

# Escribir con la boca: el dictado por voz como práctica de escritura

---

**Miriam Inza**

<https://orcid.org/0000-0001-6849-069X>

Depto. de Escultura, Arte y Tecnología, Facultad de Bellas Artes,  
Universidad del País Vasco/Euskal Herriko Unibertsitatea  
[miriam.inza@ehu.eus](mailto:miriam.inza@ehu.eus)

Recibido: 01.08.2022

Revisado: 17.10.2022

Publicado: 20.12.2022

Cómo citar este artículo

Inza,M., 2022. Escribir con la boca: el dictado por voz como práctica de escritura.

*Inmaterial. Diseño, Arte y Sociedad*, 7 (14), pp.104-116

DOI 10.46516/inmaterial.v7.149



## Resumen

Según Barthes, al intelectual se le pide constantemente que transcriba la exposición oral que ha hecho, como si eso no planteará ningún problema. Lo hace a través de un ejercicio de escritura y edición, nunca como mera transcripción. Realicé diversas ponencias en las que propuse un sistema de dictado por voz que me permitía exactamente eso: transcribir mi exposición oral a texto digital. El texto resultante era una masa informe y relegada a un segundo plano, dada la presencia ineludible de un cuerpo que habla. Para escribir este artículo, recuperé el mismo sistema, esta vez enfrentando el reto de mantener un registro propio de la escritura que favorezca al texto impreso, pero escribiendo desde la boca, alterando por completo la oralidad.

Las particularidades del dictado por voz (nivel de comprensión de la máquina, volumen, velocidad y dirección del hablante, longitud de las pausas, enunciación de los signos de puntuación...) provocan que ambas expresiones, la oral y la escrita, se vean alteradas por elementos propios de la otra: la voz da cuerpo a expresiones paralingüísticas mudas (punto, coma, guión), al tiempo que el texto registra de errores propios del habla y la escucha (lapsus, repetición, corrección).

El presente artículo da cuenta de los deslizamientos que provoca este modo de escribir.

**Palabras clave:** dictado por voz, escritura, error, audio-texto, performatividad.

## Writing with the mouth: voice dictation as writing practice

### Abstract

According to Barthes, the intellectual is constantly asked to transcribe the oral presentation he has made, as if that did not pose any problem. He does it through a writing and editing exercise, never as a mere transcription. I have made various presentations in which I proposed a voice dictation system that allowed me to do exactly that: transcribe my oral presentation into digital text. The resulting text have been a formless mass relegated to the background, due to the unavoidable presence of a speaker right there. To write this article, I recover the same system, this time facing the challenge of maintaining a writing register that favors the printed text, but writing it with the mouth, completely altering orality.

The particularities of voice dictation (comprehension level of the machine, volume, speed and diction of the speaker, length of pauses, enunciation of punctuation marks...) cause both expressions, oral and written, to be altered by elements of the other: the voice embodies silent paralinguistic expressions (dot, comma, hyphen), while the text registers errors of speech and listening (lapsus, repetition, correction).

This article gives an account of the slippings caused by this way of writing.

**Keywords:** voice dictation, writing, error, audio-text, performativity.

## Nota

Este artículo se compone de dos textos: el que puede leerse escrito y el que puede escucharse a continuación. El audio ha sido grabado en el momento en que se realizó el dictado por voz que ha escrito el presente texto. Los dos textos difieren. Leer preferentemente los dos a la vez.



Este audio se reproducirá en un enlace externo

Escribir este artículo, o cualquier otro, o una tesis, o cualquier texto, no se enfrenta al acto de escribir, el cual intentamos solventar mediante distintas estrategias. La estrategia que es cojo y ahora, que me permite escribir sin pulsar ninguna tecla, es la misma que utilicé en conferencias anteriores a este texto<sup>1</sup>. En aquellas conferencias orales, habría un mismo documento en red en mi teléfono móvil y en el ordenador que proyectaba lo que veían los asistentes y comenzaba a hablar. Con el dictado por voz activado, en el documento presentado se escribía todo lo que iba diciendo. Más o menos. La redacción de este artículo ocurre de manera similar, aunque lo que se recoge no es el pulso ahora de una ponencia, sino el dictado de un texto en registro escrito.

