



Deslizamientos progresivos del sentido. La cuestión de la congruencia audio-visual  
Udo Jacobsen C.  
Arkadin (N.º 14), e063, 2025. ISSN 2525 085X  
<https://doi.org/10.24215/2525085Xe063>  
<https://papelcosido.fba.unlp.edu.ar/ojs/index.php/arkadin>  
Facultad de Artes. Universidad Nacional de La Plata  
La Plata. Buenos Aires. Argentina

# DESLIZAMIENTOS PROGRESIVOS DEL SENTIDO

La cuestión de la congruencia audio-visual

**Progressive Slippages of Meaning**

The Issue of Audiovisual Congruence

UDO JACOBSEN C. | [udo.jacobsen@uv.cl](mailto:udo.jacobsen@uv.cl)

Escuela de Cine, Universidad de Valparaíso, Chile

Recibido: 27/05/2025 | Aceptado: 28/07/2025

## RESUMEN

Las próximas líneas intentan delimitar con mayor precisión un concepto de análisis audio-visual propuesto por Michel Fano, referido a la congruencia entre la imagen y el sonido. Principalmente referido al sincronismo, pero también relativo a la iconicidad sonora, el concepto es verificable ampliamente en diversas películas, pero presenta algunas particularidades que pueden ser reconocidas como códigos genéricos o estímulos autorales. En este sentido, la cuestión de la congruencia audio-visual, se presenta como un componente diferencial que puede formar parte de un estudio más extenso sobre una posible retórica audio-visual.

## PALABRAS CLAVE

cine; imágenes; sonido; sincronización; audiovisual

## ABSTRACT

The following lines attempt to define more precisely a concept of audio-visual analysis proposed by Michel Fano, which refers to the congruence between image and sound. Primarily related to synchronization, but also concerning sonic iconicity, the concept is widely verifiable in various films. However, it presents certain particularities that can be recognized as generic codes or authorial stylistic features. In this sense, the issue of audio-visual congruence is presented as a differential component that can be part of a more extensive study on a possible audio-visual rhetoric.

## KEYWORDS

cinema; images; sound; synchronization; audiovisual



Esta obra está bajo una Licencia  
Creative Commons Atribución-  
NoComercial-CompartirIgual 4.0  
Internacional

En 1975 Michel Fano presentó, en los coloquios dedicados a Alain Robbe-Grillet en Cerisy, una teoría de las relaciones entre imágenes y sonidos que presenta ciertas particularidades. En primer lugar, considera que la materia sonora sólo es divisible operativamente en el ámbito de la realización, pero que a la hora de ser percibida en una película debe considerarse como un solo continuum sonoro. Las diferencias tradicionales, es decir, los enunciados fonéticos, los sonidos análogos y los sonidos musicales, quedarían de este modo subsumidos en una sola banda determinada por su espesor semántico. Visto así, no habría sonidos distinguibles por su materialidad específica sino por su grado de significación y, por lo tanto, quedarían sujetos a un proceso de reconocimiento por parte del espectador que debiese identificar tal o cual sonido según su iconicidad. No obstante, es importante considerar que la función referencial de un sonido depende en buena medida del conocimiento y experiencia del espectador, en especial cuando no se encuentra asociado a una imagen concreta que, por otro lado, presenta normalmente una relación más directa con su referente. El sonido se presenta, entonces, como algo aproximado, como un punto de referencia para una interpretación más o menos certera de un referente ausente. Algo que Fano denomina plasticidad semántica.

Es justamente en este contexto que nos interesa abordar la cuestión de la iconicidad y su correspondencia con la imagen con que se asocia. Para Fano, las funciones del sonido en el cine se asocian a tres ejes que permiten comprender cómo se entrelaza el dato sonoro con la configuración visual que lo acoge. En primer lugar, está el sincronismo o no-sincronismo del sonido con la imagen. Un sonido puede converger al mismo tiempo que la imagen y asociarse directamente al referente expuesto por esta última. Vemos a alguien hablar y escuchamos su voz, vemos una botella quebrarse en el piso y escuchamos el ruido de los cristales estallando. En el caso de los sonidos musicales, la convención indica una suerte de correspondencia entre algunos rasgos de la composición, como la armonía o el tempo, y la acción representada. En un extremo, podríamos considerar el efecto de *mickeymousing* tan característico de algunos dibujos animados o de ciertas películas cómicas. Lo mismo sucede en sentido contrario. Escuchamos una voz que habla, pero no vemos a quien emite las palabras o escuchamos el ruido de la rotura de vidrios, pero no vemos la botella estrellándose en el suelo.

