



TC/50/14

ORIGINAL: Inglés

FECHA: 13 de marzo de 2014

# UNIÓN INTERNACIONAL PARA LA PROTECCIÓN DE LAS OBTENCIONES VEGETALES

Ginebra

## COMITÉ TÉCNICO

**Quincuagésima sesión  
Ginebra, 7 a 9 de abril de 2014**

### DENOMINACIONES DE VARIEDADES

*Documento preparado por la Oficina de la Unión*

*Descargo de responsabilidad: el presente documento no constituye un documento de política u orientación de la UPOV*

1. La finalidad del presente documento es facilitar un resumen de las eventuales novedades relativas a la revisión de las "Notas explicativas sobre las denominaciones de variedades con arreglo al Convenio de la UPOV" (documento UPOV/INF/12/4), a la elaboración de un instrumento de la UPOV de búsqueda de denominaciones similares a los fines de la denominación de variedades y a los posibles ámbitos de cooperación entre la Comisión Internacional para la Nomenclatura de Plantas Cultivadas de la Unión de Ciencias Biológicas (Comisión de la UICB), la Comisión de Nomenclatura y Registro de Cultivares de la Sociedad Internacional de la Ciencia Hortícola (Comisión de la ISHS) y la UPOV.

### ÍNDICE

I.	EVENTUAL REVISIÓN DEL DOCUMENTO UPOV/INF/12 "NOTAS EXPLICATIVAS SOBRE LAS DENOMINACIONES DE VARIEDADES CON ARREGLO AL CONVENIO DE LA UPOV" .....	2
	LABOR QUE ESTÁ REALIZANDO EL CAJ-AG EN MATERIA DE DENOMINACIONES DE VARIEDADES.....	2
	PROPUESTA ACERCA DE LA ORIENTACIÓN SOBRE DENOMINACIONES DE VARIEDADES .....	2
II.	POSIBLE ELABORACIÓN DE UN INSTRUMENTO DE LA UPOV DE BÚSQUEDA DE DENOMINACIONES SIMILARES A LOS FINES DE LA DENOMINACIÓN DE VARIEDADES.....	3
	ANTECEDENTES .....	3
	CREACIÓN DE UN GRUPO DE TRABAJO .....	4
III.	NOVEDADES RELATIVAS A POSIBLES ÁMBITOS DE COOPERACIÓN CON LA COMISIÓN DE LA UICB Y LA COMISIÓN DE LA ISHS .....	6
	ANEXO: PROCEDIMIENTO DE BÚSQUEDA DE LA CPVO	

2. En el presente documento se utilizan las siguientes abreviaturas:

CAJ: Comité Administrativo y Jurídico

CAJ-AG: Grupo Asesor del Comité Administrativo y Jurídico

TC: Comité Técnico

Comisión de la UICB: Unión Internacional de Ciencias Biológicas

Comisión de la ISHS: Comisión de la Sociedad Internacional de la Ciencia Hortícola

I. EVENTUAL REVISIÓN DEL DOCUMENTO UPOV/INF/12 “NOTAS EXPLICATIVAS SOBRE LAS DENOMINACIONES DE VARIEDADES CON ARREGLO AL CONVENIO DE LA UPOV”

Labor que está realizando el CAJ-AG en materia de denominaciones de variedades

3. En su octava sesión celebrada en Ginebra el 21 y el 25 de octubre de 2013, el CAJ-AG analizó el documento CAJ-AG/13/8/6 “*Matters concerning variety denominations*” (Asuntos relativos a las denominaciones de variedades)”

4. El CAJ-AG acordó que se elaboraran orientaciones relativas a la petición de un obtentor solicitando el cambio de la denominación registrada de una variedad en casos distintos de los de cancelación de la denominación de la variedad tras la concesión del derecho, sobre la base de que tal petición debe denegarse. No obstante, el CAJ-AG convino en que dicho cambio sería conveniente en las siguientes situaciones:

a) si se descubriera que había un derecho anterior sobre la denominación que habría tenido por resultado el rechazo de la denominación (véanse el artículo 20.4) y 7) del Acta de 1991, el artículo 13.4) y 7) del Acta de 1978, y la nota 7 del documento UPOV/INF/12/4);

b) si la denominación resultara inadecuada por ser contraria a las disposiciones del artículo 20.2) del Acta de 1991 y del artículo 13.2) del Acta de 1978; y

c) si la denominación fuera rechazada posteriormente por otro miembro de la Unión y, a petición del obtentor, la autoridad decidiera cambiar la denominación por la registrada en ese otro miembro de la Unión.

