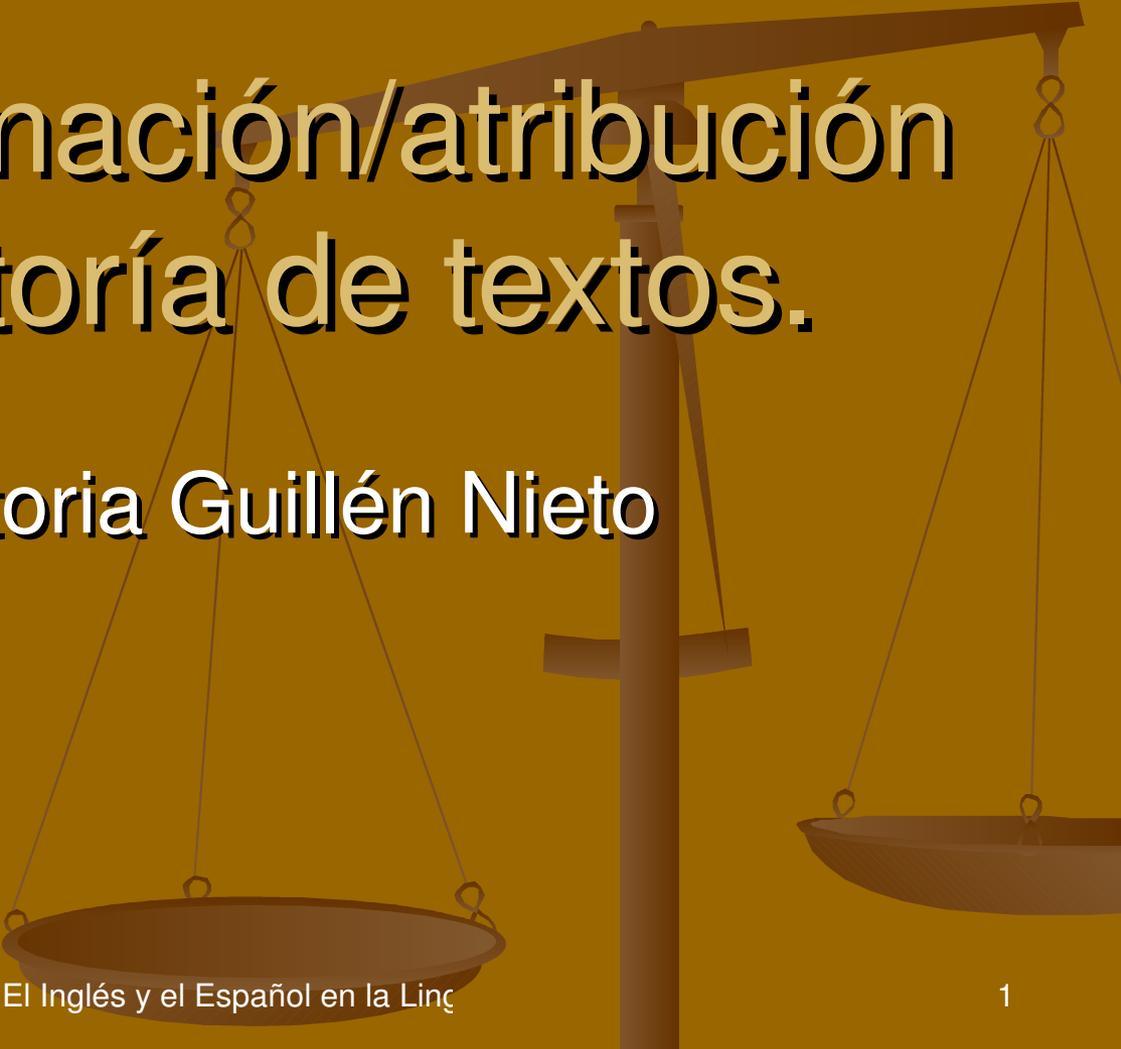


# La determinación/atribución de la autoría de textos.

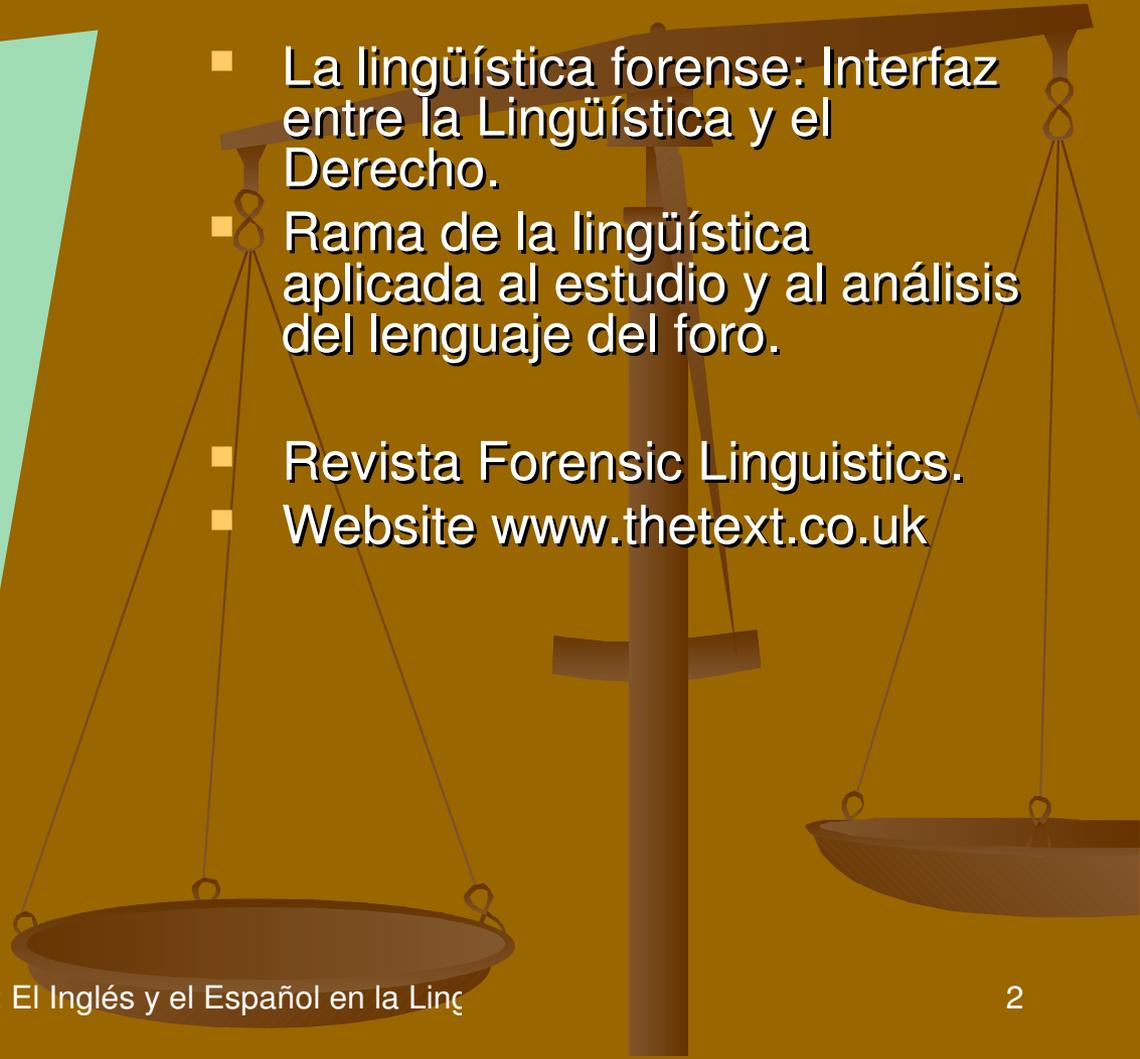


Dra. Victoria Guillén Nieto

# La Lingüística Forense.



- La lingüística forense: Interfaz entre la Lingüística y el Derecho.
- Rama de la lingüística aplicada al estudio y al análisis del lenguaje del foro.
- Revista Forensic Linguistics.
- Website [www.thetext.co.uk](http://www.thetext.co.uk)



# Disciplinas de la Lingüística Forense.



- Determinación/atribución de la autoría.
- Determinación de la modalidad de producción del texto.
- Traducción e interpretación jurídica.

# Disciplinas de la Lingüística Forense.



- Transcripción de declaraciones verbales.
- El lenguaje y el discurso del foro.
- Derechos lingüísticos del individuo procesado.

# Disciplinas de la Lingüística Forense.



- Análisis de las declaraciones hechas por testigos para determinar su índice de credibilidad.
- Fonética Forense.
- Estatus de las pruebas documentales.