## **Hablar no es lo mismo que escribir**

La herramienta de dictado por voz transforma la voz hablada en texto digital y pretende facilitar la tarea de insertar mecánicamente y con los dedos cada una de las palabras que lo conforman. Cómo, en la mayoría de los casos como a las personas hablamos más rápido que escribimos o mecanografía vamos, el dictado por voz se suma a las diferentes estrategias en pos de la rapidez de la escritura, como son la cursiva, la propia letra minúscula toma la abreviatura y la ubicación de las letras en los teclados, por nombrar algunas. El uso de esta herramienta, al contrario que el resto de estrategias citadas, cambia el texto de lugar, trasladándose de la mano a la boca, dejando por ello de precisar la rigidez del cuerpo asociada a la práctica de la escritura. El texto se sitúa así más cerca de aquello que los lingüistas siempre quisieron que fuera, aunque muchos escritores nunca quisieron creer: una mera transcripción del lenguaje hablado (Barthes, 1986, p. 159). Sin embargo, el lenguaje hablado pertenece al registro oral, mucho menos reglado que el escrito, de modo que hablar no es lo mismo que escribir y viceversa; lo sabe bien quién debe transcribir una entrevista y crear con ella un artículo conciso, tanto como lo sabe quién quiere leer con naturalidad y habló con ellas como si fuera hablado” el discurso escrito, así como lo he sabido yo después de aquellas conferencias en las que obtuve un Registro textual, pero no en registro de escrito. La voz y la letra, lo oral y lo textual, pertenece naturalezas bien distintas (Barthes, 2002, p. 98; Lechte, 2010, p. 191).

Dice Enzensberger que casi todas las personas habla mejor de lo que escriben, debido a que el habla se aprende a edad bastante temprana y bajo unas circunstancias más bien favorables y la escritura, por el contrario, constituye la socialización autoritaria a través de la escuela fundada en dos más y formalización (Enzensberger, 1972, pp. 63–64). Cómo

---

<sup>1</sup> *GKA Arts*, Complutense de Madrid, 2019; *Jornadas de investigación*, Bilbao Arte, 2019; *Ikasgelatik espazio publikora*, Artium Gasteiz, 2020.

recuerda la antropología, el lenguaje oral está ligado a la audición y la locución (a la cara) como libros mientras que el lenguaje escrito está ligado al grafismo y al gesto (a la mano), por lo que concierne a dos zonas diferentes del cortex. Pensar que la escritura no es más que un procedimiento para fijar el lenguaje articulado, que constituye nada más que el instrumento de un instrumento, es un mito de la escritura, una ilusión alfa bética (Barthes, 2002, p. 98). En esta situación de dictado, podría decirse que nos encontramos en una posición intermedia donde existe la audición y la locución, tanto como la lectura, y en cierta medida un gesto, que no es el de la mano sino el de la máquina. Escribir con la voz es, por tanto, una práctica híbrida que, si viene gente a un texto, permanece ligeramente desplazada de lo que podría considerarse escritura.

## Hablar con las máquinas

Lo que se hace en el dictado por voz es funda mentalmente hablarle a una máquina. Técnicamente, la operación consiste en una máquina que capta el sonido de la voz para que luego la inteligencia artificial la reconozca como lenguaje y adjudica esos sonidos, con mayor o menor acierto, palabras de un diccionario. Esta operación, al contrario de lo que pudiera parecer, me impone el mismo objetivo que cuando se habla a otro individuo: la finalidad última del acto comunicativo es que el receptor me comprenda. Para que esto ocurra efectivamente deben cumplirse las cualidades de una expresión oral adecuada, a saber: volumen oportuno, ritmo fluido (no demasiado rápido ni excesivamente lento) y dicción clara. La atención debe fijarse en que la máquina comprenda lo que se le está diciendo. Si bien los esfuerzos pueden paralizarse hasta adquirir una locución propia del mejor comentarista, de poco sirve la entonación, puesto que como en las herramientas a disposición, el dictado por voz no reconoce dónde colocar los signos de puntuación. Para ello debe estudiarse el lenguaje específico de los comandos que permite introducir los signos paralingüísticos relativos al ritmo y la entonación. Los comandos “punto sierra”, “coma” como “:” y “punto y coma”, así como “interrogación” y “!”, deben ser dictados a continuación de la palabra que los preceda y haciendo una pausa después, que permita que sean colocados adecuadamente. Cuando se escribe mediante el dictado por voz, por tanto, la paralingüística toma cuerpo, se hace voz, y se dice lomudo.

