



Trabajo Fin de Grado:
**“Traducción al español de términos en inglés
usados en Genética general analizados en
diferentes contextos”**

Granada, Junio 2016

Grado en Traducción e Interpretación

Rafael Navajas-Pérez



Declaración de Originalidad del TFG

(Este documento debe adjuntarse cuando el TFG sea depositado para su evaluación)

D. **Rafael Navajas Pérez**, con DNI ..., declaro que el presente Trabajo de Fin de Grado es original, no habiéndose utilizado fuente sin ser citadas debidamente. De no cumplir con este compromiso, soy consciente de que, de acuerdo con la Normativa de Evaluación y de Calificación de los estudiantes de la Universidad de Granada de 20 de mayo de 2013, esto conllevará automáticamente la calificación numérica de cero [...] independientemente del resto de las calificaciones que el estudiante hubiera obtenido. Esta consecuencia debe entenderse sin perjuicio de las responsabilidades disciplinarias en las que pudieran incurrir los estudiantes que plagie.

Para que conste así lo firmo el

Firma del alumno

Índice

I.- Agradecimientos.....	6
II.- Resumen	8
III.- Índice de tablas	10
IV.- Índice de figuras	12
1.- Introducción	15
1.1.- Generalidades	15
1.2.- El inglés como lengua franca en textos científicos	19
1.3.- Tipos de discurso: especializado, didáctico y divulgativo.....	23
1.4.- Terminología genética: categorías, conceptos y relaciones conceptuales	24
2.- Objetivos.....	29
3.- Materiales y métodos	32
3.1.- Delimitación temática de los textos analizados y creación de corpora.....	32
3.2.- Minería de textos	33
4.- Resultados y Discusión.....	36
4.1.- Criterios de selección de términos y textos paralelos.....	36
4.2.- Análisis de problemas.....	38
5.- Conclusiones	63
6.- Bibliografía	68
7.- Anexos	73

I.- Agradecimientos

Me gustaría agradecer a la Dra. Pamela Faber por la supervisión del presente trabajo. Al Dr. Manuel Ruiz Rejón por sugerir algunos de los términos estudiados en esta memoria. A D. Jonatan Cogolludo por la ayuda en la conversión de formato de algunos archivos usados para la creación de los *corpora* usados en el análisis. A la Universidad de Granada por sufragar mis estudios de traducción.

II.- Resumen

Gran parte del lenguaje especializado utilizado en Genética proviene del inglés, lengua franca usada entre científicos de distintas nacionalidades y por la mayor parte de revistas científicas. Al tratarse de una ciencia relativamente reciente, existe una cierta ambigüedad e incluso imprecisión en la traducción de algunos términos genéticos. A pesar de ser correspondencias poco acertadas, estos términos acaban proliferándose en artículos, manuales y libros de texto. Su traducción poco acertada repercute en todos los niveles textuales.

En el presente trabajo se propone llevar a cabo un estudio de la traducción al español de los términos más comunes en la materia mediante el análisis de *corpora* de textos paralelos en contexto educativo, divulgativo y científico.

III.- Índice de tablas

Tabla 1.- Términos con mayor frecuencia de uso en los *corpora* en inglés analizados.

Tabla 2.- Términos en inglés y las variantes detectadas en español correspondientes a entidades.

Tabla 3.- Términos en inglés y las variantes detectadas en español correspondientes a procesos.

Tabla 4.- Términos en inglés y las variantes detectadas en español correspondientes a atributos.

Tabla 5.- Frecuencia de aparición de 'ADN/DNA' en los *corpora* analizados. **Nota:** valores expresados en número de coincidencias/número de coincidencias por millón de palabras.

Tabla 6.- Frecuencia de aparición de 'ARN/RNA' en los *corpora* analizados. **Nota:** valores expresados en número de coincidencias/número de coincidencias por millón de palabras.

Tabla 7.- Frecuencia de aparición de las distintas variantes de 'PCR' en los *corpora* analizados. **Nota:** valores expresados en número de coincidencias/número de coincidencias por millón de palabras.

Tabla 8.- Verbos que utilizan 'ADN/DNA' como sujeto o como objeto en los *corpora* de textos en español. En número se representan las coincidencias encontradas.

Tabla 9.- Verbos que utilizan 'DNA' como sujeto o como objeto en los *corpora* de textos en inglés. En número se representan las coincidencias encontradas.

Tabla 10.- Frecuencia de aparición de las distintas variantes de 'splicing' en los *corpora* analizados. **Nota:** valores expresados en número de coincidencias/número de coincidencias por millón de palabras.

Tabla 11.- Frecuencia de aparición de las distintas variantes de 'deletion' en los *corpora* analizados. **Nota:** valores expresados en número de coincidencias/número de coincidencias por millón de palabras.

Tabla 12.- Frecuencia de aparición de las distintas variantes de 'cromosómico/cromosomal', 'ribosómico/ribosomal', 'vírico/viral' en los *corpora* analizados. **Nota:** valores expresados en número de coincidencias/número de coincidencias por millón de palabras.

Tabla 13.- Frecuencia de aparición de las formas *zig* con respecto a las formas *cig* en los *corpora* en inglés y español analizados. **Nota:** valores expresados en número de coincidencias/número de coincidencias por millón de palabras.

Tabla 14.- Frecuencia de aparición de las formas *zig* con respecto a las formas *cig* en los *corpora* analizados. **Nota:** valores expresados en número de coincidencias/número de coincidencias por millón de palabras.

Nótese que aparecen coincidencias en el *corpus* divulgativo (que carece de las formas heteroc(z)igoto, homoc(z)igoto, hemic(z)igoto) por la presencia de la forma cigoto/zigoto.

Tabla 15.- Frecuencia de aparición de 'recombinación' y 'entrecruzamiento' en los *corpora* analizados. **Nota:** valores expresados en número de coincidencias/número de coincidencias por millón de palabras.

Tabla 16.- Frecuencia de aparición de 'crossing over', 'crossover' y 'recombination' en los *corpora* analizados. **Nota:** valores expresados en número de coincidencias/número de coincidencias por millón de palabras.

IV.- Índice de figuras

Figura 1.- Escenificación de la Teoría Sintética de la Evolución, que aúna las teorías sobre selección natural de Darwin y las leyes de la herencia de Mendel. (Fuente: Pasaje a la Ciencia, 2009; Autor: Fran Moyano).

Figura 2.- Análisis del histórico (1949-2007) de la frecuencia de uso de la palabra DNA. Datos expresados en coincidencias por millón de palabras por año. (Fuente: Wolfram Alpha).

Figura 3.- Evolución anual del número de publicaciones científicas recogidas en el SCI Expanded y porcentaje de España sobre el total. (Fuente: Web of Science. Gráfico extraído de Bordons, 2004).

Figura 4.- Mapas conceptuales de la clasificación del ADN utilizando como criterio fundamental (A) su localización o (B) su función.

Figura 5.- Esquema categorial para el término 'splicing'.

Figura 6.- Esquema categorial para el término 'contig'.

Figura 7.- Esquema categorial para el término 'deletion'.

Figura 8.- Nube de palabras construida para el término 'deletion' a partir de las coincidencias encontradas en el *corpus* educativo.

Figura 9.- Esquema comparativo de los términos modificados por las voces 'génico', 'genético' y 'genómico' comparadas por pares.

Figura 10.- Esquema comparativo de los términos modificados por las voces 'genético' y 'genetista'.

Figura 11.- Esquema comparativo de los términos modificados por y de los términos que modifican las voces 'entrecruzamiento' y 'recombinación'.

1.- Introducción

1.- Introducción

*English is THE language of communication
and it never occurred to me that anybody
who knows anything about the dynamics of science today
would even question the issue.*

(Pere Alberch, director del Museo Nacional de Ciencias Naturales de Madrid, 1994)

1.1.- Generalidades

1.1.1.- *¿A qué se dedica la Genética? Breve introducción a esta disciplina*

Desde tiempo inmemorial, el hombre es consciente de la similitud que existe entre padres e hijos, siendo evidente que, al igual que ocurre en plantas y animales, la descendencia suele mantener, al menos, algunos de los rasgos presentes en los parentales. Los primeros intentos de explicar la transmisión de caracteres de generación en generación fueron llevados a cabo durante el siglo IV a.c., en la antigua Grecia. La teoría que sustituyó al concepto de generación espontánea de Aristóteles fue la de la *herencia particulada*, que postulaba que las distintas partes del cuerpo de uno o ambos progenitores se fragmentaban y unían posteriormente para dar lugar a la progenie (Science Encyclopedia, 2016). Si bien, estos postulados nunca encontraron evidencia científica alguna que los soportara.

No fue hasta mucho tiempo después que se le comenzó a dar una explicación a la herencia. Tras la publicación de *El origen de las especies* de Darwin (1859), se había generalizado en la comunidad científica la idea de que los individuos con características más ventajosas tenían más posibilidades de reproducirse, y por tanto de hacer pasar a la descendencia sus rasgos, que serían cada vez más abundantes en la población. Sin embargo, aún faltaba la pieza clave del puzzle que explicara cómo se producía esa transferencia de información de padres a hijos. ¿Qué hacía que un rasgo fuera heredable? ¿Cómo funcionaba esta herencia? El propio Darwin trató sin éxito de explicar este fenómeno (adaptado de Navajas-Pérez & Aznarte-Mellado, 2014).

El monje agustino Gregor Mendel, a través del enunciado de sus leyes, se encargó de encontrar una explicación experimental consistente (1866). Aprovechando su habilidad para la jardinería y sus conocimientos en análisis matemático, llevó a cabo una serie de experimentos usando como modelo distintas variedades de guisante que diferían en una o pocas características. Con meticulosidad se encargaba de cruzarlas, para posteriormente contar, clasificar y estudiar las proporciones en las que aparecía cada una de esas características en la descendencia. Sólo así fue capaz de darse cuenta de la reproducibilidad del modo en que se heredaban algunas características. Llegó a la conclusión de que los rasgos venían determinados por dos factores

hereditarios (que hoy llamamos genes), uno aportado por el padre y otro por la madre, y que esto sería extrapolable al resto de seres vivos (adaptado de Navajas-Pérez & Aznarte-Mellado, 2014).

Probablemente el nacimiento *de facto* de la Genética como disciplina coincide con el enunciado de las Leyes de Mendel, lo que le sirvió para pasar a la historia como el *Padre de la Genética*. Un éxito relativo, ya que moriría en 1884 sin que sus trabajos hubieran alcanzado todavía repercusión. De hecho, pasaron desapercibidos durante mucho tiempo, hasta que fueron redescubiertos de forma casi simultánea por los investigadores Hugo de Vries (1900), Carl Correns (1900) y Erich von Tschermak (1900). El renacer de las teorías de Mendel, el hecho de que sus resultados fueran aplicables a un número creciente de organismos y el enunciado de la *Teoría Cromosómica* por parte de Thomas H. Morgan (1910), que correlacionaba el comportamiento cromosómico con las nuevas ideas sobre la herencia de Mendel, hizo que a finales de la primera década del siglo XX la teoría gozara ya de buen predicamento entre la comunidad científica. La Biología moderna ha unido los estudios complementarios de Darwin y Mendel en la llamada *Teoría Sintética de la Evolución*, de gran vigencia en la actualidad con ciertas modificaciones (**Figura 1**).



Figura 1.- Escenificación de la Teoría Sintética de la Evolución, que aúna las teorías sobre selección natural de Darwin y las leyes de la herencia de Mendel. (Fuente: Ruiz Rejón, 2009; Autor: Fran Moyano).

Lo descrito hasta el momento coincide a la perfección con la cuarta acepción del término '*Genética*' del DRAE:

4. f. parte de la biología que trata de la herencia y de lo relacionado con ella.

Sin embargo, esta definición es reduccionista, ya que esta ciencia abarca hoy día muchos más ámbitos aparte del puramente relacionado con la herencia de los caracteres. De hecho, con el esclarecimiento de la estructura molecular del ADN por parte de Franklin, Watson y Crick (1953) y el desarrollo de nuevas técnicas, la Genética se ha hecho hueco en muchos campos de nuestras sociedades modernas, y se habla ya de *Genética moderna o molecular* en contraposición a la antigua *Genética mendeliana* (Allen, 2003). La Genética es, de hecho, la pieza fundamental en la mayor parte de investigaciones biomédicas y biotecnológicas, en la medicina forense, la criminalística, la producción animal y agrícola o la bioética, por citar sólo algunos ejemplos. Eco de ello se hacen los planes de estudio de enseñanzas obligatorias y superiores y los medios de comunicación, como a continuación detallaremos.

1.1.2.- La enseñanza de la Genética y su discurso

La enseñanza de la Genética en distintos niveles educativos ha centrado multitud de estudios metodológicos (revisado en Abril Gallegos et al., 2003). Esto es reflejo de su importancia en muchos ámbitos de nuestra vida cotidiana. Hay un consenso sobre los principales motivos que impulsan a mejorar la comprensión de los conceptos relacionados con este campo. Así, existe una *razón utilitaria*, que permitiría la aplicación práctica del conocimiento científico, una *razón democrática*, que permitiría posicionarnos ideológicamente en la sociedad, (siendo capaces, por ejemplo, de diferenciar ciencia de pseudociencia), y una *razón cultural*, entendida como un avance de la sociedad moderna (Turney, 1995). En el sistema educativo español, los alumnos entran en contacto por primera vez con la Genética a los 15-16 años (4º curso de la E.S.O.), aunque hay autores que sugieren que esta edad podría adelantarse hasta incluso los 10 años (Donovan & Venville, 2012).