# ¿Qué se entiende por la determinación de autoría de textos?

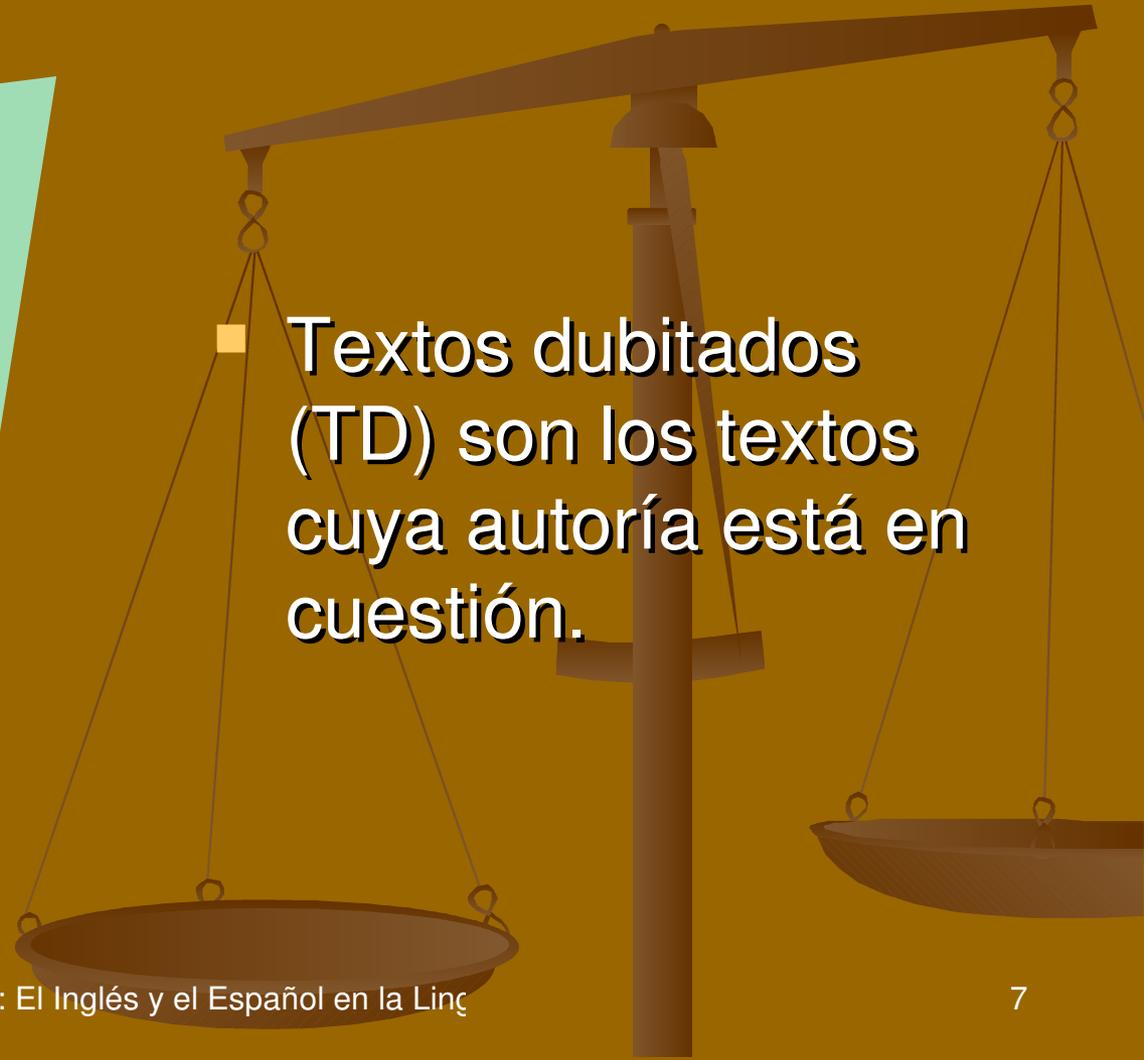


- A partir de un conjunto de individuos sospechosos se busca evidencia lingüística para determinar quién de ellos podría ser el autor de un conjunto de textos dubitados (TD).

# ¿Qué son Textos dubitados (TD)?



- Textos dubitados (TD) son los textos cuya autoría está en cuestión.