La segunda función es aquella a la que nos hemos referido al inicio. Se trata de una función referencial que implica que el sonido que escuchamos podemos identificarlo, al menos, dentro de un espectro más o menos acotado. Podríamos escuchar el sonido de la botella rompiéndose, pero no podemos asegurar, si no hay sincronismo o no ha habido una imagen previa que permita suponer una asociación precisa, que se trata, en efecto, de una botella o de algún otro objeto con características similares, como un vaso o una jarra. Por otro lado, algunos sonidos no son identificables debido a que no existe una experiencia auditiva de su fuente de emisión. Un caso interesante es el de *El territorio de los otros* (Bel, Vienne & Fano, 1970), donde escuchamos un sonido curioso, algo sordo y seco, al que no logramos atribuirle una fuente de emisión precisa [Figura 1]. Más adelante lo escuchamos nuevamente, pero esta vez asociado a su emisor, un pájaro aparentemente típico de ciertos paisajes agrestes y poco accesibles de Europa. Esto implica que, al no tratarse del sonido más genérico característico de un pájaro, no es posible reconocer su origen sino para aquellos pocos que en ese momento habitan el mismo territorio donde habita el ave.

Esto nos lleva al tercer eje, el de la congruencia. Retomando el ejemplo anterior, el canto del pájaro no puede ser atribuido a las imágenes que vemos cuando lo escuchamos por primera vez —imágenes de árboles donde no vemos la presencia de ningún pájaro—. Sin embargo, se vuelve congruente en el momento en que se produce el sincronismo entre el sonido que escuchamos y el pájaro que vemos —y que desconocíamos—, integrándose al acervo del espectador y volviéndose icónico por efecto de este nuevo conocimiento. Por el contrario, la no congruencia dependería mayormente de nuestro



Figura 1. *El territorio de los otros* (1970) Dirigida por François Bel, Michel Fano y Gérard Vienne



Figura 2. *Duck Amuck* (1953), Dirigida por Chuck Jones

conocimiento de uno de los dos referentes: el sonoro o el visual. Si vemos la imagen de algo reconocible, esperamos lógicamente el sonido que debiese producir, al menos de manera genética, de modo que no crearía otra cosa que una expectativa de confirmación del hecho sonoro correspondiente. Pero, si el sonido que escuchamos en ese momento, no se corresponde con la imagen que observamos, independientemente de su iconicidad, será no-congruente. Esta es una cuestión que reviste gran interés para comprender las posibilidades creativas de una retórica cinematográfica. Tomemos algún ejemplo.

Un caso evidente es el del cortometraje de animación de Chuck Jones, *Esto está de locos* (1953). En él, el pato Lucas lucha por sus condiciones de existencia contra un lápiz que redibuja su entorno y a él mismo constantemente [Figura 2]. Pero no se trata solo de una intervención en el ámbito de lo visual solamente, sino que, en un momento, llega a alterar su voz reemplazándola por otros sonidos como el canto de un gallo, un rebuzno o un claxon. El efecto que se busca es expresivo y persigue una elaboración de gag propiamente audio-visual. Conocemos la voz del personaje y reconocemos que los sonidos emitidos, a pesar de su sincronismo y referencialidad, no resultan congruentes. En este caso, podríamos plantear que se trata de una reflexión metalingüística de la llamada cocinería cinematográfica, que alude a los efectos de posproducción sonora. El pato emite una muestra de catálogo de archivos sonoros típicos, no para decir algo sobre las características del personaje sino para referirse a la modalidad de sincronización de sonidos en la etapa de posproducción. Habla, en el fondo, de la «arbitrariedad» del procedimiento desplegando sus posibilidades.