5. El CAJ-AG acordó examinar la posibilidad de elaborar orientaciones adicionales como parte de una eventual revisión de las “Notas explicativas sobre las denominaciones de variedades con arreglo al Convenio de la UPOV” (documento UPOV/INF/12/4) (véase el documento CAJ-AG/13/8/10 “Informe”, párrafos 69 a 71).

6. El informe antes mencionado sobre la labor que está realizando el CAJ-AG en materia de denominaciones de variedades se presentará al CAJ en su sexagésima novena sesión que se celebrará en Ginebra el 10 de abril de 2014 (véanse los párrafos 56 a 59 del documento CAJ/69/2 “Elaboración de material de información sobre el Convenio de la UPOV”).

Propuesta acerca de la orientación sobre denominaciones de variedades

7. Desde el 11 de noviembre al 15 de diciembre de 2013, se comprobó el funcionamiento de la versión en inglés del curso de enseñanza a distancia “Examen de las solicitudes de derechos de obtentor” (DL 305), contando con la participación, en calidad de alumnos, de nueve expertos que a su vez son tutores del curso DL-205. Uno de los alumnos hizo el siguiente comentario relativo al párrafo 2.3.3.a)i) del documento UPOV/INF/12/4 (reproducido a continuación para facilitar su consulta):

“2.3.3 Identidad de la variedad

a) Como recomendación general, sólo una diferencia de una letra o un número podrá considerarse susceptible de inducir a error o prestarse a confusión en relación con la identidad de la variedad, salvo cuando:

i) la diferencia de una letra indique una diferencia visual o fonética clara; por ejemplo, si se trata de una letra al comienzo de una palabra:

*Ejemplo 1:* en el idioma inglés, “Harry” y “Larry” no se prestarían a confusión; sin embargo, “Bough” y “Bow” podrían causar confusión (en términos fonéticos).

*Ejemplo 2:* en los idiomas japonés y coreano, no hay diferencia entre los sonidos de las letras “L” y “R”, por lo que “Lion” y “Raion” suenan exactamente lo mismo, aunque sean palabras que las personas de habla inglesa pueden distinguir perfectamente; [...]”

Comentario:

“La diferencia de una letra del subapartado i) indica una diferencia visual o fonética clara; por ejemplo, si se trata de una letra al comienzo de una palabra, se ofrece el ejemplo 1. El caso de ‘Harry’ y ‘Larry’ ilustra

bien el argumento. Lo que me preocupa es el caso de 'Bough' y 'Bow', que sin duda podrían confundirse entre sí por motivos fonéticos. Sin embargo, no son un ejemplo de denominaciones que podrían causar confusión debido a la diferencia de una letra e indicar que lo son podría distraer y confundir a los alumnos.

¿Se describiría mejor el caso de 'Bough' y 'Bow' en un subapartado diferente dedicado a la confusión fonética? Me imagino que mi reparo afecta al documento UPOV/INF/12 y a la redacción del módulo."

8. En su sexagésima novena sesión que se celebrará en Ginebra el 10 de abril de 2014, se invitará al CAJ a considerar si sería conveniente modificar el documento UPOV/INF/12, tal como se propone en el párrafo 7 del presente documento.

9. Las observaciones realizadas por el TC en su quincuagésima sesión se presentarán al CAJ en su sexagésima novena sesión.

10. *Se invita al TC a:*

*a) tomar nota de la labor que está realizando el CAJ-AG en materia de elaboración de orientaciones sobre las denominaciones de variedades, expuesta en los párrafos 3 a 6 supra;*

*b) tomar nota de que se invitará al CAJ, en su sexagésima novena sesión que se celebrará en Ginebra el 10 de abril de 2014, a considerar si sería conveniente modificar el párrafo 2.3.3.a)i) del documento UPOV/INF/12, como se propone en el párrafo 7 del presente documento; y*

*c) hacer observaciones, si procede, sobre la posible revisión del documento UPOV/INF/12.*

## II. POSIBLE ELABORACIÓN DE UN INSTRUMENTO DE LA UPOV DE BÚSQUEDA DE DENOMINACIONES SIMILARES A LOS FINES DE LA DENOMINACIÓN DE VARIEDADES

