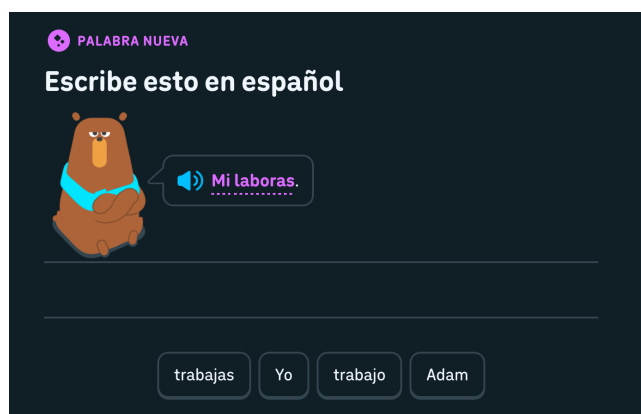


IMAGEN 2 Captura de muestra del curso de español de Duolingo.

(la de *Avatar*) más o menos al mismo tiempo (parece que últimamente hay más salida laboral creando lenguas ficticiales en Hollywood que dando clases en la universidad). Cuando le pidieron al lingüista Marc Okrand que se encargara del klingon, eligió sonidos y estructuras que reforzaran la imagen de una civilización guerrera, agresiva³, algo bien distante de cualquier lengua conocida. Cada decisión gramatical era también una decisión narrativa sobre quiénes eran los klingon y cómo organizaban su experiencia del mundo. Lo mismo ocurrió con el dothraki y el alto valyrio de Juego de Tronos, donde David Peterson construyó sistemas completos a partir de los escasos fragmentos que aparecían en las novelas de George R. R. Martin, y con el na'vi de *Avatar*, desarrollado por el lingüista Paul Frommer, que diseñó una lengua con rasgos tipológicos poco frecuentes en las lenguas europeas (como la evidencialidad, que es la marcación gramatical de si el hablante vio algo con sus propios sentidos o lo infirió de otra fuente, algo que el quechua⁴, entre otras lenguas, hace con total naturalidad) para reforzar la imagen de una civilización con una relación con un conocimiento radicalmente distinta a la humana. Este trabajo es uno que la mayoría de los espectadores no percibe conscientemente, pero que sostiene la credibilidad de universos enteros y alimenta las comunidades de fans que mantienen vivas las creaciones mucho después del estreno.

3 | LA BASE ESTÁ

Lingüistas muy importantes y el mismísimo Tolkien fueron citados en el apartado anterior. Asumo que esto te puede dar la pauta de la complejidad que trae la creación de una lengua y la absurda cantidad de decisiones que se deben tomar. Las preguntas podrían continuar abriéndose hasta el punto en que es difícil ver el fin. Pensá que en algún momento te enteraste de que no todas las lenguas tienen el orden de palabras Sujeto-Verbo-Objeto (SVO) como tenemos en español (quizá te estás enterando ahora), ya sea porque estudiaste alemán, japonés o viste un TikTok hace unos días. Decidís que querés algo distinto para tu *conlang* y proponés un orden SOV (que, de hecho, **es el más frecuente translingüísticamente**). ¿Se acabaron las decisiones en torno al orden? ¿El adjetivo lo ponés antes que el sustantivo o después? ¿O depende del tipo de adjetivo o del nombre? ¿Qué hacés con la preposición? ¿Antes de los nombres o después (que ya serían posposiciones)? Y si querés darle un criterio de naturalidad, estaría bueno saber que generalmente las lenguas con orden OV tienden a ser más posposicionales

³Asociar sistemáticamente ciertos sonidos con texturas, tamaños y características específicas se conoce como fonosimbolismo. Las dos celebridades del campo son *kiki* y *bouba*, pero si querés conocer un poco más, te sugerimos visitar el siguiente enlace: <https://theconversation.com/orcos-rugidos-y-musica-elfica-el-fonosimbolismo-en-accion-275479>.

