Reconsideration Request Form

Version of 11 April 2013

ICANN's Board Governance Committee is responsible for receiving requests for reconsideration from any person or entity that has been materially affected by any ICANN staff action or inaction if such affected person or entity believes the action contradicts established ICANN policies, or by actions or inactions of the Board that such affected person or entity believes has been taken without consideration of material information. Note: This is a brief summary of the relevant Bylaws provisions. For more information about ICANN's reconsideration process, please visit http://www.icann.org/en/general/bylaws.htm#IV and http://www.icann.org/en/committees/board-governance/.

This form is provided to assist a requester in submitting a Reconsideration Request, and identifies all required information needed for a complete Reconsideration Request. This template includes terms and conditions that shall be signed prior to submission of the Reconsideration Request.

Requesters may submit all facts necessary to demonstrate why the action/inaction should be reconsidered. However, argument shall be limited to 25 pages, double-spaced and in 12 point font.

For all fields in this template calling for a narrative discussion, the text field will wrap and will not be limited.

Please submit completed form to reconsideration@icann.org.

1. Requesters Information

Name: Booking.com B.V.

Address: Contact Information Redacted

Email: Contact Information Redacted

Name: Travel Reservations SRL (formerly, Despegar Online SRL)

Address: Contact Information Redacted

Email: Contact Information Redacted

The Requesters are both represented by:

Name: Flip Petillion, Crowell & Moring LLP

Address: Contact Information Redacted

Email: Contact Information Redacted

Phone Number (optional): Contact Information Redacted

(Note: ICANN will post the Requester's name on the Reconsideration Request page at http://www.icann.org/en/committees/board-governance/requests-for-reconsideration-en.htm. Requestors address, email and phone number will be removed from the posting.)

2. Request for Reconsideration of (check one only):

_X__ Board action/inaction

___ Staff action/inaction

3. Description of specific action you are seeking to have reconsidered.

(Provide as much detail as available, such as date of Board meeting, reference to Board resolution, etc. You may provide documents. All documentation provided will be made part of the public record.)

The Requesters seek reconsideration of both actions and inactions of ICANN's Board of Directors. The specific actions/inactions of the Board are set forth in more detail below, specifically in response to Questions 8 and 10, and relate to the ICANN Board's Resolutions 2015.04.26.14 to 2015.04.26.16, approved on April 26, 2015 and published on April 28, 2015 (hereinafter, the 'Decision'), attached as **Annex 1**.

4. Date of action/inaction:

(Note: If Board action, this is usually the first date that the Board posted its resolution and rationale for the resolution or for inaction, the date the Board considered an item at a meeting.)

On April 28, 2015, the Board published the Decision, which had apparently been taken on April 26, 2015 (**Annex 1**).

5. On what date did you became aware of the action or that action would not be taken?

(Provide the date you learned of the action/that action would not be taken. If more than fifteen days has passed from when the action was taken or not taken to when you learned of the action or inaction, please provide discussion of the gap of time.)

The Requesters learned of the Decision on April 29, 2015.

6. Describe how you believe you are materially affected by the action or inaction:

The first Requester, Booking.com, is the applicant for the '.hotels' gTLD. The second Requester, Travel Reservations, is the applicant for the '.hoteis' gTLD. The Decision impacts the Requesters. In the Decision, the Board directed ICANN's President and CEO, or his designee(s), to move forward with processing the .hotels/.hoteis contention set. It appears that ICANN is unlikely to approve both the application for '.hotels' and the application for '.hoteis'.

This would directly impact the Requesters. As a result, either: one of the Requesters, Booking.com or Travel Reservations, would not have access to its desired gTLD (together with the attendant opportunities to improve their service offer); or both Requesters would be obliged to share the same gTLD, reducing differentiation between them and potentially causing customer confusion.

7. Describe how others may be adversely affected by the action or inaction, if you believe that this is a concern.

The decision to put .hotels and .hoteis in a contention set is not consistent with ICANN's goals of increasing competition and making the domain name system more global and understandable through the use of local languages.

The Requesters are competitors who target the same group of customers. Both offer online search and reservation services, which are free of charge for the customer. Were the Requesters to have access to the separate .hotels or .hoteis gTLDs, this would create new and significant opportunities to further differentiate and identify their services based on clear independent identities. Forcing these two competitors to fight for control of a gTLD, or share a single gTLD, would reduce these opportunities for greater competition in the market and may result in communications to customers becoming more confused.

In the light of customer demand for online travel services and the customer's preference to

compare the prices of multiple online travel agents, separate gTLDs would enable the Requesters to develop their own distinct and reliable platforms for online travel search and reservation services. This would promote competition, to the benefit of Internet users across the globe.

Indeed, the Requesters have every incentive to maintain their current strong and differentiated brand identities and therefore to operate the .hotels and .hoteis gTLDs in a way that ensures the continued distinctiveness of their respective brands.

In addition, Internet users would benefit from having information on hotels in their own/preferred language and accessible through domain names in that language. As the Decision appears to imply that ICANN will only allow either the .hotels or the .hoteis gTLD, this means that either the English language community or the Portuguese language community would be deprived of a gTLD related to hotels in their own language. This is not consistent with ICANN's goal to make the domain name system more global.

As a result, resolving the .hotels/.hoteis contention set by allowing only one of the gTLDs will limit competition and be detrimental to the public interest and the interests of the global Internet user.

8. Detail of Board or Staff Action – Required Information

Staff Action: If your request is in regards to a staff action or inaction, please provide a detailed explanation of the facts as you understand they were provided to staff prior to the action/inaction presented to the staff and the reasons why the staff's action or inaction was inconsistent with established ICANN policy(ies). Please identify the policy(ies) with which the action/inaction was inconsistent. The policies that are eligible to serve as the basis for a Request for Reconsideration are those that are approved by the ICANN Board (after input from the community) that impact the community in some way. When reviewing staff action, the outcomes of prior Requests for Reconsideration challenging the same or substantially similar action/inaction as inconsistent with established ICANN policy(ies) shall be of precedential value.

Board action: If your request is in regards to a Board action or inaction, please provide a detailed explanation of the material information not considered by the Board. If that

information was not presented to the Board, provide the reasons why you did not submit the material information to the Board before it acted or failed to act. "Material information" means any information that are material to the decision.

If your request is in regards to a Board action or inaction that you believe is based upon inaccurate, false, or misleading materials presented to the Board and those materials formed the basis for the Board action or inaction being challenged, provide a detailed explanation as to whether an opportunity existed to correct the material considered by the Board. If there was an opportunity to do so, provide the reasons that you did not provide submit corrections to the Board before it acted or failed to act.