En la actualidad, aparte del gran número de empresas farmacéuticas y del sector de la Biotecnología, existen en España 54 centros públicos que tienen como actividad principal a la investigación Genética, 40 de ellos pertenecientes a universidades que se encargan de la enseñanza superior de esta ciencia (Sociedad Española de Genética, 2016). En concreto, en el Departamento de Genética de la Universidad de Granada se imparten, a día de hoy, un total de 11 asignaturas en los Grados en Biología, en Bioquímica, en Ciencias Ambientales y en Biotecnología. Entre ellas, se encuentran las materias de Biología Evolutiva, Genética de la Conservación y Medio Ambiente, Genética Humana, Genómica e Ingeniería Genética, por citar algunos ejemplos. Además, este mismo departamento coordina el máster universitario en Genética y Evolución (Departamento de Genética UGR, 2016).

Salvo en los programas bilingües, la enseñanza obligatoria se imparte en la lengua nativa de los estudiantes. Incluso en educación superior existe una tendencia cada vez mayor a utilizar recursos online en la lengua local (Holmes, 2004), si bien muchos materiales de apoyo y lecturas científicas se encuentran escritos en inglés. No obstante, los libros de texto y las aulas no son las únicas fuentes de donde los estudiantes obtienen información relacionada con la Genética. Es cada vez más frecuente encontrar alusiones a esta disciplina en radio, prensa y televisión, en forma de programas enfocados a distintas audiencias, noticiarios, revistas de divulgación, documentales y ficciones de cine o televisión (Abril Gallegos et al., 2003). Según apuntan estos mismos autores, el hecho de que los medios de comunicación traten con frecuencia este tema, da a los estudiantes más noveles la falsa impresión de dominar los términos. Esto no siempre es cierto; en primer lugar porque en muchas ocasiones la información está sesgada (hacia la investigación criminal, las ciencias forenses o la medicina, por ejemplo -Donovan & Venville, 2012), y en segundo lugar porque en estos medios es normal encontrar errores de concepto importantes (Abril Gallegos et al., 2003). Todo ello influye en la percepción que tiene el estudiante de su conocimiento sobre Genética, y podría redundar en dificultades posteriores durante su etapa de aprendizaje formal.

1.1.3.- La Genética en los medios de comunicación

Como propone el eminente catedrático de Genética D. Juan Ramón Lacadena (2011), *la revolución del ADN está produciendo [...] una "Biocracia" a través de la Biotecnología*, y los medios de comunicación son un fiel reflejo de ello. Es frecuente encontrar diariamente noticias en las que la Genética ocupa un papel importante: resolución de crímenes mediante Genética forense, organismos transgénicos, clonación, terapia génica, células madre o alusiones a la evolución, podrían ser algunos ejemplos. Así, por ejemplo, desde su primera mención documentada en 1944, el uso de la palabra 'DNA' ha sufrido un incremento exponencial hasta nuestros días, con millones de coincidencias (**Figura 2**).

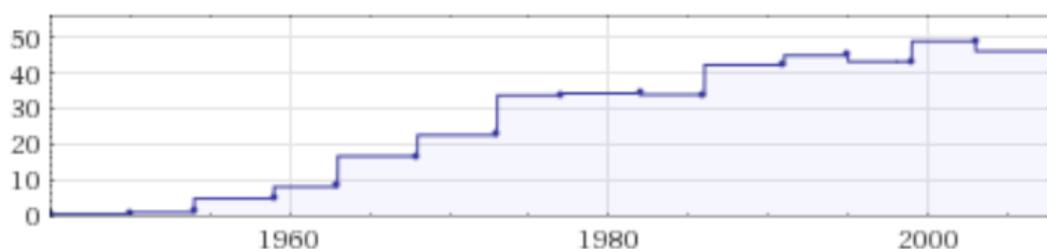


Figura 2.- Análisis del histórico (1949-2007) de la frecuencia de uso de la palabra DNA. Datos expresados en coincidencias por millón de palabras por año. (Fuente: Wolfram Alpha).

Aunque existe una tendencia cambiante (ver por ejemplo los artículos de J.M. Mulet en el diario El País 2016), la ciencia es comunicada mayoritariamente por periodistas o especialistas en Ciencias de la Comunicación (esto incluye, además, a *bloggers*, *community managers*, etc...). Aunque estos profesionales son capaces de plantear las preguntas que potencialmente interesan a un ciudadano de a pie, en ocasiones, su mensaje es confuso, equivocado y/o sesgado. Así, por ejemplo, los espectadores más jóvenes tienden a asociar el ADN con la sangre y los actos delictivos, y la mutación con enfermedades y malformaciones (Donovan & Venville, 2012).

Parte de estos errores puede deberse a la temática tratada cuando se hace referencia a la Genética, pero también a malas traducciones hechas por los comunicadores o los propios científicos. Se estima que hasta un 87% de los documentos multimedia en red están escritos en inglés, a pesar de que un 90% de los usuarios preferirían usar una lengua distinta (Tercedor, 2005), y de que un porcentaje considerable de usuarios no son nativos de habla inglesa (Marcus, 2002).

Es interesante pensar que no hace más de 50 o 60 años se desconocía casi totalmente la composición química del ADN. Sin embargo, en la actualidad somos capaces de secuenciar, clonar y modificar organismos, diferenciar tejidos celulares específicos, datar procesos que ocurrieron hace millones de años, diagnosticar enfermedades o caracterizar molecularmente organismos, incluso a nivel de individuo. El lenguaje científico de algunas disciplinas se ha ido desarrollando al ritmo (muchas veces frenético) marcado por los distintos descubrimientos e investigaciones. La rápida expansión de la Genética y el hecho de que se comunique mayoritariamente en inglés, ha contribuido a modelar un lenguaje científico cuyas particularidades se detallan a continuación.

1.2.- El inglés como lengua franca en textos científicos

1.2.1.- *Motivos históricos de la hegemonía del inglés en textos científicos*

Aunque en los siglos XVIII y XIX hubo intentos (infructuosos) de crear un lenguaje científico artificial, no cabe duda de que, en la actualidad, el inglés es la lengua oficial de la ciencia. Probablemente, no exista un único factor que permita explicar este hecho, aunque el dominio económico y político de Estados Unidos ha sido decisivo en este proceso. El declive de

la ciencia alemana después de la Primera Guerra Mundial, lengua en la que se comenzaron a escribir algunos ensayos científicos importantes del siglo XIX, parece haber favorecido el avance del inglés (Gordin, 2015).

Tras la Segunda Guerra Mundial, un número significativo de estudios científicos vieron la luz en ruso. De hecho, en este periodo se inicia un proceso de profesionalización de los traductores (en 1953, por ejemplo, se crea la Federación Internacional de Traductores –Marchio & Mackenbach, 2012). Aquellos científicos hablantes nativos de terceras lenguas, se decantaron finalmente por los estudios traducidos al inglés, probablemente a consecuencia de la menor dificultad del aprendizaje de éste con respecto al ruso (Gordin, 2015).

Desde entonces, el desarrollo científico ha dado lugar a la evolución de la lengua general, de tal forma que ésta sirve de vehículo para la transmisión del conocimiento científico. Esto ha originado diferencias entre ambos lenguajes a nivel semántico, funcional y estructural (Alcina Caudet, 2001):

1. *Diferencias semánticas*: el significado de los términos del lenguaje científico es preciso y objetivo. Esto contrasta con lo que ocurre en la lengua común, en la que se admiten connotaciones subjetivas y cierta ambigüedad
2. *Diferencias funcionales*: el objetivo de ambos tipos de lenguajes es radicalmente distinto, siendo el del lenguaje científico el de intercambiar información objetiva sobre un tema determinado, y el de la lengua general, la interacción social y la afectividad.
3. *Diferencias estructurales*: algunos elementos lingüísticos típicos del lenguaje científico son: definiciones, clasificaciones, enumeraciones, cálculos, razonamientos, argumentaciones y citas, mientras que en la lengua general abundan las narraciones, los diálogos y las interrogaciones.

La escritura científica ha desarrollado, por tanto, su propio estilo, sus propios usos y normas, y sobre todo se ha adaptado a un público con capacidades no nativas de escribir y entender el inglés (Burnett, 1999). Esta tendencia ha originado el *inglés científico*, una especie de *pseudolengua* usada específicamente en este ámbito. Sin embargo, aún el nivel de inglés de los manuscritos sometidos a revisión en revistas científicas es un motivo principal para la no aceptación de los mismos. Esto es especialmente cierto cuanto mayor es el prestigio y/o el índice de impacto de la publicación, y podría dar lugar a investigaciones de primera e investigaciones de segunda sólo por el mero hecho de no ser comunicadas por un hablante nativo de inglés (Meneghini & Packer, 2007).

1.2.2.- Consecuencias de la hegemonía del inglés en el lenguaje científico

La hegemonía actual del inglés como *lingua franca* en los textos científicos tiene consecuencias puramente formales (cómo se hace y se comunica la ciencia), pero también de índole económica (quién controla las editoriales, las revistas científicas o disfruta de proyectos de investigación) y socio-cultural (educación científica, colaboraciones internacionales, repercusión internacional) (Gordin, 2015). Además, tiene también una consecuencia directa sobre el lenguaje científico del idioma meta, ya que se tiende a reproducir convenciones típicas de la lengua origen que son errores ortotipográficos en la lengua meta: abreviaturas, signos de puntuación, fórmulas matemáticas, uso de las mayúsculas y las cursivas, sangrías (Claros, 2008). El papel del traductor es aquí fundamental para producir una traducción correcta no sólo del punto de vista semántico, sino también gramatical. Para Domínguez (2007), los diez errores más frecuentes en la traducción de un artículo científico son:

1. Anglicismos en la adjetivación.

2. Abuso de los adverbios terminados en –mente.

3. Empleo de frases demasiado largas.

4. Uso incorrecto de los gerundios.

5. Notación incorrecta de los decimales.

6. Conversión de las unidades de medida.

7. Equivalencia de siglas.

8. Orden incorrecto en la sintaxis.

9. Supresión de los artículos.

10. Abuso de la voz pasiva.

Pueden encontrarse también diferencias según la lengua meta empleada. En teoría, sólo aquéllas que pueden representar el pensamiento abstracto tendrían la capacidad de comunicar el pensamiento científico de forma eficiente. Esto es cierto en la mayoría de casos, salvo que la lengua conviva con otra de mayor prestigio social, que se encuentre en claro retroceso o en fase de desaparición (Bernárdez, 1999), por lo que se pueden argüir razones sociales y de evolución

de las propias lenguas a su adecuación al lenguaje científico, más que motivos intrínsecos del propio idioma (Alcina Caudet, 2001).

Así por ejemplo, las publicaciones locales escritas en lenguas más minoritarias suelen carecer de prestigio, reputación y/o índice de impacto por lo que, en líneas generales, comunicar en inglés beneficia al investigador, ya que le permite aumentar su visibilidad y su prestigio internacional, al tiempo que hace accesible su trabajo a un espectro más amplio de público, propiciando el intercambio, la colaboración y la comunicación con otros investigadores (Meneghini & Packer, 2007). Aparte de los inconvenientes propios de comunicar en una lengua ajena, usar el inglés como lengua franca puede causar además que el mensaje tarde en llegar al entorno del propio investigador y que las comunidades locales no obtengan beneficio inmediato de las mismas (esto es especialmente interesante si pensamos que la mayor parte de las investigaciones se llevan a cabo con fondos públicos). Valga como ejemplo el estudio desarrollado por investigadores alemanes en los años 30, que asociaba el cáncer de pulmón al tabaquismo. Estos trabajos pasaron inadvertidos fundamentalmente por estar escritos en alemán. De hecho, los programas antitabaco no se comenzaron a desarrollar hasta los años 60 cuando investigadores británicos y americanos llegaron a la misma conclusión (Proctor, 1999).

Existe, por tanto, un equilibrio de fuerzas contrapuestas. Por un lado, se encuentra la necesidad del científico de comunicar con otros colegas y colaboradores, alcanzando una mayor visibilidad. Por otro lado, se encuentran los intereses de los profesionales de la salud, los periodistas, los divulgadores, los ciudadanos de a pie y en definitiva el beneficiario último de las investigaciones, que encuentran más accesible la información si está escrita en su lengua nativa (Meneghini & Packer, 2007). No obstante, incluso textos escritos en la lengua nativa no inglesa, en los que se incluyen anglicismos, pueden dar lugar a confusión y problemas entre los hablantes (ver **Anexo I**).

Este dilema tiende a resolverse de dos formas opuestas. Existen corrientes mayoritarias que abogan por la consolidación del inglés como única lengua de la ciencia, promoviendo la educación para una sociedad bilingüe (Ofori-Adjei et al., 2006), o incentivando la comunicación de la ciencia en inglés, como en China o Japón (Meneghini & Packer, 2007). Por el contrario, se advierte una cierta tendencia a dar mayor importancia a otras lenguas. Así, grupos editoriales como Plos (*Public Library of Science*), pionero en la implantación de publicaciones *Open Access*, anima a los autores a proporcionar una versión complementaria a los artículos publicados en inglés (Editorial, 2006). En algunos países, como España, ciertas revistas se publican en dos versiones, la inglesa y la de la lengua local (Bordons, 2004). Además, no hay que perder de vista

la pujanza que en algunos países tienen las investigaciones realizadas y comunicadas en la lengua local (**Figura 4**).