⁴En quechua, por ejemplo, encontramos el ítem *mi* cuando el hablante quiere reforzar la veracidad y validez de lo que está expresando (*noqa-pas-mi* 'yo también (ciertamente)'). Pueden ver trabajos de Mayra Juanatey (2024) para conocer más sobre los evidenciales en esta lengua.

que preposicionales en comparación con las lenguas VO. Por supuesto, este jardín de senderos que se bifurcan no se acaba con la muy interesante decisión del orden de palabras, sino que abarca también los sonidos, cómo esos sonidos se combinan en palabras, cómo esas palabras se organizan en oraciones y, finalmente, qué cosas nombra la lengua y cómo las nombra. Como si fuera un curso de gramática universitario, cada una de esas capas tiene su propia lógica y sus propias (casi que a veces infinitas) posibilidades.

3.1 | A caballo regalado se le miran los dientes

La fonología es el punto de partida porque es la materia prima de todo lo demás. Antes de tener palabras hay que tener sonidos y antes de tener sonidos hay que decidir cuáles. Todas las lenguas orales del mundo usan un subconjunto de los sonidos que el aparato fonador humano puede producir y ese subconjunto puede variar enormemente, incluso si se comparan variedades del español: el español medieval tenía fonemas que hoy desaparecieron, como la distinción entre una *s* sorda [s] y una *s* sonora [z], que en algunos contextos se escribían con el mismo grafema *s* pero sonaban distinto. Además, es sabido que el español peninsular estándar conserva hasta hoy la diferencia entre [s] y la pronunciación de las sílabas *ce*, *ci*, *za*, *zo*, *zu* [θ] (bueno, en Argentina también, solo cuando algunas personas con apellido con Z tienen que exagerar la pronunciación de su nombre⁵). Algunos casos aún más ilustrativos pueden encontrarse revisando las gramáticas de lenguas no europeas occidentales. El hawaiano es conocido por tener solo ocho sonidos consonánticos, mientras que lenguas del Cáucaso como el *abjasio* superan los cincuenta. No solo la cantidad de sonidos puede distar de lo que estamos acostumbrados: son hartos conocidos los casos de lenguas que tienen chasquidos consonánticos o clics, como el *xhosa* o el zulú. En español, hacemos un sonido parecido para expresar desaprobación o incredulidad cuando pegamos la lengua contra el paladar (justo detrás de los dientes de arriba) y después la despegamos rápido, como “chupando” un poco de aire para adentro) (te imagino ahora tratando de seguir estos pasos para emular el sonido, muy bien) (2).

(2) – ¡Me dijo que se recibió de médico!

– ¡Tsk! ¡Qué va a ser médico ese mentiroso!

Otras lenguas podrían usar el tono para distinguir palabras o morfemas, de manera que la misma secuencia de sonidos significa cosas completamente distintas según el tono, como ocurre en el vietnamita o en el mandarín. Para esta última lengua, podemos ilustrar con el infame poema *El poeta comeleones en la guarida de piedras* (施氏食獅史), romanizado como Shīshì shí shī shǐ, y publicado por Yuen Ren Chao, en los que la única diferencia de pronunciación en los caracteres radica en el tono elegido, ya sea plano, descendente, ascendente o bien descendente-ascendente⁶.

⁵ Apenas termines de leer este artículo, te pido que revises el que escribieron Gonzalo Espinosa y María Mare (2025) dedicado a lo mucho que podemos aprender a partir de la reflexión ortográfica en nuestra lengua: <https://revela.uncoma.edu.ar/index.php/lingustica/es/article/view/6594> (obviamente, en *Quintú Quimün*).

⁶ Los lingüistas tenemos dos o tres trucos nada más para mostrar, así que debo decirte que Noelia Stetie (2025) explicó esto con más detalle en una edición pasada –también, dónde si no– en *Quintú Quimün*: <https://revela.uncoma.edu.ar/index.php/lingustica/es/article/view/6811>.