Provide the Required Detailed Explanation here:

(You may attach additional sheets as necessary.)

As will be demonstrated in greater detail below, the ICANN Board (1) failed to consider material information, (2) relied on inaccurate material information in the Decision, and (3) took action in contravention of ICANN's own Articles of Incorporation, Bylaws and Affirmation of Commitments.

I. The ICANN Board failed to consider material information

The ICANN Board evidently did consider the final declaration of March 3, 2015 in the Independent Review proceeding ("IRP") filed by Booking.com ("Final IRP Declaration"). The Requesters appreciate that the ICANN Board took into account the IRP Panel's comments with respect to ways in which the New gTLD Program processes might improve in future rounds.

However, the ICANN Board failed to consider the fact that it had the discretion to improve the New gTLD Program processes and implementation for the current round. The IRP Panel explicitly encouraged the ICANN Board in the Final IRP Declaration to consider whether, notwithstanding the result of the string similarity review of .hotels and .hoteis, approval of both the Requesters' proposed strings would be in the best interest of the Internet community. The Decision and the rationale of the Decision show that ICANN did not consider this material information.

Reading between the lines, the Final IRP Declaration advocated that Booking.com and ICANN resolve the issue amicably. As one of the entities that is most affected by the outcome of the string similarity review, Booking.com has expressed its willingness to engage in a discussion with ICANN, its constituents and the ICANN Board, as to how the issue can be resolved in the best interests of the Internet community. The other entity most affected by the outcome of the string similarity review, Travel Reservations, is also willing to engage in these discussions, and shares the position of Booking.com that the Internet community would best be served by delegating both the .hotels and .hoteis gTLDs.

There are no indications that the ICANN Board considered the Requesters' willingness to engage in these discussions.

There are also no indications that the ICANN Board considered the evidence submitted by Booking.com in the context of the IRP, showing that there is no possible visual confusion. The evidence also showed that ICANN has allowed applications with at least equally serious visual string similarity concerns – such as .parts/.paris, .maif/.mail, .srt/.srl, .vote/.voto and .date/.data (**Annex 2**, p. 11) – to proceed while singling out .hotels/.hoteis. The ICANN Board did not consider this material information when making the Decision.

Finally, the ICANN Board did not consider the fact that it has previously approved changes to the New gTLD Program during its implementation, where those changes were justified by the public interest, according to the ICANN Board. These facts are material to the Decision, as it creates disparate treatment between the Requesters and other applicants advocating the public interest argument (*infra*, Section III).

II. The ICANN Board relied on inaccurate material information

The IRP Panel in the Final IRP Declaration considered that Booking.com was time-barred from raising its objections to the string similarity review. The IRP Panel reasoned that

Booking.com should have objected to the string similarity review process at the time the Guidebook was first implemented.

These findings of the IRP Panel, *i.e.* that Booking.com was time-barred, are flawed:

- The IRP Panel ignored the fact that neither the string similarity review process nor the string confusion objection procedures had been established and implemented in their entirety at the time the Guidebook was adopted. As a result, neither the Requesters nor any other interested party were in a position effectively to challenge these, as yet unfinalized, processes. Indeed, at that time, ICANN still had every opportunity to correctly implement the string confusion objection procedure in accordance with both the Guidebook and the fundamental principles in ICANN's Articles of Incorporation ("AoI") and Bylaws. The opportunity for the Requesters to challenge ICANN's erroneous application of the Guidebook, which was in violation of ICANN's fundamental obligations, only arose when the flaws in ICANN's implementation of the Guidebook became apparent. Therefore, at the time of the adoption of the Guidebook, the Requesters were effectively barred from challenging the Guidebook, because the harm had not yet become manifest.

- Further, to raise an issue at that time would have required the Requesters to reveal that they were contemplating making an application for a new gTLD. This would have encouraged opportunistic applications from third parties seeking to extract monetary value from an application through a private auction.

- The IRP Panel did not draw a distinction between the adoption of the general principles and their subsequent implementation. The IRP Panel limited its review to ICANN's compliance to the letter of the Guidebook. It refrained from reviewing the Board's actions in relation to the implementation of the Guidebook, asserting that the

ICANN Board has ultimate discretion whether or not to intervene in the string similarity review.

Other IRP panels have recognized the inaccuracy of the findings in the Final IRP Declaration with respect to the issue of timing. Even if a decision is made entirely pursuant to the Guidebook, that decision must remain subject to possible review concerning its compliance with ICANN's AoI and Bylaws (*See* Interim Declaration on Emergency Request for Interim Measures of Protection in ICDR Case No. 01-14-0002-1065 (February 12, 2015), para. 79).

However, the Decision shows that the ICANN Board relied on the IRP Panel's inaccurate reasoning concerning the timing of Booking.com's objection, without considering those inaccuracies or the fact that other IRP panels have expressed themselves differently on the issue.

III. The ICANN Board contravened ICANN's Articles of Incorporation, Bylaws and Affirmation of Commitments

The ICANN Board rightfully considered the transparency and fairness issues identified in the Final IRP Declaration. However, it only did so for future new gTLD rounds and not with respect to the Requesters' applications. Instead of resolving the contraventions of ICANN's AoI and Bylaws present in the current new gTLD round, the ICANN Board relied on the erroneous reasoning that Booking.com was time-barred from raising these issues.

The Requesters fail to understand why the ICANN Board did not consider the contraventions of the AoI and Bylaws with respect to the implementation of the string similarity review in the current round. The point is all the stronger in view of the fact that the ICANN Board has previously made changes to the New gTLD Program during implementation. For example, the ICANN Board has introduced a review mechanism for certain specific string confusion objection expert determinations. During the implementation of the New gTLD Program, the ICANN Board also introduced a public interest commitment specification to the standard

registry agreement as well as specific contractual provisions for .brand TLDs. These ICANN Board decisions are clear policy changes, which alter previously established policies.

The Requesters fail to understand why the ICANN Board has approved these policy changes, but has not considered changing the implementation of the string similarity review process in order to bring it into line with ICANN's AoI and Bylaws, not only for the future but also for the current new gTLD round.

9. What are you asking ICANN to do now?

(Describe the specific steps you are asking ICANN to take. For example, should the action be reversed, cancelled or modified? If modified, how should it be modified?)

The Requesters ask ICANN to reverse the decision in which '.hotels' (Application ID 1-1016-75482) and '.hoteis' (Application ID 1-1249-87712) were put in a non-exact match contention set.