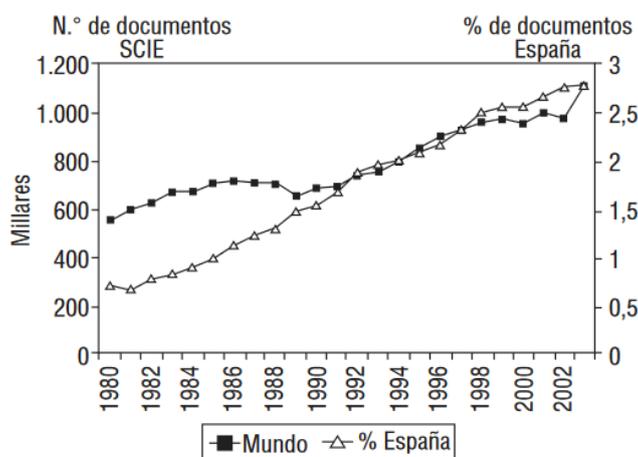


Figura 3.- Evolución anual del número de publicaciones científicas recogidas en el *SCI Expanded* y porcentaje de España sobre el total. (Fuente: Web of Science. Gráfico extraído de Bordons, 2004).

1.3.- Tipos de discurso: especializado, didáctico y divulgativo

De acuerdo con Halliday (1978), y su clasificación de factores que afectan al registro, el contexto en el que se desarrolla una situación comunicativa viene determinado por tres factores básicos: el campo, el modo y el tenor:

1. El *campo* establece el grado de especificidad de un texto, y puede estar influenciado por: a) la temática: en un discurso especializado, la temática puede presentar distintos niveles de complejidad, pero siempre se centra en un tema específico cuyo mensaje se pretende comunicar (Montero Martínez et al., 2011), y b) por el marco social en el que se desarrolla la práctica comunicativa: actas de un congreso, artículo de investigación o con fines divulgativos (Centro Virtual Cervantes, 2016).
2. El *modo* tiene en cuenta el medio o el canal utilizado para la comunicación y determina el grado de planificación del discurso (Centro Virtual Cervantes, 2016). Así, puede tratarse de una comunicación oral a un congreso, una charla divulgativa en un colegio o la lectura comentada de un artículo de investigación.
3. El *tenor* considera las características de los usuarios (tenor interpersonal o tono) y el final último del acto comunicativo (tenor funcional). En relación al tono, se ha de tener en cuenta que los usuarios del discurso especializado tienen generalmente un cierto conocimiento del tema, o al menos estar en disposición de descodificar el mensaje (Montero Martínez et al., 2011). En cualquier caso, de ello dependerá el grado de

formalidad y el grado de subjetividad (Centro Virtual Cervantes, 2016). Por su parte, el *tenor funcional*, determina la función principal del texto, que en los textos científicos puede ser argumentativa, explicativa, descriptiva, e incluso persuasiva.

A pesar de que no existe una noción universal de contexto, es posible establecer unos criterios basados fundamentalmente en el fin último de la comunicación y el tipo de información que se transmite, tanto a nivel sintáctico como semántico y pragmático (Faber & León-Araúz, 2016). Así, en este trabajo asumiremos que el emisor es un especialista o un comunicador especializado que se dirige a un receptor con distintos grados de formación, cuyo conocimiento quiere modificar. Así, basándonos en la clasificación de Montero Martínez y colaboradoras (2011), hablaremos de tres tipos de discurso:

1. El *discurso especializado*, hecho por y para especialistas: artículos de investigación, comunicaciones a congresos y reuniones científicas, *workshops* o memorias de proyectos de investigación.
2. El *discurso didáctico*, hecho por especialistas y dirigido a estudiantes o personal en formación: guías, apuntes, manuales y libros de texto.
3. El *discurso divulgativo*, hecho por especialistas o mediadores y destinado a un público más general, de alguna manera interesado en un tema en cuestión: medios de comunicación en general, artículos y charlas de divulgación y talleres.

1.4.- Terminología genética: categorías, conceptos y relaciones conceptuales

Tal y como ocurre con la clasificación de los seres vivos, agrupar la realidad en clases o categorías hace que la percepción del mundo sea más comprensible. Este modelo se corresponde con construcciones mentales fruto de la observación, la percepción, la abstracción y la comparación (revisado en Montero Martínez et al., 2011).

Se suelen considerar cuatro tipos de categorías (Sager, 1990).

1. Las *entidades* (u objetos), que se usan para describir elementos que son percibidos como una unidad diferenciable e independiente. En Genética, podríamos considerar como entidades, moléculas o estructuras físicamente independientes (ADN, ARN, cromosoma), personas (genetista), aparatos (secuenciador) y elementos mentales (evolución) (ver **Tabla 1**).

2. Los *procesos* (o eventos), son propios de la ocurrencia de un fenómeno natural o de una operación artificial. Las entidades pueden ser agentes (llevan a cabo la acción) o pacientes (la acción recae sobre ellas). En Genética, podríamos considerar como procesos, fenómenos que ocurren de forma natural en las células (entrecruzamiento) y su consecuencia (mutación), técnicas empleadas por los investigadores (secuenciación) y nombre de procedimientos (huella genética) (ver **Tabla 2**).
3. Los *atributos*, que sirven para diferenciar a las entidades y/o procesos entre sí, asignándoles cualidades y características o un valor escalar o literal. En Genética, podríamos considerar como atributos, distintos tipos de moléculas u organismos por su naturaleza (transgénico), localización (ribosómico) o efecto (epigenético) (ver **Tabla 3**).
4. Las *relaciones*, que representan las relaciones que se establecen entre entidades, entre procesos o entre entidades y procesos, por lo que se basan en relaciones jerárquicas (superordinación y subordinación). Para categorizar las relaciones es necesario establecer correlaciones del tipo «X es un tipo de Y», «X es una consecuencia de Y», «X es una parte de Y».

(Modificado de Montero Martínez et al., 2011)

Las entidades suelen estar representadas por sustantivos, mientras que los procesos suelen venir determinados por verbos y los atributos por adjetivos. Sin embargo, esto no es así siempre de ninguna manera, ya que las categorías gramaticales no se corresponden con las ontológicas, de mayor complejidad (Viegas et al., 1999).

Otro aspecto interesante que hay que destacar es la multidimensionalidad de este sistema de categorización, en los que los conceptos pueden ser ordenados según una o varias variables (Montero Martínez et al., 2011). Esto permite que un mismo término se clasifique de distinta forma, según el sistema conceptual utilizado (**Figura 4**) (Bowker, 1997). Este tipo de clasificación se caracteriza por su gran dinamismo, por lo que es, además, muy útil para analizar lenguajes especializados, que se encuentran en constante cambio.

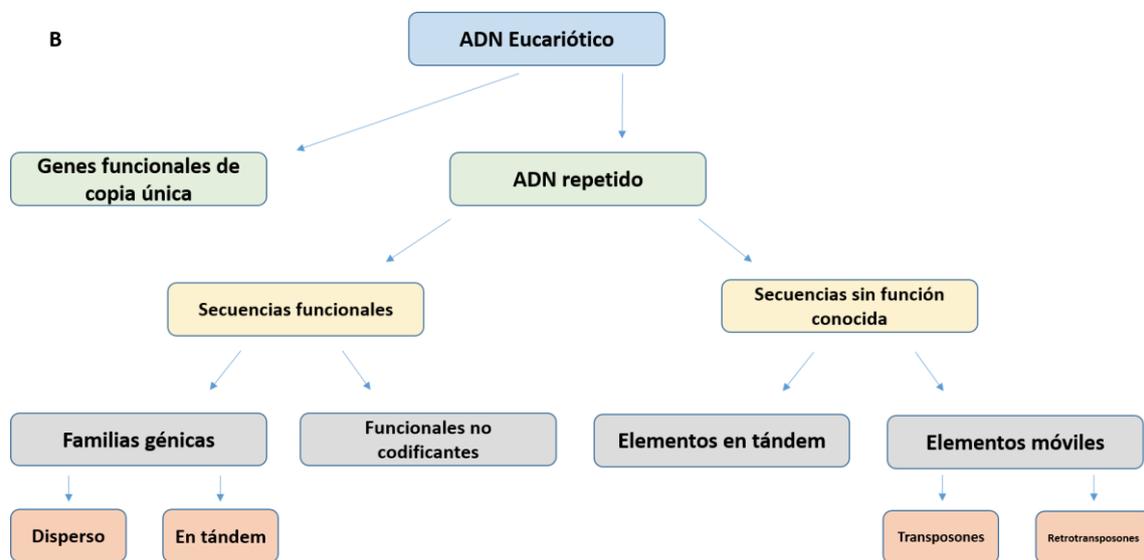
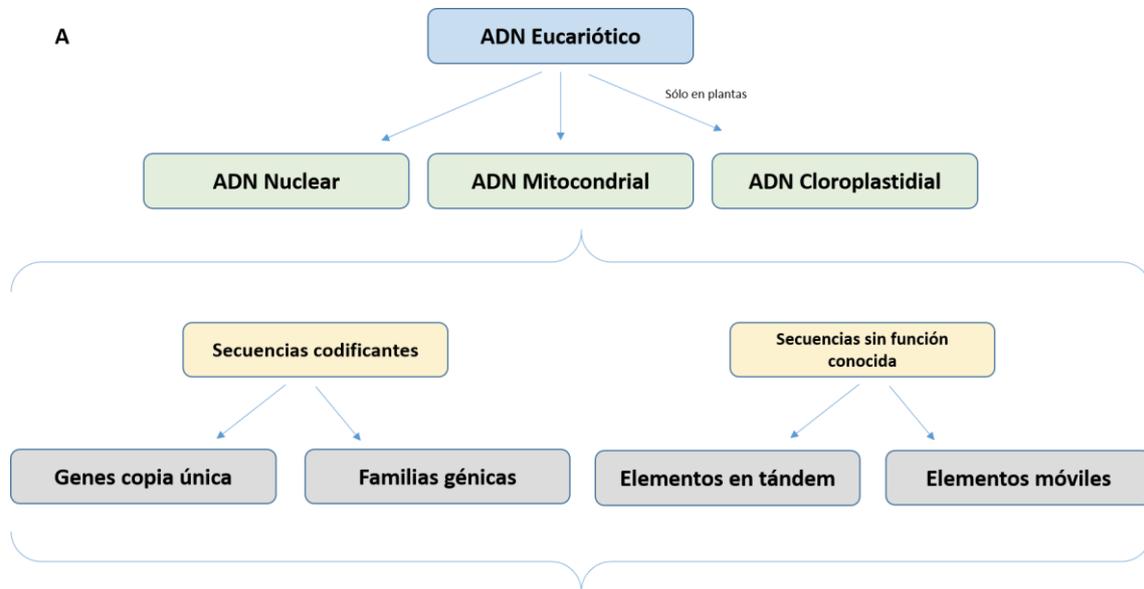


Figura 4.- Mapas conceptuales monodimensionales de la clasificación del ADN utilizando como criterio fundamental (A) su localización o (B) su función.

2.- Objetivos

2.- Objetivos

- (1)** Crear tres *corpora* de textos paralelos inglés-español cuya temática principal sea la Genética, en contexto educativo, divulgativo y científico.
- (2)** Mediante un estudio de frecuencia de uso, relaciones semánticas, variantes terminológicas y concordancias, analizar la ocurrencia de una lista de términos (divididos en entidades, procesos y atributos) en dichos *corpora*.
- (3)** Analizar las traducciones más frecuentes según el contexto y analizar su grado de validez y corrección.

3.- Materiales y métodos

3.- Materiales y métodos

3.1.- Delimitación temática de los textos analizados y creación de corpora

Los *corpora* utilizados como material para la ejecución del presente estudio, fueron elaborados a partir de textos paralelos pertenecientes a los tres contextos mencionados con anterioridad: divulgativo, educativo y de científico.

Para la elaboración del *corpus* divulgativo se utilizaron textos pertenecientes a los números publicados entre los años 2000 y 2002 de la revista mensual *Scientific American* (Nature Publishing Group) y de su homóloga en español *Investigación y Ciencia* (Prensa Científica, S.A.). Dado que se trata de una publicación de carácter generalista, fue necesario llevar a cabo una selección manual de los artículos, eligiendo aquéllos cuya temática estuviera relacionada con la Genética. Además, la correspondencia de los números de la publicación en inglés y en español no es exacta, ni en el contenido ni en el tiempo. También se usaron todos los artículos con la etiqueta "Genética" de la plataforma OpenMind (<https://www.bbvaopenmind.com>), que publica reseñas de investigadores sobre temas de actualidad en inglés y español. Así, un total de 118 textos paralelos fueron analizados (389 701 palabras totales -201 029 y 188 672 en español e inglés, respectivamente) (**Anexo I**).