3.2 | Importarán las formas y la piel que te pongas

¡Pero /oxota/! El inventario de sonidos es solo una parte de la fonología⁷. También hay que decidir cómo esos sonidos pueden combinarse, qué secuencias son posibles y cuáles no. En español, por ejemplo, una palabra puede empezar con *tr* o con *pr*, pero no con *tl*, aunque esa combinación existe en náhuatl (donde originalmente al chocolate se le dice *xocolātl*) y en otras tantas lenguas. Esas restricciones de combinación, llamadas fonotácticas, son las que le dan a una lengua su perfil sonoro característico, esa sensación de que algo “suena” a una lengua aunque no entiendas ni una palabra (y estas restricciones propias de nuestras lenguas a veces inconscientemente las trasladamos a otra que estamos aprendiendo y por eso nuestro “oído” está entrenado para saber si alguien que habla español como segunda lengua viene de Rusia, Italia o Brasil). Otros fenómenos interesantes son la armonía vocálica, en la que las vocales de una palabra se “contagian” rasgos entre sí y terminan pareciéndose, presente en el turco (*ev-ler* ‘las casas’, *kapı-lar* ‘puertas’, siendo *l_r* el sufijo de plural que modifica su vocal dependiendo de la base); y el debilitamiento, el proceso por el cual las consonantes se vuelven más frágiles en ciertos contextos, que en español podemos ver en la diferencia entre la *b* de “bota” y la *b* de “la bota” (repetilo en voz alta y te vas a dar cuenta; muy bien, sé que otra vez me hiciste caso).

La morfología es el nivel en que el trabajo de construcción se vuelve más intrincado. En fonología, el espacio de decisiones tiene un techo relativamente claro: los sonidos que el aparato fonador humano puede producir están más o menos identificados en su totalidad. En cambio, en morfología el horizonte se abre mucho más: las combinaciones posibles de estrategias, categorías y tipos de marcación son tan vastas que es difícil siquiera saber por dónde empezar. Las lenguas del mundo ofrecen un abanico extraordinario de estrategias en la formación de palabras y en cómo se combinan las piezas que las componen. En el universo de las *conlangs*, el klingon es un caso particularmente ilustrativo de hasta dónde puede llegar la complejidad al interior de la palabra. Tomemos la palabra *Qagh-Hom-mey-Hey-llj-mo'* (separamos con guiones para que sea más fácil de leer), que significa aproximadamente “debido a tus aparentes errores menores”: *Qagh* es “error”, *Hom* lo hace menor, *mey* marca el plural, *Hey* indica que es aparente, *llj* es el posesivo de segunda persona y *mo'* expresa causa (ejemplo tomado de Richards 2018). Seis morfemas apilados con precisión, cuya traducción al español requiere de la presencia de por lo menos seis palabras distintas (quizás cinco si decidimos traducir como *errorcitos* o considerar a *debido a* como una única palabra).

Finalmente, la sintaxis determina cómo se organizan las palabras en oraciones y qué relaciones se establecen entre ellas. El orden de palabras (como ya te conté) es la decisión más visible, pero hay muchas otras: cómo se forman las preguntas, cómo negamos cosas, cómo hacemos para unir varias ideas en una misma oración, qué pasa si una lengua marca cómo terminan las palabras para marcar su función sintáctica (como hacemos cuando pasamos de *yo* a *me* y *mí*), si existen fenómenos de concordancia (como en español, que no podemos decir **el mesa rojos*) y tantos otros rasgos más.

⁷ Otro dato no menor que estamos obviando es que estamos utilizando la notación ortográfica latina para marcar los sonidos y grafemas de las letras, ¿pero qué impediría a una lengua desarrollar su propio sistema de escritura? Ahí también deberíamos pensar en las posibles asociaciones entre grafemas y fonemas que tendría la lengua construida si es alfabética. Otros problemas surgirían si decidimos seguir un tipo de escritura más silábica (como en algunos sistemas de escritura del japonés) o logográficos (como en el mandarín). En el ámbito de las *conlangs*, se conocen específicamente como *neografías* (el más famoso siendo el Tengwar, diseñado —nuevamente— por Tolkien, y que seguro viste bien con los primeros planos al anillo cada vez que Frodo se mandaba alguna). Es el sistema de escritura que aparecía grabado en el anillo. De hecho, se encuentran en algunas versiones de Unicode y en algunas tipografías disponibles en línea (aunque no en la que usamos aquí en *Quintú Quimiin*).