### Antecedentes

11. El CAJ, en su sexagésima séptima sesión, celebrada en Ginebra el 21 de marzo de 2013, asistió a una ponencia presentada por la Delegación de la Unión Europea sobre la experiencia de la Oficina Comunitaria de Variedades Vegetales (OCVV) en la utilización de su instrumento de búsqueda de denominaciones similares en el examen de las propuestas de denominación de variedades. Durante la presentación, la OCVV propuso estudiar la posibilidad de elaborar un instrumento de la UPOV de búsqueda de denominaciones similares a los fines de la denominación de variedades, sobre la base del instrumento de búsqueda de la OCVV.<sup>1</sup> El CAJ acogió con agrado el ofrecimiento de la OCVV y acordó incluir un punto dedicado a considerar esa propuesta en su sexagésima octava sesión, que se iba a celebrar en Ginebra el 21 de octubre de 2013 (véase el documento CAJ/67/14 "Informe sobre las conclusiones", párrafos 49 y 50).

12. Actualmente la pestaña de búsqueda por denominación de la base de datos sobre variedades vegetales (base de datos PLUTO) (<https://www3.wipo.int/pluto/user/en/index.jsp>) proporciona las siguientes posibilidades de búsqueda para encontrar denominaciones similares:

<sup>1</sup>

La elaboración del factor de similitud estuvo a cargo del Grupo de Estudio y Control de Variedades y Semillas (GEVES) de Francia, en colaboración con la Oficina Comunitaria de Variedades Vegetales (OCVV) de la Unión Europea.

“a) Factor de similitud [Instrumento de búsqueda de la OCVV] Se efectúa un análisis de la denominación que se haya introducido sobre la base de una combinación de factores, incluidas las letras en común, la longitud relativa de las palabras y la ubicación de las letras en común. Se trata del método de comparación más complejo y puede que la búsqueda tarde varios segundos en realizarse. El factor de similitud fue ideado por el Grupo de Estudio y Control de Variedades y Semillas (GEVES) (Francia) en colaboración con la Oficina Comunitaria de Variedades Vegetales (OCVV) de la Unión Europea. No obstante, debe tenerse en cuenta que los resultados de la búsqueda efectuada con arreglo al factor de similitud en la base de datos PLUTO deben ser interpretados y no constituyen una garantía en cuanto a la idoneidad de las denominaciones de variedades, que debe ser decidida por las autoridades encargadas de la concesión de derechos de obtentor en el territorio de que se trate.

En el Anexo del presente documento se facilita una explicación pormenorizada del análisis.

“b) Búsqueda aproximada [*fuzzy*] Se buscan denominaciones que contengan palabras con uno o dos caracteres diferentes de los términos introducidos. Se parece al método de correspondencia aproximada de la pestaña de ‘Búsqueda de términos’.

“c) Búsqueda fonética [*phonetic*] Se buscan denominaciones que contengan palabras que suenen de forma similar a los términos introducidos. Se parece al método de correspondencia fonética de la pestaña de ‘Búsqueda de términos’.

“d) Contiene [*contains*] Se buscan denominaciones que contengan palabras que a su vez contengan la misma serie de letras que los términos que se hayan introducido. Se parece al método de correspondencia contiene de la pestaña de ‘Búsqueda de términos’.

“e) Empieza por [*starts*] Se buscan denominaciones que contengan palabras que empiecen con la misma serie de letras que los términos introducidos. Se parece al método de búsqueda correspondencia comienza por de la pestaña de ‘Búsqueda de términos’.

“f) Termina por [*ends*] Se buscan denominaciones que contengan palabras que finalicen con la misma serie de letras que los términos introducidos. Se parece al método de búsqueda de correspondencia termina por de la pestaña de “Búsqueda de términos”.

13. Durante las conversaciones mantenidas con la Oficina de la Unión a fin de examinar la manera de elaborar un instrumento de la UPOV de búsqueda de denominaciones similares a los fines de la denominación de variedades, la OCVV aclaró que debían contemplarse todas las opciones y que, teniendo en cuenta los avances en la tecnología de la información, era posible que el mejor instrumento no se basara necesariamente en el de la OCVV. El principal factor a tener en cuenta debía ser elaborar un instrumento que pudiesen utilizar todos los miembros de la UPOV, a fin de minimizar las diferencias en las decisiones sobre la idoneidad.