Para la elaboración del *corpus* educativo se utilizaron 10 capítulos íntegros del libro de texto *Genetics: a conceptual approach, 2nd Edition* (Editorial Panamericana) y su traducción al español, *Genética: un enfoque conceptual, 2^a Edición*. Este manual es usado como material de apoyo para la docencia de distintas asignaturas impartidas en el Grado en Biología por el Departamento de Genética de la Facultad de Ciencias de la Universidad de Granada. Un total de 20 textos paralelos fueron obtenidos (245 729 palabras totales -133 374 y 112 355 en español e inglés, respectivamente) (**Anexo II**).

Para la elaboración del *corpus* científico se utilizaron textos paralelos obtenidos del repositorio de tesis doctorales contenido en la base de datos DIGIBUG de la Universidad de Granada. Se hizo un filtrado por temática, usando como palabras clave Genética, Genómica, Gen, y posteriormente se seleccionaron aquellos documentos que contuvieran textos paralelos. Además, se usaron artículos relacionados con la Genética de la Revista Española de Cardiología, que publica sus artículos en inglés y español. De esta forma, se obtuvo un total de 79 documentos (218 049 palabras totales -08 032 y 110 017 en español e inglés, respectivamente) (**Anexo III**).

Todos los textos se inspeccionaron manualmente, se convirtieron a texto plano y se eliminaron los espacios en blanco innecesarios utilizando el programa Notepad++ (<https://notepad-plus-plus.org/>). Todos los textos están disponibles en la dirección web <http://www.rafaelnavajas.eu/TFG>.

3.2.- Minería de textos

Las listas de palabras, la frecuencia de uso, la comparación de términos y las concordancias fueron generadas utilizando el programa SketchEngine (www.sketchengine.co.uk). Para la minería de textos no pertenecientes a los *corpora* propios, se usó el Corpus del Español Actual (CREA).

4.- Resultados y Discusión

4.- Resultados y Discusión

4.1.- Criterios de selección de términos y textos paralelos

Tomando como referencia los *corpora* de textos en inglés y usando el programa Sketch Engine se obtuvo una lista con los términos cuya frecuencia de uso era mayor. Para ello, fue necesario quitar artículos, pronombres, determinantes y palabras no específicas del campo (**Tabla 1**). Como era esperable, se encontraron algunas diferencias cuando los *corpora* fueron tratados por separado. Así, en el *corpus* educativo encontramos una alta frecuencia de términos de Genética básica (*chromosome, gene, cell, locus, phenotype*), mientras que en los *corpora* divulgativo y científico predominaban términos relacionados con la salud humana y los nuevos retos de la Genética (*cancer, tumor, patient, cell, mutation*).

Atendiendo a los criterios de clasificación utilizados por Sager (1900) y Montero Martínez et al., (2011) descritos en el apartado 1.4 de la Introducción, y considerando los términos de esta lista (**Tabla 1**) un total de 17 unidades terminológicas han sido analizadas en este trabajo. La **Tabla 2** muestra las entidades seleccionadas y las variantes detectadas en español, la **Tabla 3** muestra los procesos y sus respectivos en español, mientras que la **Tabla 4** hace lo mismo con los atributos.

Término	Frecuencia	Término	Frecuencia
gene/genetic(s)/genome	3432	cardiovascular	155
chromosome	1347	heterozygous	155
protein	1309	virus	155
cell	968	recessive	154
mutation	921	research	154
disease	857	development	151
DNA	768	much	151
sequence	562	tissue	149
allele	558	bacterial	147
factor	525	Mendel	144
phenotype	468	principle	144
cancer	457	evolution	142
plant	379	muscle	142
bacterium	350	crose	140
locus	310	fiber	139
progeny	303	inherit	136
cardiac	291	child	132
male	286	copy	132
drug	285	dominant	132
molecular	275	pattern	132
molecule	268	crossover	127
animal	266	immune	123
specie	249	mitochondrial	122
tumor	244	diagnosis	119
researcher	232	offspring	118
genotype	223	antibody	117
gamete	222	species	116
sex	216	enzyme	115
scientist	215	biology	114
map	192	death	114
probability	192	frequency	114
meiosis	191	homologous	114
syndrome	184	encode	110
polymorphism	173	chromosome	109
region	172	fragment	105
variant	171	function	104
vaccine	170	homozygous	101
recombination	167	myocardium	101
inheritance	157	treatment	101
pair	157	clon	100

Tabla 1.- Términos con mayor frecuencia de uso en los *corpora* en inglés analizados.

<i>Término en inglés</i>	<i>Variantes en español</i>
DNA (y derivados mtDNA, cpRNA,...)	ADN/DNA, mtDNA/ADNmt/mtADN,...
RNA (y derivados mRNA, tRNA,...)	ARN/RNA, mRNA/ARNm/mRNA,...
Genetist	Genético/a, Genetista
Genetic Code	Código genético/ADN
Contig	Cóntigo/contig
EST (Expressed sequence tags)	ESTs, librería/genoteca de ESTs, secuencias ESTs

Tabla 2.- Términos en inglés y las variantes detectadas en español correspondientes a entidades.

<i>Término en inglés</i>	<i>Variantes en español</i>
Crossing-over	Sobrecruzamiento/entrecruzamiento, recombinación
Genomic imprinting	Impronta genética, imprinting, marcaje genómico
PCR	la/el PCR
Deletion	Delección/delecionar, delección (*)
(Alternative) Splicing	Splicing/Empalme, corte y empalme (alternativo)
DNA fingerprinting	Huella (dactilar) genética, perfil genético

Tabla 3.- Términos en inglés y las variantes detectadas en español correspondientes a procesos.

<i>Término en inglés</i>	<i>Variantes en español</i>
Heterozygous	Heterocigoto/a, Heterozigoto/a, Heterocigótico
Genetic/Genomic	Uso intercambiable de genético/genómico
Ribosomal	Ribosómico/ribosomal
Chromosomal	Cromosómico/cromosomal
Viral	Vírico/viral

Tabla 4.- Términos en inglés y las variantes detectadas en español correspondientes a atributos.

4.2.- Análisis de problemas

Los términos que aparecen en las tablas 2-4 fueron analizados en los distintos contextos. Los problemas que se han encontrado en relación a los mismos, se pueden dividir en los siguientes apartados:

4.2.1.- Siglas

ADN/DNA

Una sigla es un «signo lingüístico formado generalmente con las letras iniciales de cada uno de los términos que integran una expresión compleja» (RAE, 2010). En ocasiones, incluyen letras iniciales o intercaladas que facilitan su pronunciación, en cuyo caso pasan a denominarse

acrónimos. Los acrónimos con más de cuatro letras que se usan para designar nombres propios pueden opcionalmente llevar tilde y escribirse sólo la primera letra en mayúscula. Por el contrario, las siglas se escriben en mayúsculas, sin cursiva, sin puntos, ni entrecomillado y no deben hacer el plural con una 's' final (ni mayúscula ni minúscula) (Fundéu, 2011; 2013). El núcleo del desarrollo de la sigla es el que determina el género de la misma.

Según el *Diccionario Panhispánico de Dudas*, en líneas generales es recomendable traducir las siglas extranjeras, salvo cuando su uso esté muy asentado o se trate de nombres comerciales, en cuyo caso es preferible mantener la grafía original. Si se trata de una sigla no muy común, se recomienda indicar su traducción y su desarrollo, en minúscula si se trata de nombres comunes o en mayúscula en el caso de los nombres propios (Fundéu, 2011; 2013).

El uso de siglas y abreviaturas en el lenguaje científico es muy frecuente. No de manera inesperada, ya que se trata de la molécula central de la disciplina que nos ocupa, la sigla con mayor frecuencia de uso en los textos analizados en este trabajo fue 'DNA' (deoxyribonucleic acid) (**Tabla 1**).

A pesar de que el *Diccionario de la RAE* acepta la forma inglesa de esta sigla ('DNA'):

DNA

Del ingl. DNA, sigla de deoxyribonucleic acid 'ácido desoxirribonucleico'.

1. m. Biol. ADN.

y la adaptación española ('ADN')

ADN

Sigla de ácido desoxirribonucleico.

1. m. Biol. ácido desoxirribonucleico.

La voz detectada con mucha mayor frecuencia en los textos en español analizados es 'ADN'. Significativamente, los textos de contexto educativo optaron por la voz 'DNA', los textos científicos por la voz 'ADN', mientras que los divulgativos usaron ambas (**Tabla 5**).

	Divulgativo	Educativo	Científico
ADN	339/1 489,44	0/0	107/840,32
DNA	5/21,97	493/3 249,66	0/0

Tabla 5.- Frecuencia de aparición de ADN/DNA en los *corpora* analizados. **Nota:** valores expresados en número de coincidencias/número de coincidencias por millón de palabras.

Además de la forma convencional de esta sigla, en inglés existen diversas variantes que hacen referencia a distintos tipos de moléculas de ADN; esto es, 'cDNA' para *complementary DNA*, 'mtDNA' para *mitochondrial DNA*, 'cpDNA' para *chloroplastidial DNA*. En los *corpora* de textos en español analizados en este trabajo, se han encontrado las siguientes variantes: 'cADN', 'ADNc', 'ADN-c', 'ADNmt', en el *corpus* científico; 'ADNmt', 'ADNc', en el *corpus* divulgativo; 'cDNA', 'mtDNA', 'cpDNA', en el *corpus* educativo. Los textos divulgativos y los de ámbito educativo son coherentes, ya que, cuando aparece una variante, adaptan completamente la sigla al español, los primeros ('ADNmt' por 'mtDNA') o mantienen sin modificar la sigla inglesa ('cDNA', por ejemplo). En cambio, en los textos científicos observamos variantes que mezclan ambos criterios ('cADN'* en lugar de 'ADNc' o 'cDNA'). Por último, se ha encontrado en contexto científico y divulgativo, las formas plurales 'ADNs', 'DNAs' desaconsejadas en textos escritos como hemos visto anteriormente.

ARN/RNA

De forma análoga, se ha estudiado el comportamiento de la voz 'RNA' (*ribonucleic acid*), que es una molécula también central en Genética, y cuyas formas española e inglesa son aceptadas por la RAE. Al contrario de lo que ocurría con ADN/RNA, en este caso, la voz más frecuente fue 'RNA' (Tabla 6). Es interesante resaltar que el tándem ARN/RNA repite el patrón observado anteriormente para el par ADN/DNA, esto es, los textos de contexto educativo optaron por la voz 'RNA', los textos científicos por la voz 'ARN', mientras que los divulgativos usaron ambas (Tabla 6).

	Divulgativo	Educativo	Científico
ARN	70/307,55	0/0	32/251,31
RNA	11/48,33	66/435,05	0/0

Tabla 6.- Frecuencia de aparición de ARN/RNA en los *corpora* analizados. **Nota:** valores expresados en número de coincidencias/número de coincidencias por millón de palabras.

Además de la forma convencional de esta sigla, en inglés existen distintas variantes que hacen referencia a distintos tipos de moléculas de 'ARN'; esto es: 'mRNA' para *messenger RNA*, 'tRNA' para *transfer RNA*, 'rRNA' para *ribosomal RNA*, 'RNAi' para *RNA interference*, 'RNAase' para *ribonuclease*, 'siRNA' para *small interfering RNA*. En los *corpora* de textos en español analizados en este trabajo se han encontrado las siguientes variantes: 'ARNm', 'micro-ARN', 'siARN', 'aminoacil-t-ARN', en el *corpus* científico; 'rRNA', 'RNAi', 'ARNr', 'ARNm', 'ARNt', 'ARNasa' en el *corpus* divulgativo; 'mRNA', en el *corpus* educativo. Los textos divulgativos son nuevamente coherentes ya que adaptan completamente la sigla al español ('ARNr' por 'rRNA') o mantienen sin modificar la sigla inglesa ('rRNA', por ejemplo) y los textos del *corpus* educativo mantiene las siglas en inglés sin modificar ('mRNA'). En los textos científicos, aparecen algunas variantes parcialmente adaptadas al español como 'siARN'. La forma más correcta de esta variante sería 'ARNip', si bien aún está en proceso de aceptación por parte de la propia comunidad científica (10 300 resultados en Google frente a los de 71 800 de 'siARN'). También se aprecia una adaptación parcial en 'aminoacil-t-ARN' (en lugar de 'aminoacil ARNt'). Quizás este gran número de variantes de moléculas de ARN y la dificultad de encontrar una nomenclatura adaptada ampliamente aceptada ha propiciado el predominio de la forma 'RNA' con respecto a la forma 'ARN' en los *corpora* de lengua española. Por último, se han encontrado las formas plurales 'ARNs', 'RNAs'.

PCR/RCP

A este respecto, es interesante el caso de la sigla 'PCR' (*Polymerase Chain Reaction*). En los *corpora* analizados, su uso está completamente extendido, salvo la ocurrencia puntual de la sigla adaptada al español 'RCP' (**Tabla 7**). En cardiología, y de forma general en medicina, la sigla 'RCP' está completamente asentada y se destina a describir el término *reanimación cardiopulmonar*. Probablemente, éste sea uno de los motivos por los que la sigla 'PCR' ha permanecido invariable a pesar de lo extendido de su uso.

	Divulgativo	Educativo	Científico
PCR	1/4,39	2/13,18	28/219,90
RCP	2/8,79	0/0	0/0

Tabla 7.- Frecuencia de aparición de las distintas variantes de PCR en los *corpora* analizados. **Nota:** valores expresados en número de coincidencias/número de coincidencias por millón de palabras.