# ¿Qué se entiende por la atribución de la autoría de textos?

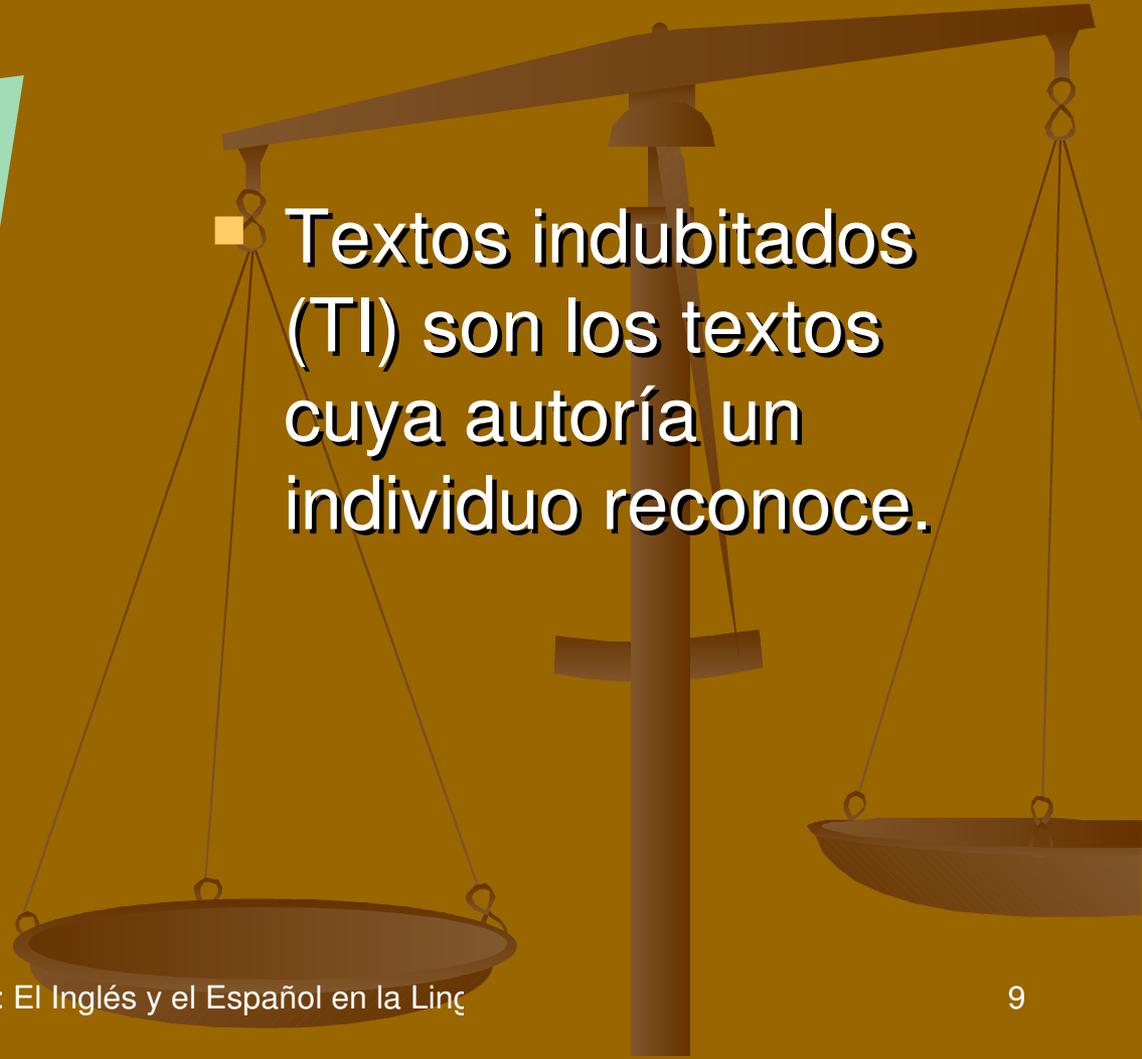


- A partir del análisis de los textos indubitados (TI) escritos por un único individuo sospechoso se busca evidencia lingüística que permita atribuirle o no la autoría de un conjunto de textos dubitados (TD).

# ¿Qué son textos indubitados (TI)?



- Textos indubitados (TI) son los textos cuya autoría un individuo reconoce.



# ¿Qué se entiende por texto forense?



# Antecedentes en los estudios de determinación de autoría: La Biblia.



- S. XVIII. Concepto Cartesiano de la individualidad y de la propiedad intelectual de las ideas.
- 1711. H. B. Witter (monje alemán) plantea la primera controversia en torno a la autoría de la Biblia hebrea, basándose en los distintos nombres que recibe la divinidad en los libros del Pentateuco, tradicionalmente atribuido a Moisés.
- Jean Astruch (médico francés) llega a las mismas conclusiones.
- Siglo XIX, J. G. Eichhorn (Profesor Luterano en la Universidad de Jena) también llegó a las mismas conclusiones.

# Antecedentes en los estudios de determinación de autoría: Las obras literarias de W. Shakespeare.



- 1785. El Reverendo James Wilmot atribuyó la autoría de las obras literarias de W. Shakespeare a R. Bacon. (Trabajo inédito).
- 1805. James Corton Cowell presentó dicha controversia en la Sociedad Filosófica de Ipswich, Inglaterra.

# Estado de la cuestión en los estudios de determinación de autoría: Primeros estudios científicos y estadísticos.



- 1851. Augustus de Morgan, primer Catedrático de Matemáticas en la Universidad College en Londres, escribió una carta al Reverendo Heald (Cambridge) acerca de la autoría de las epístolas atribuidas a San Pablo.

# Estado de la cuestión: Primeros estudios científicos y estadísticos.



- 1851. A Augustus de Morgan se le atribuye:
  - Método estadístico de cálculo del promedio de número de letras por palabra.
  - Identificación del número de letras por palabras como posible marca lingüística de autoría.

# Estado de la cuestión en los estudios de determinación de autoría: Primeros estudios científicos y estadísticos.



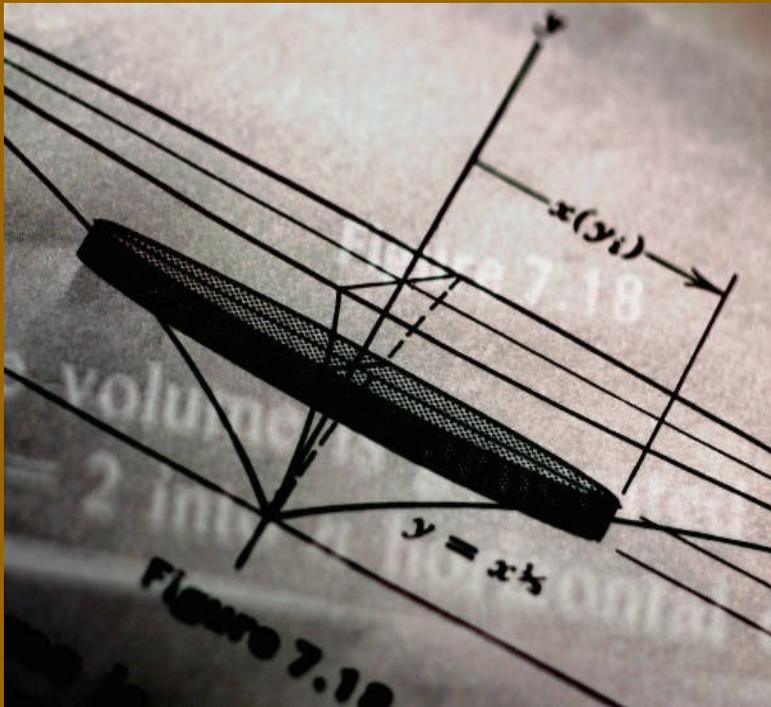
- 1887. T. C. Mendenhall, Catedrático de Física y Mecánica en la Universidad del Estado de Ohio, publicó sus trabajos sobre autoría de texto basándose en el promedio de la longitud de palabra como marca de autoría.