#### Creación de un grupo de trabajo

14. En su sexagésima octava sesión celebrada en Ginebra el 21 de octubre de 2013, el CAJ examinó el documento CAJ/68/9, “Posible elaboración de un instrumento de la UPOV de búsqueda de denominaciones similares a los fines de la denominación de variedades” y aprobó la creación de un grupo de trabajo que elabore propuestas para un instrumento de la UPOV de búsqueda de denominaciones similares a los fines de la denominación de variedades, de conformidad con lo dispuesto en los párrafos 4 a 7 de dicho documento (véase el párrafo 40 del documento CAJ/68/10 “Informe sobre las conclusiones”):

“Dicho grupo podría estar integrado por:

- “a) examinadores de denominaciones que sean miembros de la Unión (3 a 6 expertos);
- “b) el Servicio de Bases de Datos Mundiales de la OMPI (responsable de la base de datos PLUTO);
- “c) la Oficina Comunitaria de Variedades Vegetales (OCVV) de la Unión Europea); y
- “d) la Oficina de la Unión.

“El grupo de trabajo establecerá su propio plan de trabajo; sin embargo, cabe esperar que su primer paso sea examinar los tipos de búsqueda existentes actualmente en la pestaña de Búsqueda por denominación, de la base de datos PLUTO, en especial el factor de similitud (instrumento de búsqueda de la OCVV), y analizar los tipos de búsqueda utilizados en otras situaciones (p. ej. en relación con marcas) que podrían proporcionar una base alternativa al instrumento de búsqueda de denominaciones similares de la UPOV.

“En el análisis de la idoneidad de los tipos de búsqueda se deberá, en particular, tener en cuenta especialmente el documento UPOV/INF/12 “Notas explicativas sobre las denominaciones de variedades con arreglo al Convenio de la UPOV”. A ese respecto, el grupo de trabajo tendrá que solicitar al CAJ orientación adicional si su labor indicara que es necesario el análisis del documento UPOV/INF/12 para elaborar un instrumento eficaz de búsqueda de denominaciones similares para la UPOV.

“Las reuniones del grupo de trabajo se celebrarán en la Oficina de la Unión en Ginebra y las presidirá dicha Oficina. Se dispondrá que las reuniones no coincidan con las sesiones de la UPOV y se prevé la participación electrónica de los examinadores de denominación y de la OCVV. Las propuestas elaboradas por el grupo de trabajo se presentarán al CAJ y al Comité Técnico (TC) y ambos comités recibirán un informe breve de las reuniones del grupo de trabajo.”

15. En su sexagésima octava sesión, el CAJ tomó nota de la propuesta formulada por la Delegación de la Unión Europea de incluir examinadores de denominaciones de los Países Bajos y España en el grupo de trabajo y de la importancia de asegurar que haya una cobertura suficiente de expertos en aspectos lingüísticos de las denominaciones de variedades (véase el párrafo 41 del documento CAJ/68/10 “Informe sobre las conclusiones”).

16. En su sexagésima octava sesión, el CAJ convino en que debe alentarse a los miembros y los observadores a que hagan sugerencias sobre cuestiones relativas al mandato del grupo de trabajo (véase el párrafo 42 del documento CAJ/68/10 “Informe sobre las conclusiones”). La primera reunión del grupo de trabajo se organizará para junio/julio de 2014.

17. Las observaciones realizadas por el TC en su quincuagésima sesión se presentarán al CAJ en su sexagésima novena sesión.

18. *Se invita al TC a:*

*a) tomar nota del informe relativo a la posible elaboración de un instrumento de la UPOV de búsqueda de denominaciones similares a los fines de la denominación de variedades facilitado en la sección II del presente documento; y*

*b) hacer observaciones, si procede, sobre la posible elaboración de un instrumento de la UPOV de búsqueda de denominaciones similares a los fines de la denominación de variedades.*

III. NOVEDADES RELATIVAS A POSIBLES ÁMBITOS DE COOPERACIÓN CON LA COMISIÓN DE LA UICB Y LA COMISIÓN DE LA ISHS

19. La información sobre los antecedentes de ese tema figura en los párrafos 2 a 25 del documento TC/49/8, "Denominaciones de variedades".