Este es el tema central de una película como *Berberian Sound Studio* (Strickland, 2012). En ella seguimos el proceso de sonorización de una hipotética película de horror italiana, aparentemente muy cruda. Nos situamos en el punto de vista de los personajes que observan la película durante el proceso, actores y actrices, especialistas en foley, productor y, especialmente, el ingeniero de sonido. Las imágenes de la película nunca nos son mostradas, con la salvedad de la secuencia de créditos inicial que alude a los códigos típicos de este tipo de películas. A lo largo del relato observamos cómo se manipulan y refinan los sonidos que acompañarán las imágenes, poniendo énfasis en los recursos, mayoritariamente vegetales, que se utilizan para lograr los sonidos de cuchilladas y torturas varias [Figura 3]. Lo relevante aquí es comprender que la referencialidad, en tanto función, no requiere de la veracidad en la producción del sonido. Acuchillar un repollo será tan icónico como acuchillar un cuerpo. Nos adentramos, en el fondo, en el campo de la convencionalidad, donde la iconidad depende más del horizonte de experiencias del espectador que de la realidad. Es algo que podemos comprobar en el ámbito de los géneros. Cada uno tiene soluciones propias para la producción de sonidos equivalentes en situaciones similares. Más allá de las cuestiones puramente técnicas, tenemos claro que un disparo no suena del mismo modo en un western clásico que en un espagueti western.



Figura 3. *Berberian Sound Studio* (2012), dirigida por Peter Strickland

La iconicidad depende, por lo tanto, más de la convención, por ende de la repetición de ciertas fórmulas, que de la veracidad del sonido. Es algo que responde a la definición aristotélica de la verosimilitud: las reglas del género se deben corresponder con la opinión del público. Un sonido será, entonces, congruente si es aceptable como resultado de la emisión de un determinado referente en la imagen, independientemente de su veracidad. Se trata de una construcción y, como tal, puede cambiar con el tiempo. Todo dependerá de esa relación, algo elástica, entre el espectador y la película.

Pero si esta relación está construida, es posible «proponer» el sonido de la imagen. Es algo que podemos verificar fácilmente en el ámbito del doblaje. Es conocida, por ejemplo, la política de Disney de buscar los timbres más similares posibles a los de los actores originales para sus castings de doblaje en diversos idiomas, buscando minimizar al máximo las diferencias sonoras entre las distintas versiones. Pero esto no es, realmente, el caso de la mayoría de las películas y, para muchos espectadores, puede llegar a resultar chocante escuchar, por ejemplo, la versión original de una serie cuando ha estado acostumbrado al doblaje en su idioma. Pero si nos situamos en el caso de personajes fantasiosos, como un monstruo sobrenatural o un extraterrestre, las posibilidades parecen más amplias, no obstante la elección de los timbres responde habitualmente a parámetros sonoros impuestos por la verosimilitud genérica. Incluso, las voces tienden a someterse a un proceso de intervención de diversos niveles con el fin de deshumanizarlas.

Volvamos ahora a la cuestión central. La congruencia entre imagen y sonido no es algo dado previamente sino que se construye en el propio texto filmico. En aquellos casos en que nos enfrentamos a seres u objetos desconocidos, como hemos observado, la construcción es naturalmente aceptable dentro del marco genérico correspondiente, pero ¿qué sucede cuando la referencialidad del sonido como de la imagen son reconocibles independientemente, pero no resultan congruentes? Tenemos otro caso de *El territorio de los otros*. En una escena aparecen unos flamencos levantando el vuelo en bandada y escuchamos el sonido del despegue de un avión, o en otra escena vemos a unos conejos corriendo de un lado a otro por un prado y escuchamos el sonido de un tren en marcha. En ambos casos nos encontramos claramente frente a la producción de una comparación. Podríamos decir que es un equivalente audiovisual del montaje intelectual de Eisenstein, tal como lo vemos, por ejemplo, al final de *La huelga* (1925), o una muestra precisa de su propuesta de un contrapunto orquestal para el cine sonoro. No deja de ser interesante, respecto de esto, lo que anota Michel Chion cuando dice que «el contrapunto audiovisual no se advierte salvo si opone sonido e imagen sobre un punto preciso, no de naturaleza, sino de significación» (Chion, 1993, p. 36). Como sea, se trata de un recurso retórico que apunta a la complementariedad entre el sonido y la imagen en la creación del discurso, aunque nos pueda resultar un tanto inocente o directo el ejemplo.