# Estado de la cuestión en los estudios de determinación de autoría: Primeros estudios científicos y estadísticos.



- 1938. Udney Yule, Catedrático de Estadística en la Universidad de Cambridge, identificó el promedio de la longitud de oración como posible marca de autoría.
- Con posterioridad, otros estadísticos como Mosteller & Wallace (1964), y Kenny (1982) también se basaron en el promedio de la longitud de oración como posible marca de autoría.

# Estado de la cuestión en los estudios de determinación de autoría: Primeros estudios científicos y estadísticos.



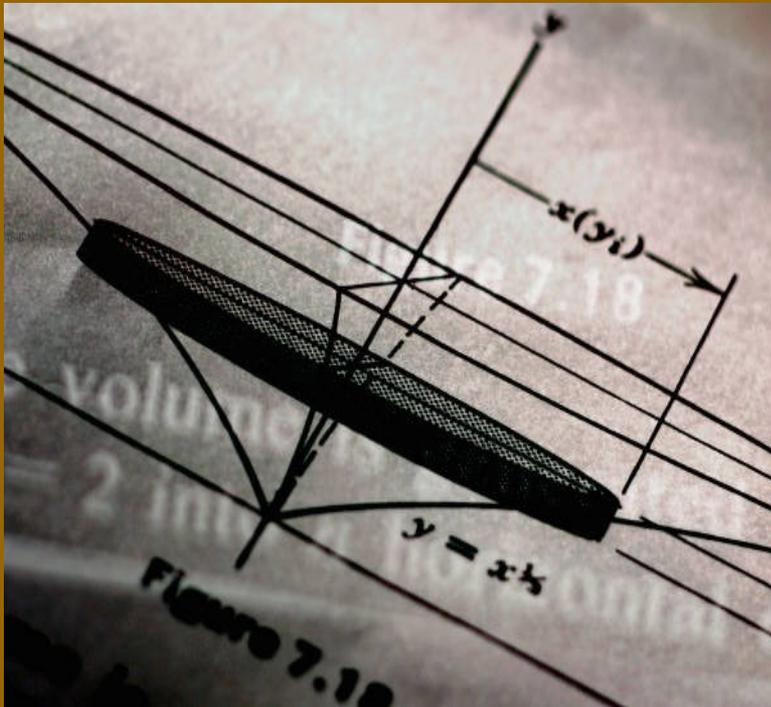
- El promedio de la longitud de oración que caracteriza el estilo de un autor varía dependiendo del género o registro en que éste se comunique por escrito o verbalmente.

# Estado de la cuestión en los estudios de determinación de autoría: La era de la Lingüística Computacional.



- 1980. Esta década marca el nacimiento de la Lingüística Computacional.
- Los estudios llevados a cabo sobre determinación de autoría en el área de la Lingüística computacional se centran en caracterizar el estilo de autor por medio de un algoritmo.
- “(...) each specific text carries the fingerprints of its creator”. (Ephratt).

# Estado de la cuestión en los estudios de determinación de autoría: La era de la Lingüística Computacional.



- Se trabaja con corpora de textos muy amplios. (Texto=17.000-50.000 palabras)
- Se aplican complejas ecuaciones matemáticas.
- Índice de fiabilidad del 87%.

# Estado de la cuestión en los estudios de determinación de autoría: El nacimiento de la Lingüística Forense como disciplina.



- 1968. Jan Svartvik: Caso Timothy John Evans.
- Análisis de cuatro declaraciones del sospechoso a la policía. Descubrió la manipulación textual al identificar dos tipos distintos de registro por medio de un análisis estadístico de naturaleza descriptiva.
- Acuñó el nombre de Lingüística Forense.

# Estado de la cuestión en los estudios de determinación de autoría: El Cusum Test de los hábitos lingüísticos del hablante.



- 1990. Michaelson y Morton. El método de análisis Cusum se basa en la hipótesis de que cada hablante muestra un conjunto distintivo de hábitos lingüísticos.
- Los hábitos lingüísticos del hablante forman patrones textuales identificables mediante procedimiento estadístico.

# Estado de la cuestión en los estudios de determinación/atribución de autoría: El Cusum Test de los hábitos lingüísticos del hablante.



- Son hábitos lingüísticos:
  - el número de palabras de 2/3 letras en una oración.
  - La longitud de oración.
  - Palabras que empiezan por vocal.