Asimismo, se ha detectado la presencia de las variantes QRT-PCR y RT-PCR en el *corpus* de textos científicos, correspondientes a los términos *real-time quantitative reverse transcription PCR* y *reverse transcription PCR*, respectivamente.

También, es necesario hacer una consideración acerca del género de las siglas 'PCR'. Atendiendo a los textos normativos, 'PCR', cuyo núcleo es femenino *reacción*, debe ir acompañada por artículos y determinantes de este género. A pesar de que no hemos encontrado incorrecciones a este respecto en los textos analizados, hay que hacer notar que es frecuente encontrar textos en los que esta sigla es tratada erróneamente como un sustantivo masculino, como en el ejemplo que sigue:

[Anderson y Relman usando técnicas de ampliación genómica usando **el PCR*** detectan DNA de la Bartonella en la biopsia del tejido.]

Fuente: Acta médica peruana, Vol. XVIII, nº 3, 09-12/2001

Por último, las tablas 8 y 9 muestran algunos de los contextos en los que se ha encontrado la voz ADN/DNA. Es destacable el hecho de que el español presente una mayor riqueza en verbos que usan este término como sujeto (**Tabla 8**). Así, para 'have' encontramos las equivalencias: *tener, contener, captar, deber, poseer, haber*; para 'encode': *sintetizar, duplicar, copiar, encadenar*; para 'indicate': *significar, evidenciar, detectar, leer*; para 'create': *producir, obtener, aislar, secuenciar, incorporar, inyectar, liberar, integrar, enfilear*, y para 'play': *jugar, insertar, cortar, extraer, agregar, alargar, utilizar, envolver, abrir, donar, intercambiar*. Mención aparte merece el verbo 'to be', sustituido por: *ser, constituir, aparecer, poder, llamarse, especializar, manipular, separar, modificar, transportar, quedar, encontrar, incluir, unir*, y cuya preponderancia numérica puede deberse a su polisemia, a su intervención en la formación de tiempos compuestos y de la voz pasiva, de mayor uso en los textos en lengua inglesa, como puede verse en el siguiente ejemplo:

[...DNA is replicated in the S stage...]

[...en la fase S tiene lugar la síntesis de DNA...]

Con respecto a los verbos que usan ADN/DNA como objeto, se observa un número mayor de verbos ingleses que de españoles, no sólo en cantidad sino también en números absolutos (**Tabla 9**). De esta forma, aunque existen coincidencias exactas de algunos verbos: ‘ser’: *be*, ‘detectar’: *detect*, ‘extraer’: *extract*, ‘obtener’: *obtain*, ‘perder’: *eliminate*, ‘crear’: *replicate*, ‘estar’: *locate*, existen equivalentes inglesas que no se encuentran en los *corpora* españoles: *splice*, *inject*, *hybridise*, *transfer*, *degrade*, *isolate*, etc.

	CIENTÍFICO	EDUCATIVO	DIVULGATIVO	TOTAL
SUJETO				
captar	0	6	1	7
ser	6	0	0	6
duplicar	0	3	2	5
contener	2	2	0	4
insertar	0	2	2	4
poder	3	0	0	3
aislar	0	2	1	3
cortar	0	0	3	3
extraer	0	2	1	3
agregar	0	0	2	2
constituir	0	0	2	2
detectar	0	1	1	2
haber	0	2	0	2
leer	0	1	1	2
obtener	0	0	2	2
sintetizar	0	2	0	2
alargar	1	0	0	1
aparecer	1	0	0	1
deber	1	0	0	1
poseer	1	0	0	1
quedar	1	0	0	1
abrir	0	1	0	1
copiar	0	0	1	1
donar	0	1	0	1
encadenar	0	0	1	1
encontrar	0	1	0	1
enfilarse	0	0	1	1
envolver	0	0	1	1
especializar	0	1		1
evidenciar	0	0	1	1
incluir	0	1	0	1
incorporar	0	0	1	1
integrar	0	1	0	1
intercambiar	0	0	1	1

inyectar	0	0	1	1
jugar	0	0	1	1
liberar	0	1	0	1
llamar+se	0	0	1	1
manipular	0	0	1	1
modificar	0	0	1	1
producir	0	1	0	1
secuenciar	0	1	0	1
separar	0	1	0	1
significar	0	0	1	1
tener	0	1	0	1
transportar	0	1	0	1
unir	0	0	1	1
utilizar	0	1	0	1
OBJETO				
ser	0	0	5	5
detectar	3	0	0	3
poder	0	0	2	2
alterar	0	0	1	1
crear	0	1	0	1
deber	0	1	0	1
desnudar	0	0	1	1
estar	0	0	1	1
extraer	1	0	0	1
formar	0	0	1	1
obtener	1	0	0	1
perder	0	0	1	1
permanecer	0	0	1	1
persistir	1	0	0	1
sobrevivir	0	0	1	1

Tabla 8.- Verbos que utilizan ADN/DNA como sujeto o como objeto en los *corpora* de textos en español. En número se representan las coincidencias encontradas.

	CIENTÍFICO	EDUCATIVO	DIVULGATIVO	TOTAL
SUJETO				
be	9	38	21	68
encode	0	4	2	6
have	0	0	2	2
indicate	0	2	0	2
play	0	0	2	2
create	1	0	0	1
OBJETO				
cut	0	6	3	9
transfer	0	8	0	8
be	0	4	2	6
obtain	4	0	2	6
replicate	0	4	2	6
analyze	0	3	2	5
synthesize	0	5	0	5
contain	0	2	2	4
extract	4	0	0	4
carry	0	3	0	3
copy	0	0	3	3
detect	3	0	0	3
have	0	3	0	3
inject	0	0	3	3
sequence	0	3	0	3
splice	0	0	3	3
complexed	0	2	0	2
degrade	0	2	0	2
isolate	0	2	0	2
make	0	2	0	2
produce	0	2	0	2
use	0	0	2	2
eliminate	1	0	0	1
hybridise	1	0	0	1
locate	1	0	0	1

Tabla 9.- Verbos que utilizan DNA como sujeto o como objeto en los *corpora* de textos en inglés. En número se representan las coincidencias encontradas.

4.2.2.- Extranjerismos

El proceso de adquisición de nuevos términos procedentes de otra lengua se conoce como préstamo lingüístico o adquisición de extranjerismos y es uno de los mecanismos principales de la ampliación del léxico de una lengua (RAE, 2010). Existen varios motivos por los que puede producirse este fenómeno; así, se puede adoptar un extranjerismo por la influencia de la lengua extranjera, en términos de prestigio, estilo o moda o cuando en la lengua propia no se dispone de un término que designe la realidad que aquél representa o se refiere a referentes de reciente aparición para los que no se cuenta con términos propios. Este último sería el mecanismo más frecuente por el que se adoptan extranjerismos en el lenguaje científico.

Como ya se ha comentado a lo largo de esta memoria, el inglés es usado como *lingua franca* en ciencia. Por ello, es frecuente encontrar términos que surgen en inglés para designar técnicas, elementos o conceptos que el rápido desarrollo de las disciplinas pone en juego sin que dé tiempo a que la lengua los asimile o cree términos específicos para designarlos.

En los *corpora* analizados en este trabajo, se han detectado tres ejemplos de extranjerismos particularmente interesantes: *splicing*, *contig* y *delección*.

SPLICING

A continuación se desarrolla el caso de la palabra inglesa 'splicing' (Figura 5).

<p>USE TO: splicing</p> <p>Contexto educativo: 3 casos, 22,79 por millón de palabras.</p> <p>Contexto divulgativo: 2 casos, 9,27 por millón de palabras.</p> <p>Contexto científico: 16 casos, 117,45 por millón de palabras.</p>	<p>DEFINICIÓN: (Science: molecular biology) The process by which introns are removed from hnRNA to produce mature messenger RNA that contains only exons. Alternative splicing seems to occur in many proteins and by alternative exon usage a set of related proteins can be generated from one gene, often in a tissue or developmental stage specific manner.</p>
<p>FUENTE: Dictionary of Biology Online</p> <p>http://www.biology-online.org/dictionary/</p>	<p>SEGMENTO DE ACTIVACIÓN:</p> <p>[...immature RNA transcript by a process called splicing, resulting in a molecule of messenger RNA...]</p> <p>[...the alteration of normal mRNA splicing patterns, the possibility that...]</p> <p>[This would represent a splicing mutation, and the patient's 30-year-old...]</p> <p>[Alternative splicing of the CLA1 gene produces 2 forms...]</p>
<p>RELACIONES:</p> <p>[ES_UN]: proceso</p> <p>[AFECTA_A]: mRNA</p> <p>[SUSCEPTIBLE_DE]: mutación</p> <p>ATRIBUTOS:</p> <p>[PROPIEDAD]: alternativo</p>	

Figura 5.- (página anterior) Esquema categorial para el término 'splicing'.

En la tabla 10 pueden encontrarse las distintas equivalencias halladas para 'splicing' en los corpora analizados en este trabajo. Este término, probablemente por ser muy especializado, no se encuentra presente en el corpus de textos divulgativos. En el corpus educativo, sólo existe el equivalente en lengua española 'corte y empalme', que probablemente sea el más extendido (de acuerdo con CREA). En el corpus de textos científicos aparece esta variante, aunque es desplazada en frecuencia por un término español de registro más elevado 'ayuste' o por el extranjerismo sin adaptación 'splicing'.

	Divulgativo	Educativo	Científico
Splicing	0/0	0/0	4/31,41
Corte y empalme	0/0	4/26,37	2/15,71
Ayuste	0/0	0/0	6/47,12

Tabla 10.- Frecuencia de aparición de las distintas variantes de 'splicing' en los corpora analizados. **Nota:** valores expresados en número de coincidencias/número de coincidencias por millón de palabras.

El sonido 's' seguido de consonante al inicio de palabra no existe en español, por lo que se observa la tendencia a adaptar este término mediante la inclusión de una 'e' inicial, 'esplicing'. Este término aún no se encuentra aceptado, en cambio 'espliceosoma' es la variante más extendida para referirse a la palabra inglesa "spliceosome".

espliceosoma

1. f. Complejo de ribonucleoproteínas que se forma durante los procesos de corte y empalme del ácido ribonucleico mensajero para eliminar los intrones que no van a ser traducidos a proteínas.

Fuente: Diccionario Médico de la Clínica Universidad de Navarra

CONTIG

Como se ha mencionado anteriormente, otro de los extranjerismos detectados es 'contig'

(Figura 6).

<p>USE TO: contig</p> <p>Contexto educativo: 7 casos, 53,18 por millón de palabras.</p> <p>Contexto divulgativo: 0 casos.</p> <p>Contexto científico: 0 casos.</p>	<p>DEFINICIÓN: Contiguous sequence of DNA created by assembling overlapping sequenced fragments of a chromosome. A group of clones representing overlapping regions of the genome.</p>
<p>FUENTE: Medical Dictionary Online</p> <p>http://www.medicinenet.com</p>	
<p>RELACIONES:</p> <p>[ES_UN]: fragmento de ADN</p> <p>[USED_TO]: construir cromosomas</p> <p>ATRIBUTOS:</p> <p>[PROPIEDAD]: solapante</p>	<p>SEGMENTO DE ACTIVACIÓN:</p> <p>[...a contiguous stretch of DNA is called a contig.]</p> <p>[This approach was used in 1993 to create a contig of the human Y chromosome consisting...]</p> <p>[...by putting together the sequences of all overlapping contigs...]</p>

Figura 6.- Esquema categorial para el término 'contig'.

Se ha detectado su presencia únicamente en el corpus educativo (3 coincidencias, 19,77 por millón de palabras) donde aparece manteniendo la forma inglesa "contig", por lo que se trataría de otro préstamo no adaptado. Sin embargo, empieza a surgir la variante 'cántigo':

cántigo (contig)

Un cántigo -de la palabra contiguo- es una serie de secuencias superpuestas de ADN utilizadas para hacer un mapa físico que reconstruye la secuencia original de ADN de un cromosoma o de una región de un cromosoma. Un cántigo también puede referirse a una de las secuencias de ADN utilizadas en la fabricación de dicho mapa.

Fuente: Glosario hablado de términos genéticos (NIH).

DELECIÓN

Por último, he detectado la presencia del término 'delección' y sus variantes, cuyo uso está muy extendido en Genética (**Figura 7**). Aunque etimológicamente la palabra provenga del latín (*deletio*), parece haber entrado en nuestro diccionario a través del inglés (*deletion*), con la única acepción para la que fue acuñada, como se puede deducir de la definición del Diccionario de la RAE:

delección

Del ingl. *deletion*, y este del lat. *deletio*, -ōnis 'destrucción'.

1. f. Biol. Mutación genética que consiste en la pérdida de uno o más nucleótidos de la secuencia del ADN.