En consecuencia, hablar con las máquinas no es exactamente igual hablar y constituye una adaptación de lenguaje propio o natural en el que las palabras tienen espaciarse y la adición a exagerar se, al tiempo que se producen cortes para introducir los comandos del paralingüaje. Con algo de práctica, el habla el hablante/escritor puede adoptar un registro—muy similar al del profesor que dicta para copiar—que permite a la inteligencia artificial comprender cada palabra—como lo haría un niño de primaria—. Este símil no es en vano,

puesto que este medio de escritura guarda muchas similitudes con el dictado tradicional, una actividad lingüística a medio camino entre lo oral y lo escrito que resulta totalmente distinta de cortar o escribir (Fabbretti and Teberosky, 1993). Gracias a la “familiarización”, la herramienta de dictado por voz genera un diccionario propio y por tanto el aprendizaje es mutua 2 puntos al tiempo que yo aprendo a dictar de esta manera y comprendo cuál es el ritmo de cua do, la máquina va entendiendo me cada vez mejor, puesto que amplía su diccionario, reconoce mi voz y escribe las palabras que utilizo con mayor frecuencia.

## Errores de transcripción

Para que este artículo tuviera sentido, para que supusiera realmente la puesta en práctica de un tipo de escritura hecha con la boca, me autoimpuesto la norma de no corregir lo que se va escribiendo<sup>2</sup>. Esto, que ya se ha manifestado a lo largo de todo el texto, contradice un rasgo primordial de la escritura académica 2 puntos el copy editing. La revisión y corrección de un texto escrito no es lineal, va debajo arriba va de abajo a arriba, delante atrás, vuelve sobre el escrito y transforma los contenidos. Este procedimiento resulta imposible en el habla y en consecuencia, también en el dictado. Cómo apunta Barthes, “la palabra (hablada) es irreversible, es decir: no se puede corregir una palabra sin decir explícitamente que se va a corregir” (Barthes, 1986, p. 315), lo que nos lleva expresiones como “digo...”, “mejor dicho puntos suspensivos”, “creo que me he expresado mal... o a la repetición y el balbuceo. En el presente ejercicio de escritura, la corrección se realiza mediante el habla y por tanto implica siempre suma, nunca borrar. Al contrario que muchos otros textos que han sido presumiblemente escritos con la voz, este no se somete al escrutinio final final del teclado punto bobo párrafo

En consecuencia, uno de los errores comunes de este texto es el error del habla (*lapsus linguae*), que se corrige a través del balbuceo y la repetición —cuando me equivoco atasco al hablar, trato de repetir, con más claridad, lo que acabo de decir—. Pero no es la única clase de error en esta transcripción, sino que aparecen también errores propios de la escritura (*lapsus calami*). La intromisión de letras indebidas es un error común de la mecanografía, así como la transposición de las letras que conforman una palabra. El *lapsus calami* produce principalmente dos tipos de palabra: palabras a semic as, fruto de una letra inoportuna que desfigura la palabra de modo que no encaja con ninguna figura semántica, y palabras que constituyen grafías identificables en el éxito, si bien no eran las que se querían escribir.

---

<sup>2</sup> A excepción de las notas al pie y los nombres de los autores, ya que estos, debido a mi mala pronunciación o al poco acervo cultural de mi teclado, pueden fácilmente confundirse con otras palabras.