ICANN is requested to modify the Decision and to decide that the '.hotels' gTLD, as applied for in the Application with ID 1-1016-75482, can co-exist with the '.hoteis' gTLD, as applied for in the Application with ID 1-1249-87712.

In the event that ICANN will not immediately reverse its decision, the Requesters ask that ICANN engage in conversations with the Requesters, and that a hearing be organized. In addition, ICANN is requested to stay the present reconsideration proceedings with a view to allowing the Requesters to further consider how best to exclude all perceived likelihood of visual confusion.

10. Please state specifically the grounds under which you have the standing and the right to assert this Request for Reconsideration, and the grounds or justifications that support your request.

(Include in this discussion how the action or inaction complained of has resulted in material harm and adverse impact. To demonstrate material harm and adverse impact, the requester must be able to demonstrate well-known requirements: there must be a loss or injury suffered (financial or non-financial) that is a directly and causally connected to the Board or staff action or inaction that is the basis of the Request for Reconsideration. The requestor must be

able to set out the loss or injury and the direct nature of that harm in specific and particular details. The relief requested from the BGC must be capable of reversing the harm alleged by the requester. Injury or harm caused by third parties as a result of acting in line with the Board's decision is not a sufficient ground for reconsideration. Similarly, injury or harm that is only of a sufficient magnitude because it was exacerbated by the actions of a third party is also not a sufficient ground for reconsideration.)

I. The ICANN Board's failure to consider material information harmed Requesters

The Requesters are harmed by the ICANN Board's failure to consider material information; the Requesters have been treated differently from other applicants in the current new gTLD round. There is no justification for this disparate treatment.

In addition, the Decision creates disparate treatment for the Requesters in comparison with applicants in future new gTLD rounds. There is no justification for the ICANN Board accepting the need to comply with its fundamental obligations of ensuring due process, transparency and fairness in a future round, while not accepting this in the current round.

The Decision directly harms the Requesters, as it blocks one of the Requesters' applications for strings that should otherwise be permitted for registration according to ICANN's policy as outlined in the Applicant Guidebook.

In addition, Booking.com has invested significant time and effort in defending its application for .hotels against the unreasoned advice by the string similarity review panel. This advice and ICANN's acceptance of it contravenes ICANN's AoI and Bylaws. As a result of ICANN's acceptance of the advice, the Requesters' respective applications for .hotels and .hoteis have suffered unnecessary delays and are currently experiencing further delays because of the Decision.

The IRP Panel recognized in the Final IRP Declaration that Booking.com's submissions contributed to the public interest. The Requesters expect the ICANN Board not to limit the public interest benefits of Booking.com's contributions to future new gTLD applicants, but

also to allow the Requesters to benefit from these contributions in the current new gTLD round. Otherwise, the parties that are harmed most severely by the lack of due process would be the only ones not to benefit from these contributions.

II. The requested relief reverses most of the harm

Although the requested relief in this Reconsideration Request does not compensate for the lost time, costs and effort, it would reverse most of the harm to the Requesters in that the relief would allow the Requesters to proceed with the safe and secure operation of gTLDs that are relevant to the Requesters' respective businesses.

The Requesters are willing to invest additional time and effort in developing solutions to alleviate the ICANN Board's concerns and to exclude as far as possible any perceived likelihood of visual confusion. The Requesters expect that ICANN contribute in good faith to this process.

11. Are you bringing this Reconsideration Request on behalf of multiple persons or entities? (Check one)

__X__ Yes

____ No

11a. If yes, Is the causal connection between the circumstances of the Reconsideration Request and the harm the same for all of the complaining parties? Explain.

Apart from differences in the time, costs and effort spent in resolving the perceived visual similarity issue, the Requesters' harm is identical, as explained in section 6 above.

Do you have any documents you want to provide to ICANN?

If you do, please attach those documents to the email forwarding this request. Note that all documents provided, including this Request, will be publicly posted at http://www.icann.org/en/committees/board-governance/requests-for-reconsideration-en.htm.

1. ICANN Board's Resolutions 2015.04.26.14 to 2015.04.26.16
2. Expert report showing that .hotels and .hoteis cannot be considered confusingly similar

Terms and Conditions for Submission of Reconsideration Requests

The Board Governance Committee has the ability to consolidate the consideration of Reconsideration Requests if the issues stated within are sufficiently similar.

The Board Governance Committee may dismiss Reconsideration Requests that are querulous or vexatious.

Hearings are not required in the Reconsideration Process, however Requestors may request a hearing. The BGC retains the absolute discretion to determine whether a hearing is appropriate, and to call people before it for a hearing.

The BGC may take a decision on reconsideration of requests relating to staff action/inaction without reference to the full ICANN Board. Whether recommendations will issue to the ICANN Board is within the discretion of the BGC.

The ICANN Board of Director's decision on the BGC's reconsideration recommendation is final and not subject to a reconsideration request.

Fiz Pothion May 13, 2015

Signature Date

crowell moring

Annex 1.

Resources
- About ICANN
- Board
- Accountability & Transparency
- Governance
- Groups
- Business
- Contractual Compliance
- Registrars
- Registries
- Operational Metrics
- Identifier

Approved Board Resolutions | Regular Meeting of the ICANN Board

26 Apr 2015

1. **Consent Agenda:**
   a. **Approval of Board Meeting Minutes**

   b. **Delegation of the .հայ ("hye") domain representing Armenia in Armenian script to the Internet Society of Armenia**
      Rationale for Resolutions 2015.04.26.02 – 2015.04.26.03

   c. **Redelegation of the .BN domain representing Brunei Darussalam to Brunei Darussalam Network Information Centre Sdn Bhd (BNNIC)**
      Rationale for Resolutions 2015.04.26.04 – 2015.04.26.05

   d. **Delegation of السودان. ("Sudan") representing Sudan in Arabic script to Sudan Internet Society**
      Rationale for Resolutions 2015.04.26.06 – 2015.04.26.07

Systems Security, Stability and Resiliency (IS-SSR)

- ccTLDs
- Internationalized Domain Names
- Universal Acceptance Initiative
- Policy
- Public Comment
- Contact
- Help

e. **Appointment of Annual Independent Auditors**
Rationale for Resolution 2015.04.26.08

f. **Next Steps for the EWG Final Report on Next Generation Registration Directory Services**
Rationale for Resolutions 2015.04.26.09-2015.04.26.12