3.3 | Sin palabras no hay paraíso

Por último, también debe tenerse en cuenta el vocabulario, que es donde aparecen aspectos propios de una cultura. Por ejemplo, los sistemas de color varían enormemente entre lenguas. En ruso y en español hay dos palabras distintas para lo que en muchas otras lenguas solo se considera azul (en ruso, *goluboy*⁸ sería el celeste y *siniy*, el azul más oscuro), mientras que muchas lenguas agrupan bajo una sola etiqueta lo que el español separa en “verde” y “azul” (estas serían las lenguas *grue*, un *portmanteau* de *green* y *blue*, como cuando cada vez que hay un *shippeo* y se arma una palabra como *Brangelina* u *Occhiamin*). Los sistemas de parentesco son otro caso clásico: el español usa la misma palabra para el hermano del padre y el hermano de la madre, mientras que lenguas como el tamil o algunas variedades del quechua tienen términos distintos o idénticos según la línea de parentesco (por ejemplo, *hermano* y *primo* comparten un único lexema). Los números también podrían contar una historia dentro de la *conlang*: es muy común que las lenguas tomen prestado el sistema numérico de otra y el resultado es a veces la coexistencia de dos organizaciones. Si revisamos el wichí, en esta lengua solo se registran como formas en desuso los equivalentes a los números uno al tres con un lexema patrimonial (por ejemplo, *hotefwasi* ‘tres’). Por el contrario, en la lengua actual, se recurre a préstamos del español adaptados a la fonología y fonotaxis del wichí (*tales* ‘tres’, *lies* ‘diez’, *sinku* ‘cinco’) (Nercesian 2014). Un buen punto de partida para no perderse en la vastedad del léxico es la lista de Swadesh, compilada por el lingüista Morris Swadesh en los años cincuenta: entre cien y doscientas palabras consideradas universalmente estables (partes del cuerpo, pronombres, números, fenómenos naturales, acciones cotidianas) que cualquier lengua debería poder nombrar. Antes de preocuparse por cómo se van a dar nombre a conceptos más complejos, conviene asegurarse de que la lengua pueda decir cosas como *mano*, *agua* y *morir*.

Todo esto (los sonidos, la morfología, la sintaxis y algunas particularidades de la elección léxica) es lo que cualquier manual de construcción de lenguas inventadas podría decirte que tenés que diseñar. Y es verdad: sin eso no hay lengua y con esto ya hay probablemente un sistema lo suficientemente robusto como para funcionar. Pero hay algo que las lenguas naturales tienen y que las construidas rara vez logran del todo y que tiene que ver menos con la gramática y más con la historia y lo que viene asociado al uso de una lengua.

4 | QUE PASE EL DESGRACIADO

Hasta hace muy poco, construir una lengua con el nivel de detalle y coherencia que describimos en la sección anterior era una tarea artesanal que podía llevar años o décadas o el trabajo mancomunado entre muchas personas. Tolkien trabajó con las lenguas del universo de El Señor de los Anillos (*The Lord of the Rings*) prácticamente toda su vida adulta.

Hoy esa ecuación cambió y de maneras que todavía estamos procesando. Los modelos de inteligencia artificial generativa pueden producir, bajo instrucción, algo que se parece bastante a una lengua completa: sonidos, gramática, vocabulario y textos en esa lengua con cierta consistencia interna. A estas lenguas generadas automáticamente se las empieza a llamar *genlangs*, aunque el fenómeno es tan reciente que todavía estamos viendo qué tan lejos puede llegar.