En el primer tipo es posible hacer asociaciones pero no se descifra, mientras que en el segundo Tito la frase sigue teniendo sentido, por excéntrico que sea, y ello da lugar a juegos de palabras, anagramas, polisemias y, en definitiva, deslizamientos el interior del código (Barthes, 1986, p. 329). En esta práctica de dictado por voz los lapsus son únicamente del segundo tipo, puesto que los errores se deben a fallas en la escucha o la comprensión de lo que digo por parte de la inteligencia artificial, que sustituye unas palabras por otras que se parecen fonéticamente. Las palabras de las que dispone el dictado por voz son siempre significativas en la medida en que proceden de un diccionario o Corpus. De cualquier modo, cabe destacar que no se trataría estrictamente de lapsus calami, puesto que, recordémoslo, aquí nadie teclea, si bien existe un procedimiento de construcción de la escritura.

A los errores comunes como la misión de palabras, la inserción de incorrecciones no emitidas o la sustitución de unas palabras por otras, hay que sumar la dificultad del sistema para discriminar el ruido ambiental y la ignorancia ante términos de construcciones no contenidos en sus gramáticas (Huang, 2007). El repositorio disponible para la escritura por voz es un corpus, pero, por algún motivo, en ocasiones ocurren faltas de ortografía. Sucede especialmente con las tildes diacríticas, aquellas que permiten distinguir palabras que se escriben igual pero se pronuncian y significan diferente. Es lógico pensar entonces que el lee todo por voz no comprende el texto en su globalidad y que va resolviendo cada palabra en la medida en que las escucha. De lo que si dispone es de un criterio de probabilidad, similar al del texto predictivo, en base al cual es mas probable que el orador diga “hoy” después de la palabra “lloverá” coma el lugar de “lloverá cemento”, y por ello existe una predisposición por ciertas combinaciones. En cualquier casco, las palabras escritas incorrectamente, o aquellas correctamente recogidas pero mal entendidas en el contexto, son subrayadas en rojo: una llamada de atención que indica claramente la demanda de una revisión up “manual” por parte del orador guión orador/escritor y que en este momento no es posible realizar mediante comandos de voz (por comandos cómo “seleccionar palabra más borrar” o similares).

Por último, existen ciertas clases conflictivas cuando se trata de dictar por voz debido al requisito de dictar los signos de puntuación (esto se ha podido resolver en algunos idiomas como el inglés y el japonés, que ya disponen de puntuación automática). Se trata de aquellas frases en las que el léxico del lenguaje y el del paralinguaje pueden confundirse coma tales como “, lo que coma, me sienta mal” o “hay: importantes dos puntos este y el otro”.

En definitiva, todo este conjunto de errores hace del escrito resultante una masa textual torpe en comparación con el texto tradicional escrito con las manos, en el que, si bien puede haber erratas, estás suelen pertenecer a versiones iniciales de los textos, previos a su corrección: un material bruto que rara vez llega mucho más lejos de las manos de quién lo

escribió y que aún con más improbabilidad veremos publicado. El texto presente continuo deslices y desvíos, tantos como el audio que lo acompañe. De este aprendizaje mutuo entre el usuario y la máquina, el habla del orador se convierte en maquina<sup>3</sup> por el obligado uso de tantos signos más allá del registro oral, y la escritura de la máquina se torna peculiar, inexacta, casi diría “personal” como repleta de deslices propios del descuido o el desaliño. Así, para exponer el funcionamiento de este sistema de escritura y analizar sus particularidades, acabo haciendo dos textos desplazados punto com un audio lleno de metalenguaje y un texto lleno de ratas.

## Aceptar el error

La práctica de la escritura, extremadamente formalizada, está sujeta a un alto grado de especialización y llena de tabús: las fallas involuntarias propias del habla como las pausas, interrupciones, equivocaciones, repeticiones y anacolutos no son tan duramente señaladas como los errores ortográficos. La mala ortografía, que es completamente indiferente para la comunicación, está castigada con el desplazamiento social del escritor (Enzensberger, 1972, p. 62) y por ende un artículo académico o cualquier otra publicación que se precie con errores de lenguaje es una desfachatez. No obstante como desde una práctica plástica que toma el texto como materia, pueden entenderse escrituras no normativas; las grafías asemejas de Mirtha Dermisache y León Ferrari o las últimas cartas de Réquichot, y, de forma menos casi gráfica y más tipográfica, algunas propuestas de Marco Giovenale y Geof Huth. Todas ellas son consideradas ilegibles pero con indiscutible textualidad, y como este ejercicio de dictado, proponen derogar las funciones puramente prácticas de la escritura, funciones de contabilidad, comunicación y registro, a favor del “simbolismo que mueve el signo escrito” (Barthes, 2002, p. 93). La puesta en marcha de esta máquina para la escritura Herrada me empuja a aceptar los errores incluso cuando eso puede suponer el rechazo de este artículo, alejando la evidente incorrección de mi escritura. No obstante, este texto no podría existir de otra manera.