<p>USE TO: deletion</p> <p>Contexto educativo: 60 casos, 455,79 por millón de palabras.</p> <p>Contexto divulgativo: 7 casos, 32,46 por millón de palabras.</p> <p>Contexto científico: 19 casos, 139,47 por millón de palabras.</p>	<p>DEFINICIÓN: (Science: genetics) a chromosome abnormality in which part of a single chromosome has been lost.</p>
<p>FUENTE: Dictionary of Biology Online</p> <p>http://www.biology-online.org/dictionary/</p>	
<p>RELACIONES:</p> <p>[ES_UNA]: mutación</p> <p>[AFECTA_A]: DNA sequence</p> <p>[DA_LUGAR_A]: proteína mutada</p> <p>[TIPO_DE]: reordenación cromosómica</p> <p>[USADO_EN]: mapeo</p> <p>ATRIBUTOS:</p> <p>[PROPIEDAD]: letal</p>	<p>SEGMENTO DE ACTIVACIÓN:</p> <p><i>[...a mutation (deletion of 21 bases) was identified in gene MEF2A...]</i></p> <p><i>[...can occur in the DNA sequence; these include large and small insertions and deletions...]</i></p> <p><i>[When a small deletion or insertion occurs in a coding region, the result is a mutated protein.]</i></p> <p><i>[The four basic types of rearrangements are duplications, deletions, inversions, and translocations]</i></p> <p><i>[...that play a role in deletion mapping and somatic-cell hybridization...]</i></p> <p><i>[The deletion acts as a recessive lethal because loss of all copies...]</i></p>

Figura 7.- Esquema categorial para el término 'deletion'.

Esta voz está muy asentada en el campo de la Biología en general, y en la Genética en particular, y se utiliza para describir uno de los procesos más frecuentes de mutación cromosómica que ocurren en la naturaleza y las consecuencias que provoca (ver **Figura 8**).

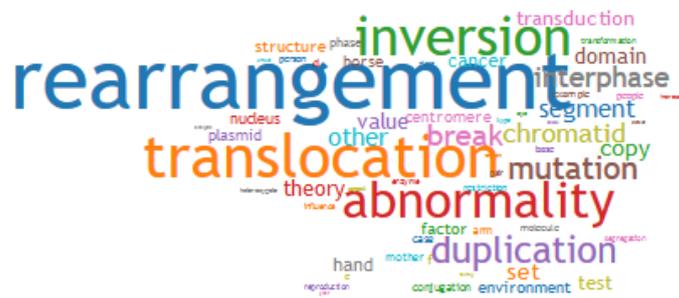


Figura 8.- Nube de palabras construida para el término ‘deletion’ a partir de las coincidencias encontradas en el corpus educativo.

Se trataría, por tanto, de un extranjerismo adaptado. En los corpora analizados, las únicas formas detectadas han sido las derivadas del lema ‘delección’ (Tabla 11).

	Divulgativo	Educativo	Científico
Delección	5/21,97	66/435,05	18/141,36

Tabla 11.- Frecuencia de aparición de las distintas variantes de ‘deletion’ en los corpora analizados. **Nota:** valores expresados en número de coincidencias/número de coincidencias por millón de palabras.

Significativamente, en el corpus del contexto educativo se encuentra la forma verbal (‘delecionar’) derivada de la forma nominal (‘delección’), lo que podría ser un indicativo del grado de inserción del término en nuestra lengua.

*En primer lugar, en el transcurso de la mitosis o de la meiosis puede perderse un cromosoma si, por ejemplo, su centrómero se **deleciona**.*

Asimismo, se detecta en la literatura la presencia incorrecta ‘delección’* que bien podría corresponderse con una adaptación de tipo ortográfico característica del idioma español.

*La **delección*** de tales genes [...] provoca la ruptura de la estructura cromatínica no nucleosómica...*

Fuente: ABC Cultural, 06/12/1996: MARCADORES TUMORALES

A modo de conclusión, podríamos indicar que mediante procesos muy diversos, ya sea por iniciativa de los hablantes o de las instituciones de normalización, se observa que los extranjerismos incorporados a la lengua española tienden a acomodarse a los patrones característicos de ésta. El resultado es una adaptación de acuerdo con la pronunciación o las normas ortográficas del español, incluyendo las flexiones nominales de género y número o la adopción de las distintas formas verbales de nuestro idioma (adaptado de RAE, 2010). Cuando esto sucede, se puede hablar de un extranjerismo adaptado, en contraposición a los extranjerismos crudos o no adaptados.

El aumento de la alfabetización en otros idiomas (predominantemente el inglés) y la globalización de la información han propiciado el aumento del número de voces extranjeras que se insertan en nuestra lengua, pero también ha contribuido a una mayor familiaridad por parte de los hablantes con grafías que no le son propias. Todo ello, ralentiza frecuentemente el proceso de acomodación e integración de los préstamos.

4.2.3.- Uso de adjetivos

CROMOSÓMICO/CROMOSOMAL, RIBOSÓMICO/RIBOSOMAL, VÍRICO/VIRAL

Alguno de los adjetivos con sufijos *-ico* y *-al* son considerados adjetivos de relación (en contraposición a los adjetivos calificativos), que como su propio nombre indica, expresan un relación entre el adjetivo y el sustantivo al que modifican (de Kock, 1994). Suelen ser de naturaleza clasificativa y se derivan de bases nominales (*-ico*) o se asocian semánticamente a ellas (*-al*, relación de pertenencia).

Por lo general, se pueden organizar en grupos de adjetivos caracterizados semánticamente en lugar de por oposiciones antónimas (ribosómico en contraposición a nuclear, por ejemplo). Además, no son graduables y suelen ir pospuestos al sustantivo.

[Las infecciones **virales** favorecen la formación de tumores al incorporar...]

Aunque existe un gran número de adjetivos españoles terminados en *-al* (ambiental, artificial, boreal, conceptual, estatal, forestal, vital, etc.), esta terminación no es tan frecuente para la formación de adjetivos especializados como en inglés. Los adjetivos ingleses terminados en *-al*

usan en español el sufijo *-ico* (chromosomal, ‘cromosómico’), *-iano* (bacterial, ‘bacteriano’), *-ario* (embryonal, ‘embrionario’), *-o* (external, ‘externo’), *-oso* (parenchymal, ‘parenquimatoso’), *-ar* (vulval, ‘vulvar’), *-ífero* (mammal, ‘mamífero’), entre otros (Navarro, 2008).

En los *corpora* analizados, se detecta el uso de las formas de los adjetivos en *-ico*, ‘cromosómico’, ‘ribosómico’ y ‘vírico’ alternado con las formas en *-al*, ‘cromosomal’, ‘ribosomal’ y ‘viral’ (**Tabla 12**). Las formas ‘cromosomal’ y ‘ribosomal’ no están aceptadas por la RAE, mientras que la forma ‘viral’ sí:

viral

1. adj. Perteneciente o relativo a los virus.

	Divulgativo	Educativo	Científico
Cromosómico	5/21,97	163/1 074,43	8/62,83
Cromosomal	0/0	0/0	2/15,71
Ribosómico	8/35,15	0/0	2/15,71
Ribosomal	0/0	0/0	5/39,27
Vírico	66/289,98	0/0	0/0
Viral	7/30,76	27/177,97	1/7,85

Tabla 12.- Frecuencia de aparición de las distintas variantes de ‘cromosómico/cromosomal’, ‘ribosómico/ribosomal’, ‘vírico/viral’ en los *corpora* analizados. **Nota:** valores expresados en número de coincidencias/número de coincidencias por millón de palabras.

La *Real Academia Nacional de Medicina de España* también señala que, aunque se trate de un anglicismo, ‘viral’ es una voz cuyo uso está completamente asentado en la práctica. Por lo tanto, la alternancia de los términos ‘vírico’ y ‘viral’ como sinónimos es correcta. Probablemente, a la consolidación de ‘viral’ haya contribuido su uso muy extendido en el ámbito de las redes sociales e internet, donde esta voz es usada como adjetivo y como sustantivo para referirse a un mensaje transmitido en distintas formas (memes, imágenes, eslóganes) de forma exponencial, valiéndose de la interacción entre usuarios (Fundéu, 2014).

GÉNICO/GENÉTICO/GENÓMICO

Los términos ‘génico’, ‘genético’ y ‘genómico’ no son sinónimos: ‘génico’ se aplica a todo lo perteneciente o relativo a los genes, ‘genético’ se usa para hacer referencia a lo relativo o

perteneciente a la Genética, y 'genómico' para referirse al genoma (Real Academia Nacional de Medicina, 2012; RAE, 2014). No obstante lo dicho, es frecuente encontrar el uso alternativo de este grupo de palabras, en muchas ocasiones de forma confusa o equivocada.

[*La cuasiespecie viral como tal es el sujeto de la selección y por tanto de la evolución vital, y no cada molécula **genómica*** en particular.*]

Fuente: Lázaro Lázaro E, Escarmís Homs C (2002) Virus emergentes. La amenaza oculta.

[*...de secuencias repetitivas de ADN, de gen contiguo-microdelección, de imprinting **genómico***, y los cuadros de etiología génica...*]

Fuente: Boletín del ECEMC: Revista de Dismorfología y Epidemiología, Serie V, nº 1, 2002

Cuando se analizan las palabras que modifican estos términos, comparativamente dos a dos (ver **Figura 9**), se observa que en los textos analizados existe una clara diferenciación entre 'génico/genético' y 'genómico'. Mientras que 'genómico' modifica a 'ADN', el término 'génico' modifica a *expresión, terapia, polimorfismo, regulación*, y 'genético' a *prueba, variante, información o estudio*, por citar sólo algunos ejemplos (**Figura 9**). Adicionalmente, se ha observado en todos los casos en los *corpora* analizados, las colocaciones 'impronta genética' y 'código genético'. En cambio, cuando se analiza el par 'génico' y 'genómico' no se observa un consenso tan amplio, existiendo varios sustantivos a los que modifican ambas voces: *mutación, polimorfismo, ligamiento, variante, marcador, regulación* (**Figura 9**).

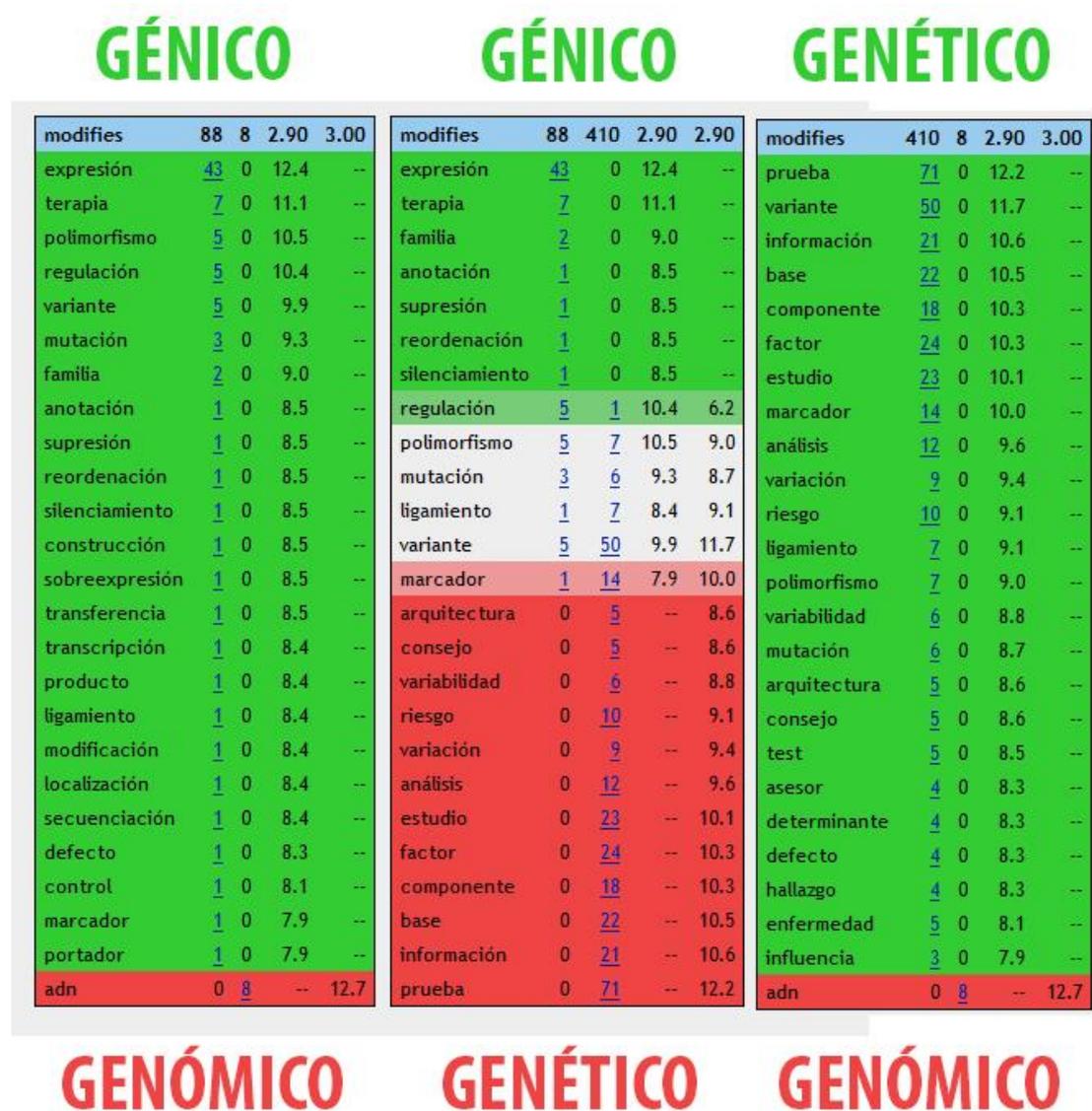


Figura 9.- Esquema comparativo de los términos modificados por las voces 'génico', 'genético' y 'genómico' comparadas por pares.