# Estado de la cuestión en los estudios de determinación/atribución de autoría: El análisis del discurso forense.



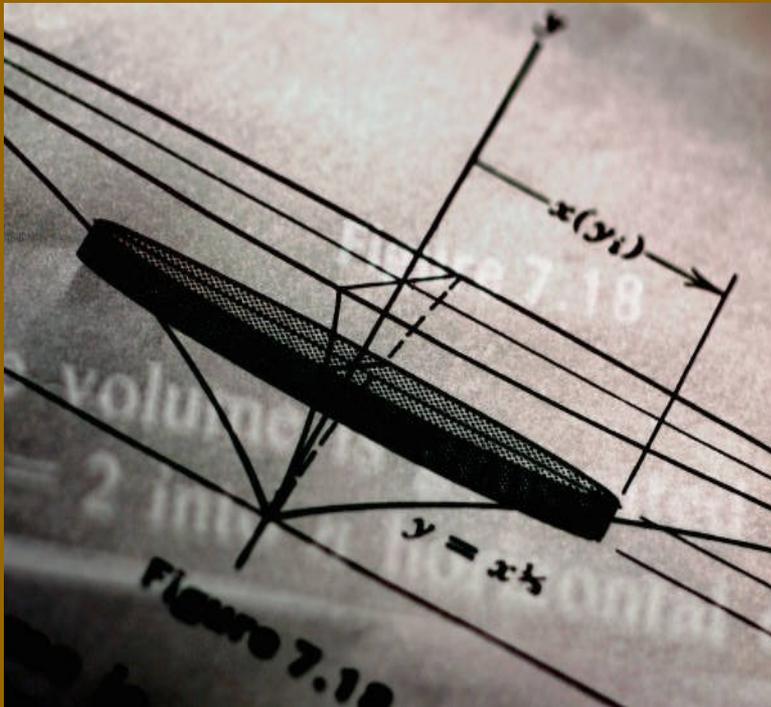
- Malcolm Coulthard (1990) analizó declaraciones tomadas por la policía aplicando herramientas lingüísticas tomadas del análisis del discurso. Identificó manipulaciones textuales, y a él se debe la exculpación a título póstumo de Derek Bentley.

# Estado de la cuestión: Estilística Forense y Atribución de Autoría.



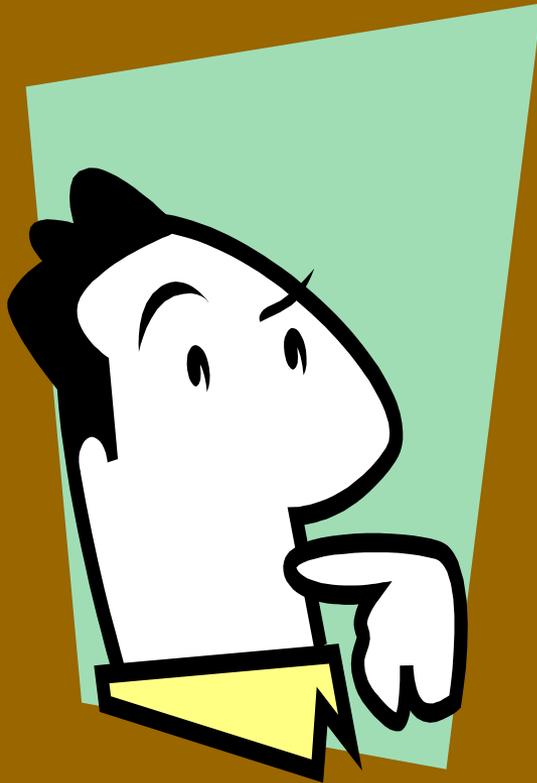
- McMenamin (1993) compara un conjunto de TD y TI sobre la base de categorías que considera marcas idiosincrásicas del estilo comunicativo del autor de los TD tras minuciosa observación empírica:
  - Diseño de la página.
  - Puntuación.
  - Ortografía.
  - Léxico.
  - Gramática.

# Estado de la cuestión: Estilística Forense y Atribución de Autoría.



- El principal problema que presentan los análisis de tipo cualitativo es el cálculo del índice de probabilidad de que un conjunto de rasgos lingüísticos sean marca de autoría.