Hay un dato que sorprende cuando uno se pone a analizarlas: las *genlangs* cumplen la ley de Zipf. Esta ley, formulada por el lingüista George Kingsley Zipf en los años cuarenta, describe un patrón que se cumple en todas las lenguas naturales: la palabra más frecuente de cualquier texto aparece aproximadamente el doble de veces

⁸Justo esta palabra resulta interesante, porque una extensión de significado hoy ha generado que este mismo lexema sea utilizado como un término peyorativo o reivindicativo equivalente a *hombre gay* en esta misma lengua.

que la segunda más frecuente, el triple que la tercera y así sucesivamente. En español, “de” es la que aplasta a todo el resto. Es un patrón tan consistente que durante décadas se usó como una especie de sello de identificación de las lenguas humanas. Que las *genlangs* lo reproduzcan es, probablemente, solo una consecuencia de que los modelos de inteligencia artificial fueron entrenados con texto humano y no podrían no reproducirlo, pero aun así es interesante tenerlo en cuenta.

Entonces, la pregunta no es si una Inteligencia Artificial (IA) puede generar una lengua artificial, sino qué es lo que en ese producto resultaría insuficiente a simple vista y qué dimensiones quedan por fuera. Así, podríamos abordar la noble tarea de hacer una *conlang* desde una perspectiva global, histórica y con conocimiento de los múltiples fenómenos que ocurren en una lengua más allá de aquellos que son exclusivamente léxicos o gramaticales.

5 | HAY MÁS TELA PARA CORTAR

¿Qué pueden hacer los *conlangers* (aficionados, docentes, escritores, creadores de mundos ficticios o simplemente curiosos como yo que estoy escribiendo esto y vos que lo estás leyendo) para que una lengua inventada se sienta verdaderamente propia? La respuesta más honesta radica en prestar atención a lo que las lenguas hacen cuando nadie las está diseñando: mirar hacia afuera, a lo que algunas comunidades lingüísticas del mundo ya resolvieron de maneras que nunca se nos hubieran ocurrido; mirar hacia atrás, a cómo las lenguas cambian con el tiempo y qué huellas dejan esos cambios en su gramática; y mirar hacia el presente, a lo que está pasando ahora mismo en las lenguas que hablamos todos los días. Es decir, mirar en términos de lo que en un programa de lingüística reconoceríamos como tipología, historia y evolución de las lenguas y la variación en curso.

Si observamos otras lenguas, podemos constatar que hay todavía espacio para seguir innovando gramaticalmente en las decisiones de nuestras *conlangs*, ya que hay algunas que codifican, por ejemplo, variación según el sexo del hablante. No se trata de variaciones de género como las que observamos en español, asociadas al género de las palabras (si es masculino, si es neutro, si es femenino...), sino que tienen realmente en cuenta el sexo de quien habla. El kadiwéu, lengua de la familia guaycurú hablada en el Mato Grosso do Sul brasileño (emparentada con lenguas indígenas locales como el toba, el pilagá y el mocoví del Chaco argentino), es una de las que marca diferencias fonológicas y léxicas sistemáticas según el sexo de quien habla. Por ejemplo, los hombres dicen *niwenigi* para “comida”, las mujeres dicen *niweenigi*, con una vocal alargada que marca la diferencia. Para preguntar “¿qué más querés?”, un hombre diría *amaleeGini anemani* y una mujer *ameeGini anemani*. Para citar un caso probablemente más conocido, en el japonés sucedería algo parecido. En esta lengua asiática, existen palabras, pronombres, prefijos y partículas diferenciadas según si habla un hombre o una mujer (o incluso, si uno quiere hablar de una manera educada o más neutra): *hara* para “estómago” en la variedad masculina, *onaka* en la femenina.