2. **Main Agenda:**

a. **Approval of Minutes**

b. **Consideration of Independent Review Panel's Final Declaration in Booking.com v. ICANN**
Rationale for Resolutions 2015.04.26.14-2015.04.26.16

c. **Reserve Fund Release – USG IANA Stewardship Transition Costs**
Rationale for Resolution 2015.04.26.17

d. **IT Services Contracting**
Rationale for Resolutions 2015.04.26.18 – 2015.04.26.19

e. **SO/AC FY16 Additional Budget Requests for FY16**
Rationale for Resolution 2015.04.26.20

f. **ICANN Five-Year Operating Plan**
Rationale for Resolution 2015.04.26.21

g. **Structural Improvements Committee Chair**
Rationale for Resolution 2015.04.26.22

h. **Funding Digital Services platforms and code-base review**
Rationale for Resolutions 2015.04.26.23-2015.04.26.24

i. **Investment management – Adjustments to the account structure**
Rationale for Resolutions 2015.04.26.25-2015.04.26.26

j. **AOB**

1. Consent Agenda:

a. Approval of Board Meeting Minutes

Resolved (2012.04.26.01), the Board approves the minutes of the 12 February 2015 Meeting of the ICANN Board.

b. Delegation of the .հայ ("hye") domain representing Armenia in Armenian script to the Internet Society of Armenia

Resolved (2015.04.26.02), as part of the exercise of its responsibilities under the IANA Functions Contract, ICANN has reviewed and evaluated the request to delegate the .հայ IDN country-code top-level domain to Internet Society. The documentation demonstrates that the proper procedures were followed in evaluating the request.

Resolved (2015.04.26.03), the Board directs that pursuant to Article III, Section 5.2 of the ICANN Bylaws, that certain portions of the rationale not appropriate for public distribution within the resolutions, preliminary report or minutes at this time due to contractual obligations shall be withheld until public release is allowed pursuant to those contractual obligations.

Rationale for Resolutions 2015.04.26.02 – 2015.04.26.03

Why the Board is addressing the issue now?

In accordance with the IANA Functions Contract, the ICANN staff has evaluated a request for ccTLD delegation, and is presenting its report to the Board for review. This review by the Board is intended to ensure that ICANN staff has followed the proper procedures.

What is the proposal being considered?

The proposal is to approve a request to the IANA Department to assign the sponsoring organization (also known as the manager or trustee) of the .հայ country-code top-level domains to Internet Society of Armenia.

Which stakeholders or others were consulted?

In the course of evaluating a delegation application, ICANN staff consults with the applicant and other interested parties. As part of the application process, the applicant needs to describe consultations that were performed within the country concerning the ccTLD, and their applicability to their local Internet community.

What concerns or issues were raised by the community?

Staff are not aware of any significant issues or concerns raised by the community in relation to this request.

What significant materials did the Board review?

[*REDACTED – SENSITIVE DELEGATION INFORMATION*]

What factors the Board found to be significant?

The Board did not identify any specific factors of concern with this request.

Are there positive or negative community impacts?

The timely approval of country-code domain name managers that meet the various public interest criteria is positive toward ICANN's overall mission, the local communities to which country- code top-level domains are designated to serve, and responsive to ICANN's obligations under the IANA Functions Contract.

Are there financial impacts or ramifications on ICANN (strategic plan, operating plan, budget); the community; and/or the public?

The administration of country-code delegations in the DNS root zone is part of the IANA functions, and the delegation action should not cause any significant variance on pre-planned expenditure. It is not the role of ICANN to assess the financial impact of the internal operations of country-code top-level domains within a country.

Are there any security, stability or resiliency issues relating to the DNS?

ICANN does not believe this request poses any notable risks to security, stability, or resiliency. This is an Organizational Administrative Function not requiring public comment.

c. Redelegation of the .BN domain representing Brunei Darussalam to Brunei Darussalam Network Information Centre Sdn Bhd (BNNIC)

Resolved (2015.04.26.04), as part of the exercise of its responsibilities under the IANA Functions Contract, ICANN has reviewed and evaluated the request to redelegate the .BN country-code top-level domain to Brunei Darussalam Network Information Centre Sdn Bhd (BNNIC). The documentation demonstrates that the proper procedures were followed in evaluating the request.

Resolved (2015.04.26.05), the Board directs that pursuant to Article III, Section 5.2 of the ICANN Bylaws, that certain portions of the rationale not appropriate for public distribution within the resolutions, preliminary report or minutes at this time due to contractual obligations shall be withheld until public release is allowed pursuant to those contractual obligations.

Rationale for Resolutions 2015.04.26.04 – 2015.04.26.05

Why the Board is addressing the issue now?

In accordance with the IANA Functions Contract, the ICANN staff has evaluated a request for ccTLD redelegation and is presenting its report to the Board for review. This review by the Board is intended to ensure that ICANN staff has followed the proper procedures.

What is the proposal being considered?

The proposal is to approve a request to the IANA department to change the sponsoring organization (also known as the manager or trustee) of the .BN country-code top-level domain to Brunei Darussalam Network Information Centre Sdn Bhd (BNNIC).

Which stakeholders or others were consulted?

In the course of evaluating a delegation application, ICANN staff consults with the applicant and other interested parties. As part of the application process, the applicant needs to describe consultations that were performed within the country concerning the ccTLD, and their applicability to their local Internet community.

What concerns or issues were raised by the community?

Staff are not aware of any significant issues or concerns raised by the community in relation to this request.

What significant materials did the Board review?

[*REDACTED – SENSITIVE DELEGATION INFORMATION*]

What factors the Board found to be significant?

The Board did not identify any specific factors of concern

with this request.

Are there positive or negative community impacts?

The timely approval of country-code domain name managers that meet the various public interest criteria is positive toward ICANN's overall mission, the local communities to which country-code top-level domains are designated to serve, and responsive to ICANN's obligations under the IANA Functions Contract.

Are there financial impacts or ramifications on ICANN (strategic plan, operating plan, budget); the community; and/or the public?

The administration of country-code delegations in the DNS root zone is part of the IANA functions, and the redelegation action should not cause any significant variance on pre-planned expenditure. It is not the role of ICANN to assess the financial impact of the internal operations of country-code top-level domains within a country.

Are there any security, stability or resiliency issues relating to the DNS?

ICANN does not believe this request poses any notable risks to security, stability or resiliency.

This is an Organizational Administrative Function not requiring public comment.

d. Delegation of سودان. ("Sudan") representing Sudan in Arabic script to Sudan Internet Society

Resolved (2015.04.26.06), as part of the exercise of its responsibilities under the IANA Functions Contract, ICANN has reviewed and evaluated the request to delegate the سودان country-code top-level domain to Sudan Internet Society. The documentation demonstrates that the proper procedures were followed

in evaluating the request.