Donde nos encontramos con elaboraciones más sofisticadas es en la larga colaboración entre Michel Fano y Alain Robbe-Grillet. Películas como *El hombre que miente* (1968), *El Edén y después* (1970) o *Deslizamientos progresivos del placer* (1974), son claros ejemplos de la voluntad de tratar los sonidos y las imágenes bajo la perspectiva de una «partitura» audiovisual. Las diversas estrategias que asumen los realizadores al «jugar» con los sonidos apuntan a generar un universo sonoro ambiguo y cambiante. Un sonido puede aparecer asincrónicamente, eventualmente como un dato del fuera de campo, generando una expectativa que luego se ve frustrada al no aparecer la fuente de emisión sonora. En otros casos, escuchamos los sonidos como anticipación de una acción, evitando siempre el sincronismo con la fuente. Ocasionalmente, un golpe sonoro coincide sincrónicamente con un gesto del personaje, que pareciese escucharlo, aunque claramente no forma parte, en principio, del universo diegético. Estas y otras variables tienden a la construcción de un mundo más formal que narrativo. Se condice, por lo tanto, con la concepción estética de Robbe-Grillet, que considera sus personajes no



Figura 4. *El Edén y después* (1970), dirigida por Michel Fano y Alain Robbe-Grillet

como proyecciones de personas reales sino, justamente, como construcciones formales. Los sonidos participarían de este mismo espíritu y habrían de ser consideradas más desde una perspectiva musical que desde el punto de vista de la restitución de un mundo habitado por personajes que viven una historia.

Uno de los sonidos recurrentes en la filmografía de Robbe-Grillet, incluso desde *El año pasado en Marienbad* (Resnais, 1963), es el de cristales rotos. Una escena frecuente en sus películas es el momento en que se quiebra una copa o una botella. De manera casi canónica, escuchamos primero la rotura del cristal y luego vemos el objeto roto. Es algo que podemos observar de película en película, como si formara parte de un catálogo permanente de situaciones (algo que tiende a suceder también, por ejemplo, con los nombres de los personajes) en las que el cineasta dispone las piezas de un juego en cada partida. Más allá de la correspondencia con la realidad, se produce una suerte de codificación del acto que adquiere invariablemente el valor de interrupción de la acción —la acción se detiene momentáneamente mientras alguien recoge los pedazos y es observado por los personajes circundantes—. El sonido de los cristales rotos adquirirá de esta manera una significación que será aludida incluso cuando no veamos el acontecimiento en la imagen. Así, la congruencia entre la acción y el sonido servirá para generar la función sonora que dispondrá el autor en diversos momentos como punto de marcación. Se trata, en el fondo, de la construcción de una figura intra e intertextual que se dispone como código reconocible para el espectador familiarizado.

La congruencia sonora puede ser observada desde el punto de vista del calce icónico con el referente como una coincidencia entre la imagen —fuente de emisión del sonido— y el sonido que escuchamos —se le atribuye como sonido «natural»—, pero es también posible, y sobre todo, considerarlo como un efecto formal que juega estructuralmente en la construcción del discurso filmico y que puede llegar a constituirse en un verdadero estilema de autor —recordemos de paso, el sonido del viento en las películas de Fellini—. En estas líneas hemos podido avizorar apenas parte de lo que podría constituir una teoría más amplia de retórica cinematográfica que incorpore tanto la imagen como el sonido, uno de esos grandes relegados de la teoría cinematográfica.

## REFERENCIAS

- François, B., Fano, M. y Lecompte, J. (1970). *Le territoire des autres* [El territorio de los otros][Película].
- Chion, M. (1993). *La audiovisión. Introducción a un análisis conjunto de la imagen y el sonido*. Paidós.
- Fano, M. (1976). *Musique et film. Robbe- Grillet: Colloque de Cerisy: Analyse et Théorie* [Música y cine. Robbe- Grillet: coloquio de Cerisy: Análisis y teoría]. Union Générale d'Editions. (trad. Carlos Núñez).
- Jones, C. (1953). *Duck amuck* [Esto está de locos][Cortometraje].
- Resnais, A. (1963). *L'année dernière à Marienbad* [El año pasado en Marienbad][Película].
- Robbe-Grillet, A. (1968). *L'homme qui ment* [El hombre que miente][Película].
- Robbe-Grillet, A. (1970). *L'Eden et après* [El Edén y después] [Película].
- Robbe-Grillet, A. (1974). *Glissements progressifs du plaisir* [Deslizamientos progresivos del placer] [Película].
- Strickland, P. (2012). *Berberian Sound Studio* [Película].