Volviendo al cambio en clave histórica, algunas irregularidades gramaticales, sobre todo en sistemas verbales como el del español, tienen una explicación histórica. Los verbos más irregulares del español (pongámosle, *ser* e *ir* como mínimo, aunque deberíamos preguntarle a algún extranjero cuáles son los que más le cuestan aprender) llegaron a serlo precisamente porque fue en su uso frecuente que resistieron la presión analógica de regularizarse. Cuando un verbo se usa miles de veces al día, sus hablantes no necesitan construir sus formas siguiendo reglas generales sino que las aprenden directamente de memoria. Ese tironeo entre memoria y regla es visible todavía hoy en casos como *quepo* y *cabe*, dos formas del mismo verbo (*caber*), o como *anduve* y *andé*, en los que la presión analógica del paradigma regular lleva a muchos hablantes a producir una forma

que la norma rechaza pero que la lógica del sistema tendería a favorecer. Los mejores *conlangers* conocen estas particularidades y pueden proponer historias para sus lenguas que hipoteticen cambios fonológicos (por ejemplo) ocurridos en su pasado ficticio. Tolkien llevó esto hasta un nivel que los lingüistas históricos reconocen como genuinamente riguroso: el quenya y el sindarin comparten un proto-élfico, con diferencias tan sistemáticas y predecibles como las que existen entre el latín y sus lenguas romances. Un ejemplo: la raíz *TÁWAR* evolucionó por síncope vocálica (la pérdida de una vocal interior no acentuada, el mismo proceso que convirtió el latín *nativitatem* en *navidad*) a *taurē*, que produjo en quenya *taurë* y en sindarin *taur*, ambas con el significado de “bosque” (Fauskanger s.f.). A diferencia del *andé/anduve*, donde observamos el cambio ocurriendo en tiempo real y desde adentro, Tolkien lo simuló retrospectivamente: inventó una historia de sus lenguas. Por supuesto, la idea de corrección (qué forma debería usarse en una lengua estándar y por eso sabemos que *anduve* es la variante “correcta”) no está planteada en el universo de Tolkien, pero no descartamos que algún fan ya esté trabajando en reunir a un consejo de elfos, hombres, hobbits y enanos para fundar la Real Academia de la Tierra Media.

Por último, está lo más cercano, lo que no requiere buscar en gramáticas de otras lenguas ni en la reconstrucción de las lenguas porque está pasando ahora mismo. En plataformas de video en las que hay moderación normalmente automatizada, los creadores de contenido desarrollaron estrategias de evasión léxica para hablar de temas sensibles sin activar filtros de inteligencia artificial y censura, que pueden perjudicar la monetización o el alcance de sus piezas de comunicación. En inglés, estas estrategias ya son bien conocidas: *seggs en lugar de sex*, el emoji de maíz para referirse a contenido adulto a través de la cadena *porn* → *corn* → 🌽, *unalive* en lugar de cualquier referencia al suicidio. Esto es conocido en lingüística hoy como *algospeak* (algo de “algoritmo”)⁹. En el mundo de habla hispana el fenómeno tiene sus propios exponentes: en la comunidad que orbita el *streamer* argentino Martín Cirio, fueron instalados por el creador de contenido términos como *petisuis* para el suicidio, *magic* para las drogas, *Canadá* para la cárcel o *a chaca* para amenazar con la muerte, todos conceptos que las plataformas de *streaming* en teoría penalizan; hoy *estos lexemas* son ampliamente conocidos por sus seguidores y se integran al léxico de la vida cotidiana de miles de personas que lo consumen o no. Lo que hace al *algospeak* lingüísticamente interesante es que reproduce exactamente los mismos mecanismos de evasión que encontramos en culturas muy anteriores a internet: una presión externa empuja a la lengua a crear rodeos, esos rodeos se sedimentan, se comparten y eventualmente saltan al uso cotidiano. El tabú y el algoritmo operan con la misma lógica y es fácil que ninguno de los dos aparezca en los manuales de construcción de lenguas.

6 | OPAMA A ESTE ENSAYO

Entonces, ya sabés: la próxima vez que te inviten a ese *slam* de poesía y escuches algo que no reconocés, ya vas a tener un poco más claro el fenómeno que estés escuchando. Quizás te animes a subir al escenario en otra fecha, quizás te quedes abajo jugando un juego de mesa con tus amigos con tu nueva lengua construida o quizás simplemente te retires convencido de que las *conlangs* no son para vos. Cualquiera de las tres es una respuesta válida. Pero al menos ahora sabés que esa lengua que escuchaste tiene, o debería tener, una historia y fenómenos más allá de los estrictamente gramaticales.