Posiblemente sea un atrevimiento afirmar que la escritura mediante dictado por voz es una herramienta de escritura creativa solo por sus errores. No es solo el error lo que expande esta escritura, sino que ponerla en práctica si tú a la escritora/oradora en un lugar donde dialoga consigo misma (porque escribe) y dialoga con la máquina (porque habla) como y esta nueva posición genera nuevas formas de denuncia. Ese lugar es diferente al que se tiene cuando

---

<sup>3</sup> «Con todas esas gramáticas estamos forzando a que la gente hable el lenguaje de la máquina y no al revés», denuncia Xuedong Huang, director de tecnologías del habla de Microsoft (Huang, 2007).

se habla en una conferencia y es diferente camina lugar que se ocupa cuando se escribe. El dictado por voz está más cerca de la situación de dictado tradicional, donde el dictante de Leganés el escribiente solo los aspectos ejecutivos de la escritura y asume todas las decisiones relacionadas con la tarea de composición del texto (Fabbretti and Teberosky, 1993). Si bien el escribiente no es un interlocutor, y por tanto el ejercicio de dictado no es, de ninguna manera, una conversación, el dictante está pendiente de como el escribiente realiza su tarea. Dicho estado de alerta, en consecuencia, sí que me hace sentir en diálogo con la máquina., nunca antes en una conferencia me hubiera preocupado tanto porque se entendieran mis palabras y nunca antes en un texto había tenido tan presente la puntuación. De modo que me es imposible escribir desde mi individualidad: la voz que adoptó es otra, y constantemente fijo la mirada en lo que el dictado por voz va recogiendo. Si cuando escribo puedo afirmar que escribo sola, siempre para otros que leen, aquí el papel escrito ahora se divide. En esta situación, donde la máquina impone sus res quesitos y realiza sus propias interpretaciones, tal vez no sea tan disparatado decir que este artículo es una forma de coautoria.

## Conclusiones

A modo de conclusión, queda patente que escribir un texto con esta herramienta es aún el ejercicio experimental cuando no se corrige. El habla matinal necesaria para el dictado por voz irá desapareciendo por latamente: los avances de la inteligencia artificial van a poder resolver la torpeza de la locución a través de las mejoras en puntuación automática o aplicando aprendizaje neuronal. Sin embargo, es interesante explorar esta cualidad ahora, no tanto como un paso intermedio hacia la especialización, sino como una posibilidad en sí misma, de gran utilidad ya para la transcripción, pero sobre todo de gran interés para prácticas más al límite de la escritura.

Hay una brecha enorme entre lo que se dice y lo que se escribe: lo reflejaban aquellas ponencias cuando el registro de la oralidad no proporcionaba un texto redactado y ocurre, de otra manera, también en este dictado, donde esporádicamente se fijan cosas que no he dicho o desaparecen las que sí dicte. Cuando se trata de hacer una transcripción del oral (registro oral), la escritura por voz es una técnica donde la relación de fijación de lenguaje pudiera parecer más directa, pero también anula una dimensión que Barthes advierte en Variaciones sobre la escritura: “hay una gran distancia entre mi cabeza y mi mano, y esa distancia la aprovecho para decir una cosa diferente de la que he encontrado al principio” (Barthes, 2002, p. 166). En otra medida en el dictado de este artículo, que en todo momento señala la existencia de un texto anterior y por tanto la transcripción va de lo escrito al oral y nuevamente a lo escrito (registro escrito) como el texto sufre variaciones y se genera una nueva manera de denunciar que ya es señalado.