20. En su cuadragésima octava sesión, celebrada en Ginebra del 26 al 28 de marzo de 2012, el TC tomó nota de la información presentada por la Delegación del Japón de que la Comisión de la UICB está iniciando la revisión del Código Internacional de Nomenclatura de Plantas Cultivadas (CINPC) y presentará propuestas en 2013 a la Comisión Internacional para la Nomenclatura de Plantas Cultivadas de la UICB. Acordó que la Oficina de la Unión debería ponerse en contacto con el CINPC para explicar las orientaciones proporcionadas por la UPOV en el documento UPOV/INF/12 "Notas explicativas sobre las denominaciones de variedades con arreglo al Convenio de la UPOV" (véase el párrafo 90 del documento TC/48/22, "Informe sobre las conclusiones").

21. En su cuadragésima novena sesión celebrada en Ginebra del 18 al 20 de marzo de 2013, el TC examinó el documento TC/49/8 "Denominaciones de variedades" y tomó nota de las novedades relativas a los posibles ámbitos de cooperación entre la UPOV, la Comisión Internacional para la Comisión de la UICB y la Comisión de la ISHS, según se expone en los párrafos 24 y 25 del documento TC/49/8 (véase el párrafo 89 del documento TC/49/41 "Informe sobre las conclusiones").

22. En su sexagésima octava sesión, celebrada en Ginebra el 21 de octubre de 2013, el CAJ examinó el documento CAJ/68/5 "Denominaciones de variedades". El 19 de julio de 2013, en Beijing (China), la Oficina de la Unión participó como conferenciante en el VI Simposio Internacional sobre la Nomenclatura de las Plantas Cultivadas (ISTCP 2013), organizado por la Universidad de Silvicultura de Beijing y el Jardín Botánico de Beijing, bajo los auspicios de la Sociedad Internacional de la Ciencia Hortícola (ISH). La Oficina de la Unión explicó la orientación que ofrece la UPOV sobre las denominaciones de variedades (véase el párrafo 22 del documento CAJ/68/10 "Informe sobre las conclusiones").

23. Los días 20 y 21 de julio, igualmente en Beijing, la Oficina de la Unión participó en calidad de observador en las reuniones de la Comisión de la UICB. En esas reuniones, la Comisión de la UICB consideró varias propuestas de modificación de la octava edición del Código Internacional de Nomenclatura de Plantas Cultivadas (CINPC). Esas propuestas fueron publicadas en junio de 2013 en el Volumen 7 de la revista "Hanburyana" <http://www.rhs.org.uk/Plants/RHS-Publications/Journals/Hanburyana/Hanburyana-issues/Volume-7--June-2013>. Las propuestas acordadas por la Comisión de la UICB quedarán reflejadas en la novena edición del CINPC, que está previsto que se publique en 2014.

24. En esa ocasión, la Comisión de la UICB formuló la propuesta de establecer un grupo de trabajo, del que formará parte la UPOV, en las etapas iniciales de la labor preparatoria para la décima edición del CINPC. De manera paralela a las reuniones de Beijing, tuvo lugar un encuentro informal entre la Oficina de la Unión y la Sra. Janet Cubey, Presidenta de la Comisión de la ISHS, con el fin de examinar las posibilidades de lograr una mayor armonización de las clases de denominación. Se sugirió estudiar los ámbitos de cooperación en materia de clases de denominación en el grupo de trabajo propuesto para la décima edición del CINPC.

25. La labor preparatoria de la décima edición del CINPC tuvo lugar en una reunión que se celebró entre los miembros de la Comisión de la UICB el 4 de marzo de 2014. En su quincuagésima sesión, el TC recibirá un informe verbal acerca de las novedades pertinentes.

*26. Se invita al TC a tomar nota de las novedades relativas a posibles ámbitos de cooperación entre la Comisión Internacional para la Nomenclatura de Plantas Cultivadas de la Unión de Ciencias Biológicas (Comisión de la UICB), la Comisión de Nomenclatura y Registro de Cultivares de la Sociedad Internacional de la Ciencia Hortícola (Comisión de la ISHS) y la UPOV, según lo expuesto en la sección III del presente documento.*

[Sigue el Anexo]



## PROCEDIMIENTO DE BÚSQUEDA

### 1. General

Como conclusión de la fase de estudio del proyecto presentado a su Comisión de Auditoría en noviembre de 2003, la Oficina propuso, en primera instancia, utilizar programas informáticos nacionales, adaptarlos a la base de datos centralizada y ejecutarlos todos para cada comprobación. Se estima que, de esta manera, la posibilidad de no dar con una denominación similar es muy escasa.