Resolved (2015.04.26.07), the Board directs that pursuant to Article III, Section 5.2 of the ICANN Bylaws, that certain portions of the rationale not appropriate for public distribution within the resolutions, preliminary report or minutes at this time due to contractual obligations, shall be withheld until public release is allowed pursuant to those contractual obligations.

Rationale for Resolutions 2015.04.26.06 – 2015.04.26.07

Why the Board is addressing the issue now?

In accordance with the IANA Functions Contract, the ICANN staff has evaluated a request for ccTLD delegation and is presenting its report to the Board for review. This review by the Board is intended to ensure that ICANN staff has followed the proper procedures.

What is the proposal being considered?

The proposal is to approve a request to the IANA department to create the country-code top-level domain and assign the role of sponsoring organization (also known as the manager or trustee) to Sudan Internet Society.

Which stakeholders or others were consulted?

In the course of evaluating a delegation application, ICANN staff consults with the applicant and other interested parties. As part of the application process, the applicant needs to describe consultations that were performed within the country concerning the ccTLD, and their applicability to their local Internet community.

What concerns or issues were raised by the community?

Staff are not aware of any significant issues or concerns

raised by the community in relation to this request.

What significant materials did the Board review?

[*REDACTED – SENSITIVE DELEGATION INFORMATION*]

What factors the Board found to be significant?

The Board did not identify any specific factors of concern with this request.

These evaluations are responsive to the appropriate criteria and policy frameworks, such as "Domain Name System Structure and Delegation" (RFC 1591) and "GAC Principles and Guidelines for the Delegation and Administration of Country Code Top Level Domains".

As part of the process established by the IANA Functions Contract, the "Delegation and Redelegation Report" will be published at http://www.iana.org/reports.

What factors the Board found to be significant?

The Board did not identify any specific factors of concern with this request.

Are there financial impacts or ramifications on ICANN (strategic plan, operating plan, budget); the community; and/or the public?

The administration of country-code delegations in the DNS root zone is part of the IANA functions, and the delegation action should not cause any significant variance on pre-planned expenditure. It is not the role of ICANN to assess the financial impact of the internal operations of country-code top-level domains within a country.

Are there any security, stability or resiliency issues relating to the DNS?

ICANN does not believe this request poses any notable risks to security, stability or resiliency. This is an Organizational Administrative Function not requiring public comment.

e. Appointment of Annual Independent Auditors

Whereas, Article XVI of the ICANN Bylaws (http://www.icann.org/general/bylaws.htm) requires that after the end of the fiscal year, the books of ICANN must be audited by certified public accountants, which shall be appointed by the Board.

Whereas, the Board Audit Committee has discussed the engagement of the independent auditor for the fiscal year ending 30 June 2015, and has recommended that the Board authorize the President and CEO, or his designee(s), to take all steps necessary to engage BDO LLP and BDO member firms.

Resolved (2015.04.26.08), the Board authorizes the President and CEO, or his designee(s), to take all steps necessary to engage BDO LLP and BDO member firms as the auditors for the financial statements for the fiscal year ending 30 June 2015.

Rationale for Resolution 2015.04.26.08

The audit firm BDO LLP and BDO member firms were engaged for the annual independent audit of the fiscal year end 30 June 2014 as a result of an extensive RFP process. Based on the report from staff and the Audit Committee's evaluation of the work performed, the committee has unanimously recommended that the Board authorize the President and CEO, or his designee(s), to take all steps necessary to engage BDO LLP and BDO member firms as ICANN's annual independent auditor for the fiscal year ended 30 June 2015 for any annual independent audit requirements in any jurisdiction.

The engagement of an independent auditor is in fulfillment of ICANN's obligations to undertake an audit of ICANN's financial statements. This furthers ICANN's accountability to its Bylaws and processes, and the results of the independent auditors work will be publicly available. There is a fiscal impact to the engagement that has already been budgeted. There is no impact on the security or the stability of the DNS as a result of this appointment.

This is an Organizational Administrative Function not requiring public comment.

f. Next Steps for the EWG Final Report on Next Generation Registration Directory Services

Whereas, in 2012, the Board adopted a two-pronged approach to address the recommendations of the WHOIS Review Team, calling for ICANN to (i) continue to fully enforce existing consensus policy and contractual conditions relating to WHOIS, and (ii) create an expert working group to determine the fundamental purpose and objectives of collecting, maintaining and providing access to gTLD registration data, to serve as a foundation for a Board-initiated GNSO policy development process (PDP).

Whereas, in 2014, the Expert Working Group on Next Generation Registration Directory Services (EWG) delivered its Final Report [PDF, 5.12 MB] to the Board with its recommended model and principles to serve as the foundation for the GNSO PDP.

Whereas, an informal group of Board members and GNSO Councilors collaborated and developed a proposed framework [PDF, 612 KB] to provide guidance to the GNSO PDP for the examination of the EWG's recommended models and principles for the next generation registration directory services to replace WHOIS.

Resolved (2015.04.26.09), the Board thanks the EWG for the significant effort and work exerted that produced the proposed model for a next generation registration directory services as reflected in its Final Report [PDF, 5.12 MB].

Resolved (2015.04.26.10), the Board reaffirms its request for a Board-initiated GNSO policy development process to define the purpose of collecting, maintaining and providing access to gTLD registration data, and consider safeguards for protecting data, using the recommendations in the Final Report [PDF, 5.12 MB] as an input to, and, if appropriate, as the foundation for a new gTLD policy;

Resolved (2015.04.26.11), the Board directs that a new Preliminary Issue Report that follows this framework [PDF, 612 KB] be prepared and delivered to the GNSO;

Resolved (2015.04.26.12), the Board commits to forming a group of Board members that will (i) liaise with the GNSO on the policy development process to examine the EWG's recommended model and propose policies to support the creation of the next generation registration directory services, and (ii) oversee the implementation of the remaining projects arising from the Action Plan [PDF, 119 KB] adopted by the Board in response to the WHOIS Review Team's recommendations. The Board directs the Board Governance Committee to begin the process for identifying a recommendation of a slate of Board members to do this work.

Rationale for Resolutions 2015.04.26.09-2015.04.26.12

Why the Board is addressing the issue?

This resolution continues the Board's attention to the implementation of the Action Plan [PDF, 119 KB] adopted by the Board in response to the WHOIS Review Team's recommendations [PDF, 5.12 MB]. The resolution

adopted today adopts a framework [PDF, 612 KB] to conduct a board-initiated GNSO policy development process to refine the purpose of collecting, maintaining and providing access to gTLD registration data, and consider safeguards for protecting data, using the recommendations of the Expert Working Group's Final Report [PDF, 5.12 MB] as an input to, if appropriate, to serve as the foundation for a new gTLD policy.