# Análisis cualitativo: El caso del tesorero del club canino (John Olsson)

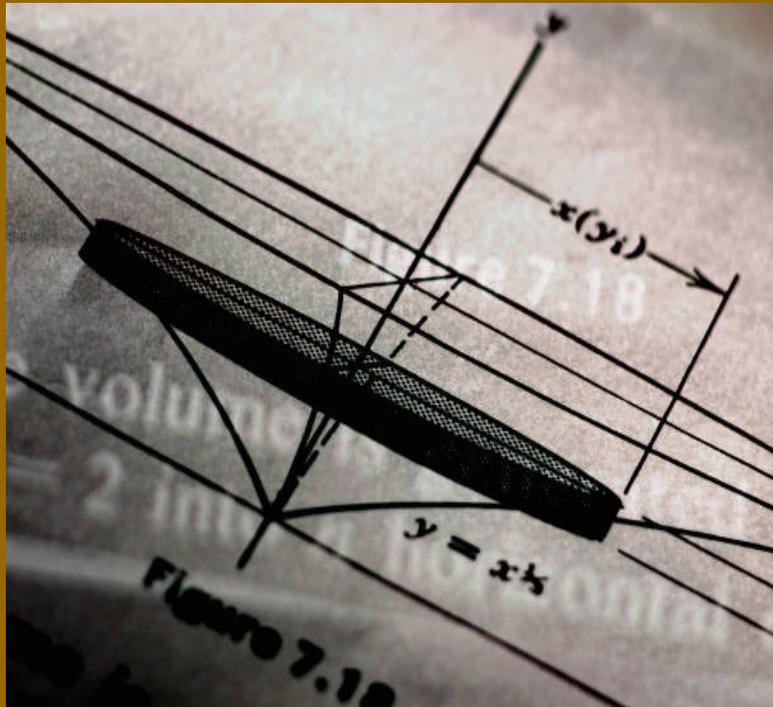


- Observación.
- Descripción cualitativa:
  - Gramática, ortografía y uso de metáforas.
  - Puntuación.
  - Uso de mayúsculas.
  - Disposición del texto de las cartas: fecha, sangrado de los primeros párrafos, y el saludo.
- Inferencia.

# Marcas Lingüísticas halladas en los TD

Ortografía	Apologies en vez de apologize
Confusión de categoría gramatical	Been en vez de being
Confusión de categoría gramatical	to en vez de too
Puntuación superflua	Punto + coma por coma
Uso de mayúsculas poco convencional	Check, Rang, etc.
Fecha	“Date: September 9 1996”
Sangrado	Diferente sangrado para los primeros párrafos.
Saludo	“Yours Sincerely” en vez de la forma más usual “Yours sincerely”

# ¿Qué se entiende por evidencia científica?



- El principal problema que presenta el análisis cualitativo de John Olsson es cómo se cuantifica la probabilidad con que se puede cometer un error de los identificados como marca lingüística de autoría.

# Estado de la cuestión: Estilística Forense y Reconocimiento de Autoría.



- Heike Hänlein (1998) ha estudiado el estilo de varios periodistas de la revista Time. El reconocimiento de autoría consiste en poner de relieve las marcas de estilo idiosincrásicas de un autor que nos permiten diferenciarle de otros.
- Hänlein considera el estilo de un autor como una proyección de su carácter o la vestidura de su pensamiento.

# Estado de la cuestión: Estilística Forense y Reconocimiento de Autoría.



- De acuerdo con Heike Hänlein (1998), el perfil estilístico de un autor se construye a partir de la selección de rasgos que ha ido haciendo en todos los niveles del lenguaje: fonofonológico, morfosintáctico, léxico-semántico y pragmático-discursivo.
- Si una selección es recurrente, ésta se convierte en una marca de estilo prominente.

# Estado de la cuestión: Estilística Forense y Reconocimiento de Autoría.



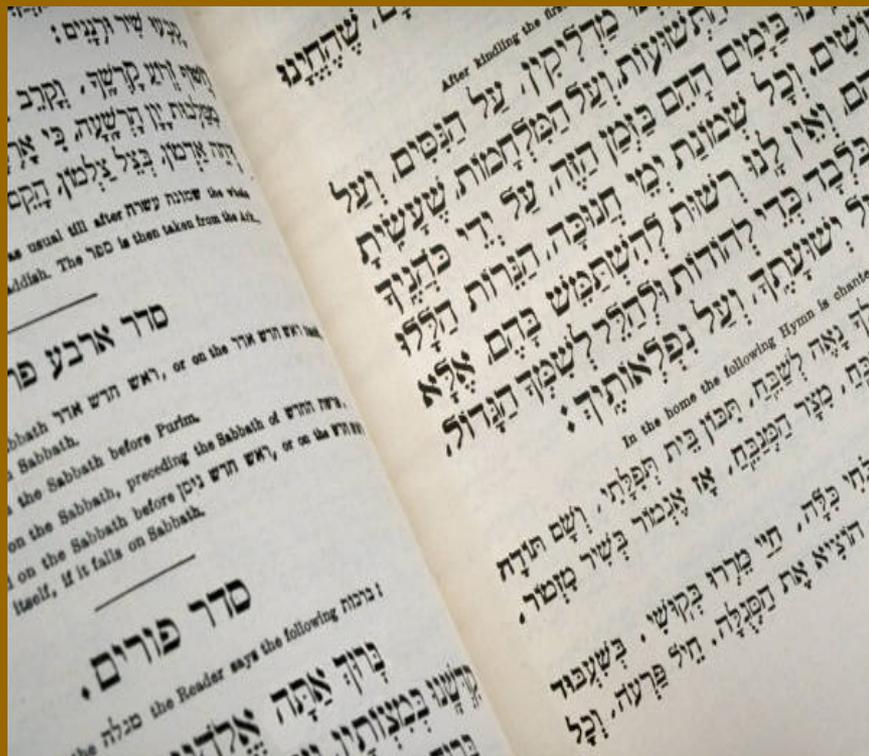
- Para construir el perfil estilístico de un autor es preciso:
  - Crear un corpus de referencia (género, período, dialecto, campo del discurso, medio, registro, y autoría).
  - Selección de categorías en su procedimiento de análisis: listado de palabras de alta frecuencia, palabras clave, y nombres propios.
  - Función comunicativa del texto.