A más largo plazo, se prevé desarrollar un programa informático de la OCVV, y la posibilidad de ampliar las funciones lingüísticas.

En la práctica, se ha partido de las especificaciones del programa informático francés para llevar a cabo el procedimiento de búsqueda en la base de datos de la OCVV.

### 2. Normas para determinar una distinción suficiente entre 2 denominaciones de variedades

Según el Reglamento de base relativo a la protección comunitaria de las obtenciones vegetales, una de las normas que debe cumplir la denominación de una variedad es que no coincida o pueda confundirse con una denominación de variedad bajo la cual figure en un registro oficial otra variedad de la misma especie o estrechamente relacionada con ella.

Esta norma se ha interpretado en las directrices del Consejo de Administración de la OCVV relativas a las denominaciones de variedades:

- ♣ Por lo general, una diferencia de tan solo una letra o un número o de una tilde en una letra debe considerarse como susceptible de prestarse a confusión.
- ♣ En general no debe considerarse que diferencias de dos o más letras inducen a confusión, excepto cuando simplemente se yuxtaponen las mismas letras.
- ♣ Además, una denominación de variedad no debe transmitir la falsa impresión de que la variedad está relacionada con otra o se deriva de ella;

El propósito del programa de búsqueda es descubrir denominaciones de la misma clase en la base de datos que pudieran estar en conflicto con una denominación propuesta.

### 3. Procedimiento de búsqueda

Las comprobaciones se realizaron mediante un programa interno de la base de datos ORACLE en el servidor de la OCVV (mayor rendimiento).

Para llevar a cabo una comprobación, el programa de interfaz (sitio web, ...) ejecuta un proceso denominado:

TESTDENOMINATION (comprobar denominación).

Como parámetros, el programa de interfaz transmite la denominación a comprobar y el código de la especie a la que pertenece la variedad.

Al finalizar el proceso, se obtiene un identificador de la comprobación efectuada en la interfaz (columna Testid). Con este identificador es posible leer los 2 cuadros que constituyen el resultado de la comprobación: los cuadros TESTS (comprobaciones) y TESTRESULTS (resultado de las comprobaciones).

En el cuadro TESTS figuran los datos generales sobre la comprobación: la fecha de la comprobación, el identificador de la persona que solicitó la comprobación, la denominación, el código de la especie, el código de la clase o el género, el tiempo de ejecución informática, las palabras excluidas, los mensajes de error,...

En el cuadro TESTRESULTS se presentan las listas de las denominaciones similares que se han encontrado mediante el proceso TESTDENOMINATION. Cada denominación similar está acompañada de un índice de similitud.

#### 4. Descripción del proceso testdenomination

##### Control de entrada

El código de la especie debe existir en la base de datos de denominaciones.

La longitud máxima de la denominación es de 100 caracteres (incluidos los caracteres vacíos).

Caracteres excluidos: caracteres de puntuación y /-\_' . No están permitidos los caracteres con tilde.

Están permitidas las denominaciones compuestas de varias palabras: 4 palabras como máximo. Deben estar separadas por un espacio en blanco.

##### a) Primera operación: dividir la denominación introducida en palabras individuales

Se divide la denominación que se va a comprobar en palabras individuales. Los caracteres vacíos se consideran separadores y se eliminan.

Se reemplazan por caracteres latinos básicos aquellos caracteres latinos que no lo son (caracteres con tilde,...). Todas las letras se transforman en mayúsculas. Véase el cuadro de transformación.

*Ejemplo: Déjà se transforma en DEJA*

Las letras dobles se reducen a una sola letra.

*Ejemplo: HELLO se transforma en HELO.*

Cada palabra individual se compara con la lista de palabras excluidas para la comprobación (es decir: YELLOW, RED, PURPLE,...).

Las palabras individuales encontradas en dicha lista se excluyen de la comprobación de similitud.