What is the proposal being considered?

Under the Affirmation of Commitments (AoC), ICANN is committed to enforcing its existing policy relating to WHOIS (subject to applicable laws), which "requires that ICANN implement measures to maintain timely, unrestricted and public access to accurate and complete WHOIS information…." The AoC obligates ICANN to organize no less frequently than every three years a community review of WHOIS policy and its implementation to assess the extent to which WHOIS policy is effective and its implementation meets the legitimate needs of law enforcement and promotes consumer trust. Under this timeline, the second WHOIS Review Team is to be convened in late 2015.

In 2012, in response to the recommendations of the first WHOIS Review Team, the Board adopted a two-prong approach that simultaneously directed ICANN to (1) implement improvements to the current WHOIS system based on the Action Plan [PDF, 119 KB] that was based on the recommendations of the WHOIS Review Team, and (2) launch a new effort, achieved through the creation of the Expert Working Group, to focus on the purpose and provision of gTLD directory services, to serve as the foundation of a Board-initiated GNSO policy development process (PDP).

The Expert Working Group's Final Report [PDF, 5.12 MB] contains a proposed model and detailed principles to serve as the foundation for a PDP to support the creation of the next generation registration directory services to

replace WHOIS. This Final Report [PDF, 5.12 MB] contains over 160 pages of complex principles and recommendations to be considered in the GNSO PDP. In order to effectively manage the PDP on such a large scale, an informal group of Board members and GNSO councilors collaborated to develop the framework [PDF, 612 KB] approved today.

What factors did the Board find to be significant?

The complex nature of the EWG's recommendations, along with the contentiousness nature of the WHOIS issue in the ICANN community over the last ten+ years, calls for a very structured approach to conducting a policy development process of this magnitude. The framework [PDF, 612 KB] provides guidance to the GNSO on how to best structure the resulting PDP(s) for success – that is, it proposes a process which leads to new policies defining the purpose of gTLD registration data and improving accuracy, privacy, and access to that data.

This framework [PDF, 612 KB] creates a 3-phased approach to conducting the PDP, with Phase 1 focusing on definition of the policy requirements, Phase 2 focusing on the functional design elements of the policy, and Phase 3 focusing on implementation of the policies and providing guidance during an expected transition period during which the legacy WHOIS system and the next generation registration directory services may coexist and both operational at the same time. The Board believes that following the framework [PDF, 612 KB] will ensure that the PDP will properly address the many significant issues and interdependencies that require consideration in order to support the creation of the next generation registration directory services.

The Board recognizes that additional resources may be needed for the conduct of this unique policy development process. The Board commits to reviewing the GNSO's proposed plan and schedule, as well as Staff's

assessment of the resources required to implement this proposed plan, and to supporting appropriate resourcing for the conduct of this PDP.

In addition, the Board believes that the importance of the WHOIS issue, along with the breadth and scope of the many WHOIS activities currently under way, support the need for a designated group of Board members dedicated to overseeing the entire WHOIS Program, including working with the community on the GNSO PDP, and any future transition to a next generation registration directory services that may emerge following the GNSO PDP. Community members participating in the informal Board-GNSO Council effort to develop the framework for the PDP also requested the Board's continued involvement in this effort.

What significant materials did the Board review?

The Board reviewed the EWG Final Report [PDF, 5.12 MB], the framework [PDF, 612 KB] developed through the informal collaboration between the Board and the GNSO Council, and the Briefing Papers submitted by Staff.

Are there fiscal impacts or ramifications on ICANN (strategic plan, operating plan, or budget)?

The initiation of focused work on WHOIS and the creation of policies to support the next generation of registration directory services are expected to have an impact on financial resources as the research and work progresses. Due to the expected complexity of this PDP, there is a potential that this PDP may have higher resource needs than other PDPs, though the full extent of those resource needs are not fully understood, particularly as to the scope of those resources in comparison to the resources proposed for allocation within the upcoming fiscal year for this effort. The Board commits to reviewing staff's assessment of resources for the conduct of this PDP (after there is a plan and

schedule developed) with a view towards providing appropriate resourcing for the conduct of this PDP.

Are there any security, stability or resiliency issues relating to the DNS?

This action is not expected to have an immediate impact on the security, stability or resiliency of the DNS, though the outcomes of this work may result in positive impacts.

Is public comment required prior to Board action?

As this is a continuation of prior Board actions, public comment is not necessary prior to adoption. A public comment period will be commenced, as required by the ICANN Bylaws, once the Preliminary Issue Report is published by Staff, thereby allowing the framework [PDF, 612 KB] approved today to be adjusted as appropriate prior to delivery of the Final Issue Report to the GNSO.

2. Main Agenda:

a. Approval of Minutes

Resolved (2012.04.26.13), the Board approves the minutes of the 11 February 2015 Meeting of the ICANN Board.

b. Consideration of Independent Review Panel's Final Declaration in Booking.com v. ICANN

Whereas, on 3 March 2015, an Independent Review Panel ("Panel") issued an advisory Final Declaration in the Independent Review proceeding ("IRP") filed by Booking.com (the "Final Declaration").

Whereas, Booking.com specifically challenged the determination of the String Similarity Panel ("SSP") to place .hotels and .hoteis in contention and the refusal of the Board to revise that determination, as well as the conduct of the Board in adopting and implementing the

entire string similarity review process.

Whereas, the Panel denied Booking.com's IRP request because the Panel determined that "Booking.com failed to identify any instance of Board action or inaction or ICANN staff or a third party (such as the ICC, acting as SSP), that could be considered to be inconsistent with ICANN's Articles of Incorporation or Bylaws or with the policies and procedures established in the Guidebook." (https://www.icann.org/en/system/files/files/final-declaration-03mar15-en.pdf [PDF, 4.76 MB].)

Whereas, while ruling in ICANN's favor, the Panel expressed sympathy for Booking.com insofar as the IRP Panel suggests that there could be future improvements to the transparency of processes developed within the New gTLD Program, and the Board appreciates the IRP Panel comments with respect to ways in which the New gTLD Program processes might improve in future rounds.

Whereas, in accordance with Article IV, section 3.21 of ICANN's Bylaws, the Board has considered the Panel's Final Declaration.