# La metáfora de la huella lingüística.



- ¿ADN?
- ¿Huella lingüística?
- ¿Es posible reconstruir la huella lingüística de un individuo?
- La adquisición del lenguaje, la cultura y la educación.
- Conjunto de marcas idiosincrásicas prominentes del estilo de un autor.
- Marcas lingüísticas que sirven para caracterizar el idiolecto de un autor.

# La reconstrucción del perfil lingüístico de un autor.



¿Cómo se creó el texto?

¿Cuántos autores intervinieron en la producción del texto?

¿Se aprecian secciones diferentes en cuanto a estilo en el texto?

# Investigaciones en Atribución de autoría: Dualidad Textual

Modalidad dual

Autoría dual

Registro dual

Cronicidad dual

Producción física dual

Objetivos discursivos duales

# Tipo 1: Análisis contrastivo

- Dualidad léxica.
- Dualidad en cuanto a los referentes empleados (nombres, nomenclatura, ortografía, etc.)
- Distribución de una o varias palabras relevantes de alta frecuencia en una sola sección del texto.

# Tipo 1: Análisis contrastivo: Caso Derek Bentley

- Discrepancia en el promedio de número de palabras por frase en las secciones inicial y final del texto, y el cuerpo del texto.
- Dualidad en cuanto al referente: Craig, Chris Craig (p.174); Chris (a partir de la p. 174).
- Análisis estadístico para calcular el índice de probabilidad: promedio de palabras por frase, desviación típica, cálculo de probabilidad, hipótesis de trabajo, hipótesis contraria, T-Test.

# Tipo 1: Análisis contrastivo: Caso carta de rescate Ramsey

- Dualidad en el referente empleado: “I”, “We”.
- Distribución localizada de los pronombres personales “We” y “I” en distintas secciones del texto.
- Distribución localizada de la cláusula “she dies” en una sección del texto.
- Distribución localizada del símbolo \$ en una sección del texto.
- Distribución localizada de “John” en una sección del texto.

# Tipo 1: Análisis contrastivo: Caso carta de rescate Ramsey.

- Observación: ¿Nos encontramos ante un texto producido por varios autores? ( $P_n$ )
- Hipótesis 1: Podríamos encontrarnos ante un texto producido por varios autores. (TS)
- Observación: ¿Existe posibilidad de error? (EE)
- Hipótesis 2: El texto podría haber sido producido por un solo autor ( $P_{n+1}$ )

Test de Popper  $P_n - TS - EE - P_{n+1}$

# Tipo 2: Investigación de la variación de registro: Caso carta de suicidio Gilfoyle.

- Observación: ¿Nos encontramos ante un texto producido por varios autores? (Pn)
- Hipótesis 1: Podríamos encontrarnos ante un texto producido por varios autores. (TS)
- Presencia de pareados de nombres (colocaciones) en algunos párrafos: “pain and heartache”, “you and your life”, “realize your goals and dreams”, etc.

# Tipo 2: Investigación aislada de la variación de registro: Caso carta de suicidio Gilfoyle.

- Se observa un uso exclusivo de “and”/”or” como conector de nombres o expresiones (pareados) en algunos párrafos, mientras que en otros aparece exclusivamente como nexos entre cláusulas coordinadas.
- Se observa mayor promedio en la longitud de palabras en los párrafos que contienen los pareados.

# Tipo 2: Investigación de la variación de registro: Caso carta de suicidio Gilfoyle.

- Se observa un menor promedio en la longitud de oración en los párrafos que contienen los pareados.
- La función comunicativa de los párrafos que contienen los pareados es acusatoria (Paula) e indirectamente exculpatoria (Eddie).
- La función comunicativa de los párrafos que no contienen los pareados es meramente informativa.

# Tipo 2: Investigación de la variación de registro: Caso carta de suicidio Gilfoyle.

Estilo de párrafo 1	Estilo de párrafo 2
Estructuras pareadas “And” (nexo de palabras/frases) Oraciones cortas Palabras largas	No hay estructuras pareadas “And” (nexo de oraciones coordinadas) Oraciones largas Palabras cortas

# Tipo 3: Investigación basada en el análisis de marcas prominentes en el estilo de un autor: Caso Lindbergh.

Confusión en función agente-paciente

Confusión en el uso de enunciados afirmativos en vez de interrogativos

Confusión en el uso de categorías gramaticales, ej. Nombres por adjetivos, adverbios por adjetivos

Confusión en el uso de pronombres personales de primera y segunda persona.

Confusión en sonidos consonánticos, ej. “s” por “l”, creación de palabras nuevas basadas en las ya existentes, ej. “impollibly” por “Impossible”.

# Determinación/Atribución de Autoría

- Test de Daubert (USA Art. 702, Testifical pericial): (a) Informe de experto, (b) Técnica empíricamente comprobada, (c) Técnica sujeta a la crítica y revisión de otros expertos, (d) Método científico, y (e) Sencillez y economía en la exposición de argumentos.
- Observación.
- Descripción (procedimiento científico):
  - Análisis cualitativo.
  - Análisis cuantitativo.
- Inferencia.

# Conclusión:

## Determinación/Atribución de Autoría

- Flexibilidad/Interdisciplinariedad.
- Peritaje lingüístico: (a) investigación preliminar, (b) dictamen.