Lo que este recorrido intentó mostrar es que las lenguas construidas son un fenómeno con más capas de lo que parece a primera vista (bueno, las lenguas naturales también): tienen una historia larga y diversa, responden a motivaciones muy distintas entre sí, implican decisiones técnicas serias sobre fonología, morfología, sintaxis y

⁹ Adiviná: esto ya fue explicado desde una perspectiva similar por Gabriela Zunino (2023) en –adivinaste– otro artículo de divulgación en *Quintú Quimün*: <https://revele.uncoma.edu.ar/index.php/linguistica/es/article/view/4720>.

léxico, y conectan con preguntas sobre el lenguaje que difícilmente surgirían de otra manera. Para el escritor que quiere que su mundo ficticio tenga un mayor peso, para el docente que busca que sus estudiantes entiendan por qué el español tiene géneros gramaticales y otras lenguas no, para el estudiante de lingüística que de repente entiende qué es la fonotaxis porque tuvo que decidir si su lengua permite grupos consonánticos al inicio de sílaba, construir una lengua es una de las formas más directas de entender cómo funcionan las que ya existen.

Las *genlangs* ponen sobre la mesa una pregunta útil: si una IA puede generar en minutos el andamiaje estructural de una lengua, ¿qué es exactamente lo que les falta? Lo que fuimos viendo a lo largo de este aporte sugiere que les faltaría todo lo que las lenguas acumulan cuando nadie las está diseñando: la variación, las irregularidades, las palabras que nacen para rodear algo que no se podía decir directamente y otros tantos fenómenos que el espacio no nos permitió nombrar y vos podés seguir buscando luego de terminar este artículo. Les falta, en definitiva, lo que hace que valga la pena estudiarlas, enseñarlas, hablarlas o simplemente escucharlas un viernes a la noche en un *slam* de poesía. Y ese trabajo no es poco ni en vano.

TRABAJOS MENCIONADOS Y OTRAS LECTURAS PARA ENTUSIASTAS

- Espinosa, Gonzalo E. y María Mare (2025) ¿Va con «v» de vaca o «b» de burro? *Quintú Quimiün* 9(1), 1-10.
- Fauskanger, Helge Kåre (s.f.) *The evolution from primitive Elvish to Quenya*. Publicación online disponible en <http://ardalambion.net/>
- Juanatey, Mayra (2024) Variación del orden de palabras en el discurso en quichua santiagueño. *Nueva Revista del Pacífico* 81, 283–317.
- Language Creation Society (s.f.) *Fiat Lingua*. Publicación online disponible en <https://fiatlingua.org>.
- Nercesian, Verónica (2014) *La gramática wichí*. Publicación online disponible en <http://www.lenguawichi.com.ar/lengua-wichi/la-gramatica>
- Peterson, David J. (2015) *The Art of Language Invention*. Penguin Books.
- Peterson, David J. (2020) *Create Your Own Secret Language*. Odd Dot.
- Richards, Norvin W. (2018) *ConLangs: How To Construct A Language*. Notas de clase. Publicación online disponible en <https://ocw.mit.edu/courses/24-917-conlangs-how-to-construct-a-language-fall-2018/>
- Rosenfelder, Mark (s.f.) Zompist resources. Publicación online disponible en <https://www.zompist.com/resources/index.html>.
- Stetie, Noelia (2025) En defensa de la ambigüedad lingüística, la octava maravilla del mundo. *Quintú Quimiün* 9(2), 1-12.
- Zunino, Gabriela Mariel (2023) Kiwi melón, Susana Sacarina, el léxic0 m3nt4l y otras yerb*s sobre el procesamiento de palabras. *Quintú Quimiün* 7(1), 1-9.