Resolved (2015.04.26.14), the Board accepts the following findings of the Independent Review Panel's Final Declaration that: (1) Booking.com has failed to identify any instance of Board action or inaction, or any action or inaction of ICANN staff or any third party (such as the ICC, acting as SSP), that could be considered to be inconsistent with ICANN's Articles of Incorporation or Bylaws or with the policies and procedures established in the Guidebook, including the challenged actions of the Board (or any staff or third party) in relation to what Booking.com calls the implementation and supervision of the string similarity review process generally, as well as the challenged actions of the Board (or any staff or third party) in relation to the string similarity review of resulting in the placement of .hotels and .hoteis in contention; (2) the string similarity review performed in the case of

.hotels was not inconsistent with ICANN's Articles of Incorporation or Bylaws or with the policies and procedures established in the Guidebook; (3) the time to challenge the Board's adoption and implementation of specific elements of the New gTLD Program, including the string similarity review process has long since passed; and (4) each party shall bear its own IRP costs.

Resolved (2015.04.26.15), the Board directs the President and CEO, or his designee(s), to move forward with processing of the .hotels/.hoteis contention set.

Resolved (2015.04.26.16), the Board directs the President and CEO, or his designee(s), to ensure that the ongoing reviews of the New gTLD Program take into consideration the following issues raised by the Panel in the Final Declaration regarding transparency and fairness:

- "The Guidebook provides no means for applicants to provide evidence or make submissions to the SSP (or any other ICANN body) and to be fully "heard" on the substantive question of the similarity of their applied-for gTLD strings to others."
- "[T]he process as it exists does [not] provide for gTLD applicants to benefit from the sort of procedural mechanisms - for example, to inform the SSP's review, to receive reasoned determinations from the SSP, or to appeal the merits of those determinations."

Rationale for Resolutions 2015.04.26.14-2015.04.26.16

Booking.com filed a request for an Independent Review Proceeding (IRP) challenging the ICANN Board's handling of Booking.com's application for .hotels, including the determination of the String Similarity Panel (SSP) to place .hotels and .hoteis in contention and the refusal of the Board to revise that determination.

Booking.com also challenged the conduct of the Board in the setting up, implementation, and supervision and review of the entire string similarity review process. On 3 March 2015, the IRP Panel (Panel), comprised of three Panelists, issued its Final Declaration. After consideration and discussion, pursuant to Article IV, Section 3.21 of the ICANN Bylaws, the Board adopts the findings of the Panel, which are summarized below, and can be found in full at https://www.icann.org/en/system/files/files/final-declaration-03mar15-en.pdf [PDF, 4.76 MB].

The Panel found that it was charged with "objectively" determining, whether or not the Board's actions are in fact consistent with the Articles, Bylaws, and Guidebook, thereby requiring that the Board's conduct be appraised independently, and without any presumption of correctness. The Panel agreed with ICANN that in determining the consistency of the Board action with the Articles, Bylaws, and Guidebook, the Panel is neither asked to, nor allowed to, substitute its judgment for that of the Board.

Using the applicable standard of review, the Panel found that objectively there was not an inconsistency with the Articles, Bylaws and Guidebook, noting that "the established process was followed in all respects" concerning the process followed by the String Similarity Panel and the BGC's [Board Governance Committee] handling of Booking.com's reconsideration request." (Final Declaration, https://www.icann.org/en/system/files/files/final-declaration-03mar15-en.pdf [PDF, 4.76 MB], at p. 41.)

Specifically, the Panel concluded:

> 144. Booking.com has failed to identify any instance of Board action or inaction, including any action or inaction of ICANN staff or a third party (such as ICC, acting as the SSP) that could be considered to be inconsistent with

> ICANN's Articles of Incorporation or Bylaws or with the policies and procedures established in the Guidebook. This includes the challenged actions of the Board (or any staff or third party) in relation to what Booking.com calls the implementation and supervision of the string similarity review process generally, as well as the challenged actions of the Board (or any staff or third party) in relation to the string similarity review of .hotels in particular.
>
> 145. More particularly, the Panel finds that the string similarity review performed in the case of .hotels was not inconsistent with the Articles or Bylaws or with what Booking.com refers to as the "applicable rules" as set out in the Guidebook.
>
> 146. To the extent that the Board's adoption and implementation of specific elements of the new gTLD Program and Guidebook, including the string similarity review process, could potentially be said to be inconsistent with the principles of transparency or fairness that underlie ICANN's Articles and Incorporation and Bylaws (which the Panel does not say is the case), the time to challenge such action has long since passed.

(*Id.* at pp. 42-43.) Accordingly, the Panel declared ICANN to be the prevailing party. (*See id.* at ¶ 152, p. 43.)

The Panel acknowledged certain legitimate concerns regarding the string similarity review process raised by Booking.com, which concerns the Panel noted were shared by some members of the NGPC. Most notably, the IRP Panel noted that while the String Similarity Review Process, as it exists does not allow for some procedural appeal mechanism, "[a]s to whether they should be, it is not our place to express an opinion, though we note that such additional mechanisms surely

would be consistent with the principles of transparency and fairness." (Id. at ¶ 128, p. 37.)

The Board appreciates the IRP Panel comments with respect to ways in which the New gTLD Program processes might improve in future rounds. ICANN will take the lessons learned from this IRP and apply it towards its ongoing assessments of the ways in which it can improve upon its commitments to accountability and transparency. In particular, the Board will include the following concerns expressed by the Panel in its review of the New gTLD Program for the next round:

- "The Guidebook provides no means for applicants to provide evidence or make submissions to the SSP (or any other ICANN body) and to be fully "heard" on the substantive question of the similarity of their applied-for gTLD strings to others."
- "[T]he process as it exists does [n]ot provide for gTLD applicants to benefit from the sort of procedural mechanisms - for example, to inform the SSP's review, to receive reasoned determinations from the SSP, or to appeal the merits of those determinations.

This action will have no financial impact on the organization and no direct impact on the security, stability or resiliency of the domain name system.

This is an Organizational Administrative function that does not require public comment

c. Reserve Fund Release – USG IANA Stewardship Transition Costs

Whereas, the Board approved the FY15 Operating Plan and Budget, which includes an amount of US$7 million for costs to be incurred related to the USG IANA Stewardship Transition initiative, which was expected to be funded by the Reserve Fund.

Whereas, ICANN is incurring ongoing costs to support the work of the ICANN Community in relation to the USG IANA Stewardship Transition initiative.

Whereas, the Board Finance Committee has recommended that the Board approve the release of funds from the Reserve Fund to cover costs incurred in FY15 related to the USG IANA Stewardship Transition initiative in an amount not to exceed US$7 million, and the Board agrees.