GENÉTICO/GENETISTA

En inglés existe una clara diferencia entre 'genetic' (adjetivo) y 'geneticist' (sustantivo), tal y como ocurre con otras voces (physic(al)/physicist, chemical/chemist). En español, existen las variantes 'física, co' y 'química, co' para indicar profesionales de dichos campos. En cambio, en español en ocasiones se confunde el uso de 'genético' como adjetivo (más extendido) con respecto a su uso como sustantivo, a pesar de que esta última posibilidad está contemplada por la RAE:

genético, ca

Del lat. mod. *geneticus* 'relativo al origen', der. del lat. *genēsis* 'génesis'; cf. gr. γεννητικός *gennētikós* 'que produce o genera'.

3. m. y f. genetista.

En los textos analizados, se observa la predominancia de 'genetista' para referirse a un profesional de la Genética (que significativamente es visto como un médico y no como un biólogo, como suele ser la generalidad –**Figura 10**). Solamente en tres casos, en el *corpus* divulgativo, se usa 'genético' con esta función:

[En 1990 los **genéticos** identificaron la zona del cromosoma Y que confiere masculinidad.]

[Los **genéticos** han creado cientos de cepas knockout de bacterias y ratones que tienen bloqueado un gen determinado.]

[Peter J. Richerson, **genético** de poblaciones, y Henry Plotkin, psicólogo.]

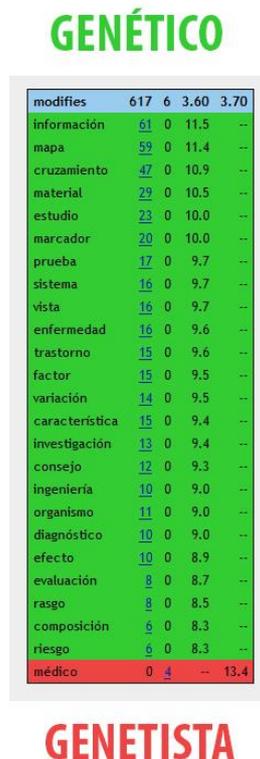


Figura 10.- Esquema comparativo de los términos modificados por las voces 'genético' y 'genetista'.

Probablemente para evitar la ambigüedad en el uso de estos términos, se observa un uso extraordinariamente variado de otras equivalencias para designar a los profesionales de esta disciplina:

[El **consejo genético** provee información y apoyo a las personas preocupadas por las patologías hereditarias en su familia.]

[Los **profesionales de la genética** estadounidenses han sido renuentes a rechazar pedidos como éste.]

[Los especialistas en **genética bacteriana** han creado técnicas para incrementar la frecuencia de transformación...]

HETEROCIGOTO/HETEROCIGÓTICO

Otro de los errores más comunes que se pueden observar en textos de Genética, es el empleo del sustantivo 'heterocigoto':

heterocigoto

De *hetero-* y del gr. ζυγωτός *zygōtós* 'uncido, unido', der. de ζυγοῦν *zygoûn* 'uncir, unir'.

1. m. Se refiere a un individuo que posee los dos alelos diferentes en un locus.

Fuente: Genética: un enfoque conceptual. Editorial Panamericana.

Y del adjetivo derivado, 'heterocigótico', como sinónimos. Esto queda ilustrado en los siguientes ejemplos extraídos de los *corpora* en español analizados en este trabajo:

Uso incorrecto:

[En resumen, se cruzan individuos **heterocigotos*** en dos o más loci genéticos...]

[En el **heterocigótico*** trans un cromosoma presenta un alelo funcional en el locus a...]

Sin embargo, al tratarse ambos términos de entidades gramaticales distintas, su uso no es intercambiable, y se propone el uso apropiado de ambas formas, siguiendo el patrón que sigue:

Uso correcto:

[En resumen, se cruzan individuos **heterocigóticos** en dos o más loci genéticos...]

[En el **heterocigoto** trans un cromosoma presenta un alelo funcional en el locus a...]

La tabla 13 muestra la alternancia de las formas adjetival ('zygous/cigótico') y nominal ('zygote/cigoto') en los *corpora* paralelos en lengua inglesa y española. Ninguna de las formas alternativas aparece en el *corpus* divulgativo. En el *corpus* educativo, ambas aparecen, siendo la forma adjetival más frecuente que la nominal. Si bien la proporción de formas adjetivales en relación a formas nominales es de 2,5:1 en inglés, mientras que es de 1,3:1 en español. Esto indica que parte de las formas adjetivales del inglés se han sustituido por formas nominales en los textos en español. Esto se pone claramente de manifiesto en el *corpus* científico, en el que se observa como todas las formas adjetivales inglesas han sido sustituidas por formas nominales en español (**Tabla 13**), tal y como se demuestra en el siguiente ejemplo:

[...whereas **heterozygous** mice develop severe dilated cardiomyopathy and the animals are viable only until 20 days.]

[...en los ratones **heterocigotos*** se desarrolla miocardiopatía dilatada severa, y son viables sólo hasta los 20 días de desarrollo.]

	Divulgativo	Educativo	Científico
*zygous	0/0	256/1 944,71	28/205,54
*zygote	0/0	102/774,85	8/58,73
*cigótico	0/0	113/744,85	0/0
*cigoto	0/0	89/586,65	12/94,24

Tabla 13.- Frecuencia de aparición de las formas *zig* con respecto a las formas *cig* en los *corpora* en inglés y español analizados. **Nota:** valores expresados en número de coincidencias/número de coincidencias por millón de palabras.

Todo esto aplica también a los pares de palabras 'homocigoto/homocigótico', 'hemicigoto/hemicigótico'.

Además, se observa la alternancia de las formas con 'z' (del inglés *zygote*) y con 'c' (de *cigoto*), ambas recogidas y admitidas por la RAE:

cigoto

Tb. **zigoto**.

Del gr. ζυγωτός *zygōtós* 'uncido, unido', der. de ζυγοῦν *zygoûn* 'uncir, unir'.

1. m. Biol. Célula resultante de la unión del gameto masculino con el femenino en la reproducción sexual de los animales y de las plantas.

Si bien, como muestra la tabla 14, las variantes con 'c' predominan claramente sobre las formas con 'z'.

	Divulgativo	Educativo	Científico
Formas *zig*	2/8,79	0/0	1/7,85
Formas *cig*	9/39,54	446/2 939,86	37/290,58

Tabla 14.- Frecuencia de aparición de las formas *zig* con respecto a las formas *cig* en los *corpora* analizados.

Nota: valores expresados en número de coincidencias/número de coincidencias por millón de palabras.

Nótese que aparecen coincidencias en el *corpus* divulgativo (que carece de las formas heteroc(z)igoto, homoc(z)igoto, hemic(z)igoto) por la presencia de la forma cigoto/zigoto.

4.2.4.- Proximidad semántica

RECOMBINACIÓN/ENTRECRUZAMIENTO

Existen dos términos en Genética que por su proximidad semántica se confunden habitualmente. Se trata de:

entrecruzamiento

1. m. Intercambio de material genético entre cromátidas homólogas pero no hermanas.

Fuente: Genética: un enfoque conceptual. Editorial Panamericana.

Que describe un evento físico que ocurre entre las moléculas de ADN durante la meiosis, y:

recombinación

1. f. Distribución de los alelos en nuevas combinaciones.

Fuente: Genética: un enfoque conceptual. Editorial Panamericana.

Que es una de las consecuencias del entrecruzamiento. El entrecruzamiento es la principal causa de la recombinación, pero no la única. Por ejemplo, mediante ingeniería genética se pueden sustituir, añadir o cambiar de posición fragmentos de ADN causando el llamado ADN recombinante.

Los textos divulgativos y científicos en español analizados utilizan exclusivamente la voz 'recombinación', mientras que en los textos educativos se observa la alternancia de las voces 'recombinación' y 'entrecruzamiento' (**Tabla 15**).

	Divulgativo	Educativo	Científico
Recombinación	22/96,66	168/1 107,39	4/31,41
Entrecruzamiento	0/0	212/1 397,42	0/0

Tabla 15.- Frecuencia de aparición de 'recombinación' y 'entrecruzamiento' en los *corpora* analizados. **Nota:** valores expresados en número de coincidencias/número de coincidencias por millón de palabras.

Cuando se analizan los modificadores de estos términos, se observa un consenso casi total en los textos analizados (**Figura 11**). Estos modificadores se refieren a las características del entrecruzamiento (simple, doble, desigual), por un lado, mientras que por otro, se refieren a las consecuencias de la recombinación (*intra- o intercromosómica, intragénica*). Con respecto a las palabras modificadas por las dos voces que nos ocupan, observamos cierto uso compartido de las palabras *episodio, tipo, probabilidad, tasa, frecuencia*, todas de carácter general y que en teoría podrían aplicar a ambos términos (**Figura 11**).

ENTRECRUZAMIENTO

n_modifier	86	27	1.50	0.60	modifies	35	100	0.60	2.30
doble	59	0	13.6	--	proceso	6	0	11.0	--
desigual	3	0	10.1	--	respecto	2	0	10.2	--
ocasional	1	0	8.5	--	progenie	4	0	9.7	--
considerable	1	0	8.5	--	ubicación	1	0	9.5	--
único	1	0	8.1	--	número	4	0	9.4	--
adicional	1	0	8.0	--	resultado	2	0	9.4	--
similar	1	0	8.0	--	nivel	1	0	9.1	--
cromosoma	1	0	6.8	--	clase	1	0	9.0	--
simple	18	1	12.2	8.9	grado	1	0	9.0	--
dna	0	1	--	7.2	través	1	0	8.9	--
genético	0	3	--	7.2	cantidad	1	0	8.2	--
génico	0	2	--	8.6	episodio	1	1	9.7	8.3
alto	0	1	--	9.3	tipo	4	4	9.0	8.7
fagos	0	1	--	9.7	probabilidad	1	3	8.8	9.5
global	0	1	--	9.9	tasa	2	14	10.0	11.8
meiótico	0	1	--	10.0	frecuencia	3	64	9.6	13.4
serie	0	1	--	10.0	prueba	0	1	--	7.6
aberrante	0	1	--	10.1	consecuencia	0	1	--	8.1
bajo	0	2	--	10.6	valor	0	1	--	8.1
intragénico	0	2	--	11.0	evidencia	0	1	--	8.2
intercromosómico	0	4	--	12.0	índice	0	1	--	8.2
intracromosómica	0	6	--	12.5	indicio	0	1	--	8.3
					suma	0	1	--	8.3
					instancia	0	1	--	8.3
					porcentaje	0	4	--	10.1

RECOMBINACIÓN

Figura 11.- Esquema comparativo de los términos modificados por y de los términos que modifican las voces 'entrecruzamiento' y 'recombinación'.

Por su parte, en los *corpora* en inglés se observa el uso exclusivo del término 'recombination' en los textos divulgativos y científicos, probablemente por la falta de especificidad de los mismos. En cambio, en los textos educativos, no sólo se observa la presencia de 'crossing over' y 'recombination', sino que aparece 'crossover' como la forma más numerosa, utilizada como sinónimo de 'entrecruzamiento' (Tabla 16).

	Divulgativo	Educativo	Científico
Crossing over	0/0	107/812,83	0/0
Crossover	0/0	127/964,76	0/0
Recombination	19/88,10	179/1 359,78	2/14,68

Tabla 16.- Frecuencia de aparición de 'crossing over', 'crossover' y 'recombination' en los *corpora* analizados. **Nota:** valores expresados en número de coincidencias/número de coincidencias por millón de palabras.

No se ha detectado la forma incorrecta, calco del inglés, 'sobrecruzamiento', que aparece en algunos textos:

[El sobrecruzamiento es responsable de variabilidad genética.]