En la descripción siguiente, las palabras individuales obtenidas se denominan Words Tested (WT)

*Ejemplo: Denominación 'Tānau TARI YELLOW'*

Se excluye la palabra YELLOW de las comprobaciones. Las palabras tenidas en cuenta (WT) son:

- TANAU,
- TARI,
- TANAUTARI,
- TARIYELLOW,
- TANAUTARIYELLOW.

## b) Segunda operación: elaboración de una lista de palabras de referencia de la clase

El programa busca la clase o el género al que pertenece la especie, para determinar el alcance de la búsqueda de similitud.

En caso de denominaciones registradas como códigos, se considera una sola cadena de caracteres, sin espacios en blanco.

El programa informático confecciona una lista de todas las palabras individuales que pertenecen a denominaciones de las variedades de la clase.

No se tienen en cuenta todas aquellas palabras que tengan más de 3 caracteres que la WT más larga y menos de 3 caracteres que la WT más corta.

Se excluyen de la comprobación de similitud todas las palabras de referencia que pertenecen a la lista de palabras excluidas para la comprobación antes mencionadas (es decir: color YELLOW, RED, PURPLE,...).

Como se comentó anteriormente, las letras dobles de las palabras de referencia se reducen a una sola letra.

En la descripción siguiente, las palabras individuales incluidas en la lista se denominan palabras de referencia (WR)

## c) Principios de la comprobación de similitud

Para cada WT, el proceso SIMILARITYTEST calcula un índice de similitud con respecto a cada WR incluida en la lista de denominaciones antes confeccionada.

La lista de WR se ordena según el valor del índice de similitud. Todas las WR con un índice de similitud superior a un umbral predefinido se excluyen de los resultados.

Ejemplo:

Denominación comprobada 'Tānau TARI YELLOW'

La lista de palabras de referencia se compone de 10.000 palabras individuales.

TANAU, TARI, TANAUTARI, TARIYELLOW y TANAUTARIYELLOW se contrastan con las 10.000 WR. Se efectúan 50.000 comprobaciones (5 palabras x 10.000 WR).

## 5. Descripción detallada de la comprobación de similitud

Para cada par (WT, WR) se siguen los siguientes pasos.

Filtro preliminar: de la lista de WR, todas las palabras con más de 3 caracteres que la WT o menos de 3 caracteres que la WT no se consideran en los cálculos siguientes.

### Primer paso: cálculo de $\chi^2$

Fórmula:  $\chi^2 = \text{sum}(di)^2 / (\text{longitud}(WT)-1)(\text{longitud}(WR)-1)$

donde di = diferencia del número de letras entre la palabra comprobada y la palabra de referencia. Se tienen en cuenta todas las letras de ambas palabras.

Ejemplo:

WT: **ALADIN** longitud: 6 caracteres.

Si comparamos esta WT con una cadena de caracteres existente: **DYLAN** (5 caracteres)

	A	L	D	I	N	Y
ALADIN	2	1	1	1	1	0
DYLAN	1	1	1	0	1	1

$$\chi^2 = ((2-1)^2 + (1-1)^2 + (1-1)^2 + (1-0)^2 + (1-1)^2 + (0-1)^2) / (6-1)(5-1)$$

$$\chi^2 = (1 + 0 + 0 + 1 + 0 + 1)(5 \cdot 4)$$

$$\chi^2 = 3/20$$

$$\chi^2 = 0,15$$

Luego, conservamos para el siguiente cálculo todas las palabras en que:

- $\chi^2 \leq 0,3$  y longitud de la WT  $\geq 5$  caracteres
- $\chi^2 \leq 0,4$  y longitud de la WT = 4 caracteres
- $\chi^2 \leq 0,5$  y longitud de la WT < 4 caracteres

## **Segundo paso: cuatro cálculos basados en la selección efectuada en el primer paso**

En este segundo paso se realizan 4 cálculos:

- El cálculo del porcentaje de las letras comunes.
- El cálculo del porcentaje de las letras NO comunes.
- El cálculo del porcentaje de la diferencia de longitud.
- El cálculo de la correlación por rangos de Kendall.

### **Cálculo del porcentaje de las letras comunes (LC)**

$$LC = 1 - (\text{Nº de letras comunes}) / (\text{longitud (WT)})$$

Ejemplo: ALADIN y DYLAN

$$LC = 1 - 4/6 = 0,33$$

LC es igual a 0 cuando todas las letras se encuentran en la WR.