Con suerte, y al igual que proponía en las conferencias en las que use el dictado por voz, el oyente/lector se enfrentará a dos textualidades: la audible (en el audio podcast) y la legible (en estas páginas) punto el sentido de todo lo que quiero decir se encontrará en la un lugar intermedio (Higgins and Higgins, 2001), no en los deslizamientos de lo que se le ni en la familiaridad de lo que se escucha, sino en aquello que el lector-oyente construye combinando ambas. Es la misma brecha la que existe entre lo que se piensa y lo que se dice, o entre lo que se quiere decir y lo que finalmente se dice. Cuando nos comunicamos, influye siempre lo que sabemos, lo que entendemos, creemos haber entendido y lo que vamos recalculando: el intercambio nunca es absoluto, también es una negociación.

## Bibliografía

ARASA, S., 2001. Entrevista a Xuedong Huang, Director mundial de tecnologías del habla de Microsoft: 'La próxima revolución será la de la voz.' *Quark. Ciencia, Medicina, Comunicación y Cultura* 21, pp.103–108. <https://raco.cat/index.php/Quark/article/view/54863/66193> (Accessed: 5 March 2022).

Barthes, R., 1986. *Lo obvio y lo obtuso*. Barcelona: Paidós.

Barthes, R., 2002. *Variaciones sobre la escritura*. Barcelona: Paidós.

Enzensberger, H.M., 1972. *Elementos para una teoría de los medios de comunicación*. Barcelona: Anagrama.

Fabbretti, D. and Teberosky, A., 1993. Escribir en voz alta. *Cuadernos de pedagogía*, No 216, pp. 54–56.

Higgins, D. and Higgins, H., 2001. Intermedia. *Leonardo*, February, pp. 49–54.

Lechte, J., 2010. *50 pensadores contemporáneos esenciales. Del estructuralismo al porthumanismo*. Madrid: Ediciones cátedra.

## **Miriam Inza**

Artista e investigadora. Egresada en Creación y Diseño (2017) y Máster en Investigación y Creación en Arte (2018) por la UPV/EHU. He participado en diversos congresos (GKA Arts, Madrid, 2019; CIVARTES, Jaén, 2020; Del aula al espacio público, Vitoria, 2020; RADAR, Tenerife, 2022), expuesto mi trabajo plástico de forma individual (sala J. de Lizarazu, Gipuzkoa, 2019; Sala J. Valera, Córdoba, 2020; Ext. version exposición itinerante por 11 localidades de la CAV) y colectiva (Bizkaia Aretoa, 2017; Ateneo Santa Cecilia, 2018; Museo de reproducciones de Bilbao, 2020; Vicerrectorado Unizar, 2021; Centro Cultural Marcos Valcárcel, 2022) y he recibido la beca de Creación-Producción en Artes plásticas y Visuales del Gobierno Vasco en el año 2019. Actualmente desarrollo mi tesis doctoral en el Depto. de Escultura y Arte y Tecnología de la Facultad de Bellas Artes (UPV/EHU) con la beca PIF (2018). Formo parte del Grupo de Investigación Consolidado IT1460-22 GizAArtea: Diálogos Críticos arte/sociedad.

Artist and researcher. Graduated in Creation and Design (2017) and Master's degree in Research and Creation in Art (2018) at the UPV/EHU. I have participated in various congresses (GKA Arts, Madrid, 2019; CIVARTES, Jaén, 2020; Del aula al espacio público, Vitoria, 2020; RADAR, Tenerife, 2022), exhibited my work both individually (J. de Lizarazu room, Gipuzkoa, 2019; Sala J. Valera, Córdoba, 2020; Ext. version itinerant exhibition through 11 locations in the CAV) and collectively (Bizkaia Aretoa, 2017; Ateneo Santa Cecilia, 2018; Museum of reproductions of Bilbao, 2020; Vice-rectorate of the Unizar, 2021; Centro Cultural Marcos Valcárcel, 2022) and received the grant for Creation-Production in Plastic and Visual Arts from the Basque Government in 2019. I am currently developing my PhD at the Dept. of Sculpture and Art and Technology of the Faculty of Fine Arts (UPV/EHU) with the PIF grant (2018). I am part of the Consolidated Research Group IT1460-22 GizAArtea: Critical Dialogues art/society.