Fuente: CREA, 1992

5.- Conclusiones

5.- Conclusiones

Las conclusiones del presente trabajo son:

1. Se ha analizado un total de 217 textos (853 479 palabras totales) y se han creado tres *corpora* de textos paralelos relacionados con la Genética en tres contextos diferentes:
 - *Contexto educativo*: el manual *Genetics: A Conceptual Approach* y su versión española.
 - *Contexto divulgativo*: artículos con la etiqueta 'Genética' de la plataforma online *OpenMind* (en versión inglesa y española) y artículos de la revista *Scientific American* y su traducción al español publicados en *Investigación y Ciencia*.
 - *Contexto científico*: extractos de Tesis Doctorales de la Universidad de Granada con la etiqueta 'Genética' y sus traducciones al inglés y artículos de la Revista Española de Cardiología (en versión inglesa y española).
2. El estudio de la frecuencia de uso de los términos en inglés y la búsqueda de las distintas equivalentes en español, no sólo dentro de los *corpora* creados *ad hoc* sino también en el *Corpus del Español Actual* (CREA) ha dado lugar a una lista de 17 términos que potencialmente podían presentar problemas de traducción.
3. Los problemas de traducción pueden dividirse en cuatro grupos principales: siglas, extranjerismos, uso de adjetivos y términos con proximidad semántica.
4. **Con respecto a las siglas:** en el *contexto educativo* se observa que tanto las siglas que están más asentadas ('ADN/DNA') como las que lo están menos o presentan potenciales problemas de traducción o comprensión ('PCR'), mantienen la forma inglesa, incluso cuando forman variantes ('mRNA', 'cpDNA'). El uso de éstas es coherente a lo largo de todo el texto y no se observa el uso incorrecto del plural terminado en 's'. En *contexto científico* se usa la traducción española en las siglas más asentadas ('ADN') y la inglesa en las que pueden crear algún tipo de problema de traducción ('PCR'). Sin embargo, se detecta cierta inconsistencia en la formación de variantes, que en ocasiones aparecen parcialmente adaptadas al español ('cADN', por ejemplo). En *contexto divulgativo* se observa en todos los casos la alternancia de las formas inglesa y española ('ADN/DNA'). Cuando se forman variantes de las siglas se observa falta de coherencia en la adaptación de las mismas al español ('siARN'). Se detecta el uso desaconsejado del plural terminado en 's' en estos dos últimos contextos ('DNAs').
5. **Con respecto a los extranjerismos:** en *contexto educativo*, se tiende a utilizar el equivalente en español ('corte y empalme' en lugar de 'splicing') o a la adaptación del mismo a nuestra lengua ('deleción', a veces llegando a utilizar formas verbales no aceptadas a partir de sustantivos 'deleccionar'*). En el caso de extranjerismos de nuevo cuño se opta por mantener la forma inglesa ('contig'). En *contexto divulgativo*, probablemente por la escasa especialización de los textos, no existen datos para dos de los extranjerismos analizados ('splicing' y 'contig'). Para el tercero, se detecta el uso de

la forma adaptada y aceptada en español ('deleción'). En *contexto científico* para el término 'splicing' conviven varios equivalentes en español ('corte y empalme' y 'ajuste') y la forma no adaptada del inglés ('splicing'). El término 'deleción' parece asentado, mientras que no se ha detectado la presencia del término 'contig'.

6. **Con respecto al uso de algunos adjetivos:** en *contexto divulgativo*, se usan los adjetivos terminados en *-ico* 'cromosómico', 'ribosómico' y 'vírico', con un uso residual de 'viral' (también aceptado por la RAE). En *contexto educativo* se detecta la presencia exclusiva de 'cromosómico' y 'viral'. En *contexto científico* se detecta la presencia de todas las formas en *-ico* y *-al* (incluyendo las formas incorrectas 'cromosomal'* y 'ribosomal'*), excepto 'vírico'. En líneas generales, los términos 'génico', 'genético' y 'genómico' no son usados como sinónimos en los *corpora* analizados. Además, se han detectado colocaciones fijas como 'impronta genética' y 'código genético'. No obstante lo dicho, la comparación de sustantivos a los que modifican 'génico' y 'genético' da lugar a una lista de algunos sustantivos de uso compartido. El empleo de 'genético' como sustantivo para describir a un profesional de la Genética se usa moderadamente y viene con frecuencia sustituido por otros sintagmas nominales equivalentes ('experto en Genética'). Excepto en el *corpus divulgativo*, donde no aparecen, se observa un uso intercambiable entre los sustantivos ('hetero-, homo-, hemicigoto') y los adjetivos ('hetero-, homo-, hemicigótico'). De hecho, muchas de las formas adjetivales en inglés de estos términos son sustituidas por formas nominales en los textos científicos y educativos en español analizados. Las formas con 'z' (zigoto) conviven con las formas con 'c' en los contextos divulgativos y científicos.

7. **Con respecto a términos con proximidad semántica:** se observa que 'recombinación' se destina a contextos poco específicos (*corpora* divulgativo y científico), en los que no aparece 'entrecruzamiento'. Ambas voces conviven en el *corpus educativo*, donde su uso parece no presentar problemas terminológicos graves, quedando 'recombinación' asociado a consecuencias del proceso ('aberrante') y 'entrecruzamiento' a las particularidades del mismo ('doble'). Los textos educativos en inglés utilizan las variantes 'crossing over' y 'crossover' como sinónimos, además del equivalente a 'recombinación', 'recombination'.

8. En general los textos paralelos aquí analizados representan buenos ejemplos de traducción científica para los problemas analizados. Se observa una mayor coherencia en las decisiones tomadas y en las estrategias utilizadas, sobre todo en los textos educativos, seguidos de los textos científicos y en menor medida, en los textos divulgativos. Detrás de este hecho, puede estar la procedencia de los mismos. Los textos del contexto educativo procedían de una sola fuente y probablemente hayan sido traducidos por el mismo equipo de traducción (usando memorias de traducción y herramientas de traducción asistida) del inglés al español. Los textos científicos procedían de varias fuentes; una de ellas, la Revista Española de Cardiología, cumple potencialmente los requisitos comentados para el caso anterior. Sin embargo, los textos de los fragmentos de Tesis Doctorales han tenido distintos traductores (con casi total seguridad no profesionales). Además, cabe la posibilidad de que haya habido bidireccionalidad en las traducciones (la mayor parte de ellas de español a inglés y en algunos casos, al contrario). Por último, los textos divulgativos proceden de las

traducciones para *Investigación y Ciencia* de la revista *Scientific American* y de las adaptaciones al inglés (en la mayoría de los casos) de textos escritos por científicos para la publicación online *OpenMind*. A juzgar por algunas inconsistencias detectadas en este último *corpus*, es posible que no se hayan utilizado herramientas de traducción asistida o que las traducciones se hayan llevado a cabo por equipos de traductores diversos.

6.- Bibliografía

6.- Bibliografía

Abril Gallego AM, Mayoral Martínez MV, Muela García FJ (2003). Los medios de comunicación social y la didáctica de la Genética y la Biología Molecular en E.S.O. *Actas del Congreso La nueva alfabetización: un reto para la educación del siglo XXI*, Madrid, 6-8 Diciembre 2003.

Alcina Caudet MA (2001). El español como lengua de la ciencia y de la medicina. *Panacea*, 4:47-50.

Allen GE (2003). Mendel and modern genetics: the legacy for today. *Endeavour*, 27:63-68.

Bernárdez E (1999). ¿Qué son las lenguas? Alianza, Madrid (España).

Bordons M (2004). Hacia el reconocimiento internacional de las publicaciones científicas españolas. *Rev. Esp. Cardiol.*, 57: 799–802.

Bowker L (1997). Multidimensional classification of concepts and terms. En: Wright SE y Budin G (Eds). 133-143.

Centro Virtual Cervantes (2016). Registro. *Diccionario de términos clave de ELE*. Instituto Cervantes (España). [online] (Fecha de consulta: 24 abril 2016) [http://cvc.cervantes.es/ensenanza/biblioteca_ele/diccio_ele/indice.htm/].

Claros MG (2008). Un poco de estilo en la traducción científica: aquello que quieres conocer pero no sabes dónde encontrarlo. *Panacea*, 28:145-158.

Correns KG (1900) G. Mendel's Regel über das Verhalten der Nachkommenschaft der Rassenbastarde. *Berichte der Deutschen Botanischen Gesellschaft*, 18: 158-168.

Darwin C (1859). El origen de las especies. Ed. John Murray. Londres (GB).

de Kock J (1994). Gramática española: enseñanza e investigación. Ediciones Universidad de Salamanca (España).

de Vries H (1900) Sur la loi de disjonction des hybrides. *Compt. Rend. Acad. Sci.*, 130: 845-847.

Departamento de Genética UGR (2016). [online] (Fecha de consulta: 22 Abril 2016) [<http://genetica.ugr.es/>]

Domínguez NA (2007). Diez errores usuales en la traducción de artículos científicos. *Panacea*, 8:121-123.

Donovan J, Venville G (2012). Exploring the influence of the mass media on primary students' conceptual understanding of genetics. *Education 3-13: International Journal of Primary, Elementary and Early Years Education*, 40:75-95.

Editorial (2006). Ich weiss nicht was soll es bedeuten: language matters in medicine. *PLoS Med*, 3:e122.

Faber P, León-Araúz P (2016). Specialized knowledge representation and the parameterization of context. *Frontiers in Psychology*, 7:196.

- Fundéu** (2011). Siglas extranjeras, claves de traducción. (online) (Fecha de consulta: 8 junio 2016) [<http://www.fundeu.es/recomendacion/siglas-traduccion/>].
- Fundéu** (2013). Siglas y acrónimos, claves de redacción. (online) (Fecha de consulta: 8 junio 2016) [<http://www.fundeu.es/recomendacion/siglas-y-acronimos-claves-de-redaccion/>].
- Fundéu** (2014a). Viral y vírico son sinónimos. (online) (Fecha de consulta: 9 junio 2016) [<http://www.fundeu.es/recomendacion/viral-y-virico-son-sinonimos/>].
- Fundéu** (2014b). Genético, génico y genómico no son sinónimos. (online) (Fecha de consulta: 9 junio 2016) [<http://www.fundeu.es/recomendacion/genetico-genico-y-genomico-no-son-lo-mismo/>].
- Gordin MD** (2015). *Scientific Babel: How Science Was Done Before and After Global English*. The University of Chicago Press Ltd., Londres (GB).
- Halliday MAK** (1978). *El lenguaje como semiótica social. La interpretación social del lenguaje y del significado*. Fondo de Cultura Económica. México D.F. (México).
- Holmes HK** (2004). An analysis of the language repertoires of students in higher education and their language choices on the internet (Ukraine, Poland, Macedonia, Italy, France, Tanzania, Oman and Indonesia). *Intern. J. Multicult. Soc.*, 6:52–75.
- Lacadena JR** (2011). *Genética y Sociedad*. Instituto de España, Real Academia Nacional de Farmacia. Realigraf S.A., Madrid (España).
- Marchio J, Mackenbach W** (2012). Miradas cruzadas/configuraciones recíprocas. Sobre la traducción, difusión y recepción de las literaturas centroamericanas en Francia y Alemania. *Centroamericana*, 22:241-268.
- Marcus A** (2002). Return on investment for usable user-interface design: Examples and statistics. Aaron Marcus and Associates, Inc. Emeryville, CA (EEUU).
- Mendel G** (1866) Versuche über Pflanzen-Hybriden. *Verh. Des Naturf. Vereines in Brünn (Abhandlungen)*, 4: 3-47.
- Meneghini R, Packer AL** (2007) Is there science beyond English? *EMBO reports*, 8 (2).
- Montero Martínez S, Faber Benítez P, Buendía Castro M** (2011). *Terminología para traductores e intérpretes*, 2ª Ed. Ediciones Tragacanto. Granada (España).
- Morgan TH** (1910). Sex-limited inheritance in *Drosophila*. *Science*, 32: 120-122.
- Mulet JM** (2016) Artículos escritos por JM Mulet. [online] (Fecha de consulta: 24 abril 2016) [http://elpais.com/autor/jose_miguel_mulet_salort/a/].
- Navajas-Pérez R, Aznarte-Mellado C** (2014). Mendelius: el acento granadino a las leyes de la herencia. *Calle Elvira*, Otoño:40-43.
- Navarro F** (2008). Adjetivos terminados en -al. Laboratorio del Lenguaje. (online) (Fecha de consulta: 9 junio 2016) [<http://medicablogs.diariomedico.com/>].
- Ofori-Adjei D, Antes G, Tharyan P, Slade E, Tamber PS** (2006) Have online international medical journals made local journals obsolete? *PLoS Med*, 3:e359.

- Proctor RN** (1999). *The Nazi War on Cancer*. Princeton University Press. Princeton, NJ (EEUU).
- RAE** (2010). *Ortografía de la lengua española*.
- Real Academia Nacional de Medicina** (2012). *Diccionario de términos médicos*. Editorial Panamericana (España).
- Ruiz Rejón M** (2009). La selección Natural de Darwin a la luz de la Genética. *Pasaje a la Ciencia*, 12:110-126.
- Sager JC** (1990). *A practical course in terminology processing*. John Benjamins Publishing Company. Amsterdam (Países Bajos).
- Science Encyclopedia** (2016) [online] (Fecha de consulta: 18 abril 2016) [<http://www.libraryindex.com/pages/2213/History-Genetics-EARLY-BELIEFS-ABOUT-HEREDITY.html>]
- Sociedad Española de Genética** (2016). [online] (Fecha de consulta: 22 Abril 2016) [<http://www.segenetica.es/departamentosBD.php/>]
- Tercedor Sánchez M** (2005). Aspectos culturales en la localización de productos multimedia. *Quaderns. Revista de traducció*, 12:151-160.
- Tschermak E** (1900) Über Künstliche Kreuzung bei *Pisum sativum*. *Berichte der Deutschen Botanischen Gessellschaft*, 18: 232-239. 5
- Turney J** (1995). The public understanding of genetics – where next? *European Journal of Genetics and Society*, 1:5-20.
- Viegas E, Mahesh K, Nirenburg S** (1999). Semantics in action. En: Saint-Dizier, P (ed). *Predicative forms in natural language and lexical knowledge bases*. Kluwer Academic Press. Berlín (Alemania). 171-204.
- Watson JD, Crick FHC** (1953) The molecular structure of nucleic acids. A structure for deoxyribose nucleic acid. *Nature*, 171: 737-738.

7.- Anexos

7.- Anexos

Anexo I