Segundo ejemplo: BANANAS y BANS

Todas las letras de BANANAS se encuentran en la palabra BANS.

$$LC = 1 - 7/7 \Rightarrow LC = 0$$

### **Cálculo del porcentaje de las letras NO comunes (LNC)**

$$LNC = (\text{Nº de letras en WR ausentes en WT}) / (\text{longitud (WT)})$$

Ejemplo: ALADIN y DYLAN

$$LNC = 1/6 = 0,16$$

LNC es igual a 0 cuando todas las letras de WR se encuentran en la WT.

Segundo ejemplo: BANANAS y BANS

Todas las letras de BANS se encuentran en la palabra BANANAS.

$$LNC = 0/7 \Rightarrow LNC = 0$$

**Cálculo del porcentaje de la diferencia de longitud (DL)**

DL = (diferencia de longitud entre las 2 cadenas)/(longitud de la WT)

Ejemplo: BANANAS y ANANAS

DL = 1/7

DL es igual a 0 cuando las longitudes de las 2 palabras son iguales.

**Cálculo de la correlación por rangos de Kendall (CK)**

**Fórmula:  $CK = 6 \cdot \sum(D_i)^2 / N \cdot (N^2 - 1)$**

dónde:

- Di es igual a la diferencia de posición de las letras comunes (Li) menos la diferencia de posición de las letras previas. Si en la WT aparece varias veces la misma letra, se utilizan las primeras letras como referencia para la posición.  
 Si en la WR aparece varias veces la misma letra, se utilizan las letras más próximas como la misma letra en la WT.
- N es igual al número de letras comunes entre la WT y la WR.

Ejemplo:

WT: **ALADIN** y **DYLAN**

	<b>A</b>	<b>L</b>	<b>D</b>	<b>N</b>
<b>ALADIN</b>	1	2	4	6
<b>DYLAN</b>	4	3	1	5
<b>Diferencia</b>	1-4 = -3	2-3 = -1	4-1 = 3	6-5 = 1
<b>Di</b>	-3-0 = -3	1-(-3) = 4	3-(-1) = 4	1-3 = -2

4 letras en común A, L, D y N

$$CK = 6 \cdot ((-3)^2 + (4)^2 + (4)^2 + (-2)^2) / (4 \cdot (4^2 - 1))$$

$$CK = 6 \cdot (9 + 16 + 16 + 4) / 4 \cdot 15$$

$$CK = 6 \cdot 45 / 60$$

$$CK = 4,5$$

WT: **ALADIN** y **BALADIN**

	<b>A</b>	<b>L</b>	<b>D</b>	<b>I</b>	<b>N</b>
<b>ALADIN</b>	1	2	4	5	6
<b>BALADIN</b>	2	3	5	6	7
<b>Diferencia</b>	1-2 = -1	2-3 = -1	4-5 = -1	5-6 = -1	6-7 = -1
<b>Li</b>	-1-0 = -1	-1-(-1) = 0	-1-(-1) = 0	-1-(-1) = 0	-1-(-1) = 0

5 letras en común A, L, D, I y N

$$CK = 6 \cdot ((-1)^2 + (0)^2 + (0)^2 + (0)^2 + (0)^2) / (5 \cdot (5^2 - 1))$$

$$CK = 6 \cdot (1) / 5 \cdot 24$$

$$CK = 1 / 20$$

$$CK = 0,05$$

En este ejemplo, se observa que si se presenta la misma secuencia de letras en las 2 palabras (LADIN en nuestro ejemplo), la diferencia entre las 2 secuencias se tiene en cuenta una sola vez.

Si las 2 palabras, la WT y la WR, son idénticas, el rango de Kendall es igual a 0.

### Tercer paso: cálculo del índice de similitud

índice de similitud = CK + LC + LNC + DL

Se **selecciona** una palabra de referencia si el valor del índice es inferior o igual a:

	CK	LC	DL	Índice de similitud
Longitud de la palabra buscada > 4 caracteres	<= 1,5	<= 0,22	<= 1,5	< 1,2
Longitud de la palabra buscada = 4 caracteres	<= 1,5	<= 0,25	<= 1,26	< 1,2
Longitud de la palabra buscada < 4 caracteres	<= 1	<= 0,34	<= 1,0	< 1,2

Se ordenan las denominaciones de referencia para presentar los resultados por índice de similitud y luego por orden alfabético.

[Fin del Anexo y del documento]