Resolved (2014.04.26.17), the Board authorizes the President and CEO, or his designee(s), to withdraw funds from the Reserve Fund to cover costs incurred in FY15 related to the USG IANA Stewardship Transition initiative in an amount not to exceed US$7 million.

Rationale for Resolution 2015.04.26.17

The USG IANA Stewardship Transition initiative is a major initiative to which the ICANN Community as a whole is dedicating a significant amount of time and resources. ICANN's supporting the Community in its work towards a successful completion of the project (including both the USG IANA Stewardship transition proposal development and accountability work) is critical for ICANN.

Considering its exceptional nature and the significant amount of costs anticipated to be incurred, the funding of this project could not be provided through the ICANN annual operating revenue. Accordingly, when the Board approved the FY15 Operating Plan and Budget, it included the anticipated funding of the project costs (US$7 million) through a corresponding withdrawal from the Reserve Fund.

The withdrawals from the Reserve Fund for the FY15 costs associated with the USG Transition Initiative will be done twice, once for the actual costs incurred from 1 July 2014 – 31 December 2014, and once for the actual costs

incurred from 1 January 2015 – 30 June 2015. The first withdrawal will be for US$1,454,287, representing US$471,438 in personnel costs, US$548,247 in travel and meeting costs, US$352,164 in professional services, and US$82,439 in administrative costs.

As costs are incurred during FY15 for this project, ICANN is proceeding with the planned withdrawals of funds from the Reserve Fund to cover the actual costs incurred in FY15 related to USG IANA Stewardship Transition initiative, up to the amount of US$7 million included in the Board approved FY15 Operating Plan and Budget.

This is an Organizational Administrative Function that does not require public comment at this stage. In particular, the anticipated costs of US$7 million was included in the FY15 Operating Plan and Budget that was subject to public comment before it was approved by the Board.

d. IT Services Contracting

Whereas, ICANN sources IT services from multiple different vendors and wish to consolidate its sourcing of such services to improve efficiency, quality and costs.

Whereas, ICANN staff has undergone an extensive request for proposal process involving 28 potential service providers, which led, after multiple reviews, demonstrations and interviews to the identification of one preferred candidate, Zensar.

Whereas, ICANN staff has undergone further due diligence of Zensar by organizing pilot projects for approximately four months to determine the effective ability to obtain timely quality services from Zensar, which have proven highly conclusive.

Whereas, ICANN staff considers that Zensar has demonstrated the ability to provide ongoing services and project development support durably.

Resolved (2015.04.26.18), the Board authorizes the President and CEO, or his designee(s), to take all actions necessary to contract with, make payments to, and carry out any additional necessary actions with Zensar for a period of up to three years, involving expenses of up to [*amount redacted for negotiation purposes*].

Resolved (2015.04.26.19), specific items within this resolution shall remain confidential for negotiation purposes pursuant to Article III, section 5.2 of the ICANN Bylaws until the President and CEO determines that the confidential information may be released.

Rationale for Resolutions 2015.04.26.18 – 2015.04.26.19

ICANN has been using the services of multiple vendors for its IT related needs, either for ongoing activities or for development projects. The management of multiple vendors is inefficient and generally leads to a higher cost for the value of services received.

ICANN staff has investigated alternative solutions to obtain the IT services that it requires, and the solution of obtaining several services on a long-term basis from a single external vendor with a knowledgeable and competent pool of resources is the preferred approach.

ICANN staff has therefore conducted an extensive request for proposal (RFP) process by defining the list of potential services it requires, obtaining proposals from 28 different vendors, conducting in-depth reviews, selecting a shortlist of five capable firms, interviewing each of the five firms, identifying two shortlisted candidates, and conducting deep-dive analyses of the two organizations to ultimately select Zensar as the preferred candidate. ICANN staff then conducted several pilot projects with Zensar to establish through live services and projects the ability of the company to put in place the adequate resources to provide timely quality services.

This extensive selection and testing process has provided a high confidence that Zensar is a capable partner for a durable period and ICANN Staff has recommended to engage the services of Zensar for a period of three years, up to [*amount redacted for negotiation purposes*].

This is an Organizational Administrative Function that does not require public comment.

e. SO/AC FY16 Additional Budget Requests for FY16

Whereas, prior discussions between community members and ICANN staff members identified the need for an earlier decision on the funding of additional budget requests from ICANN's Supporting Organizations (SO) and Advisory Committees (AC).

Whereas, the staff created an SO/AC additional budget requests process, to collect, review and submit for Board approval funding requests from the SOs and ACs.

Whereas, in accordance with the process, requests were submitted by the ICANN Community by the set deadline, and were reviewed by a panel of staff members representing the Policy, Stakeholder Engagement and Finance departments.

Whereas, the staff panel recommended the approval of requests representing $657,300 for approval.

Whereas the Board Finance Committee, reviewed the process followed and the staff's proposal, and has recommended that the Board approve staff's recommendation.

Resolved (2015.04.26.20), the Board approves committing $657,300 during Fiscal Year 2016 to cover the costs associated with the adopted SO/AC additional budget requests.

Rationale for Resolution 2015.04.26.20

Approving commitments within the FY2016 budget in advance is a reasonable accommodation within the established budget approval process and timeline in order to facilitate the work of the ICANN community and of the ICANN staff, and does not create additional expenses. The amount of the committed expenses resulting from this resolution is considered sufficiently small to not require that funding resources are specifically identified and approved by the Board. Information on the process for consideration of additional budget requests from ICANN's Supporting Organizations (SO) and Advisory Committees (AC) is available here. The list of FY2016 requests received and the disposition of the requests is available here [DOCX, 145 KB].

There is no anticipated impact from this decision on the security, stability and resiliency of the domain name system as a result of this decision.

The approval process is an Organizational Administrative process that has already been subject to significant input from the community.

f. ICANN Five-Year Operating Plan

Whereas, the ICANN Five-Year Operating Plan provides: (i) a five-year planning calendar; (ii) strategic goals with corresponding key performance indicators; (iii) dependencies; (iv) a five-year phasing; (v) a list of portfolios; and (vi) a five-year financial model.

Whereas, together with the ICANN Five-Year Strategic Plan, the ICANN Five-Year Operating Plan will serve as a foundation for the annual operating plans and budgets.

Whereas, the ICANN Five-Year Operating Plan for FY16-FY20 is the result of an extensive, collaborative, bottom-up, multistakeholder and multilingual process using the Board adopted Five-Year Strategic Plan

FY16-FY20 as its foundation.

Whereas, the ICANN Five-Year Operating Plan will be maintained and updated on an annual basis per ICANN's planning process.

Resolved (2015.04.26.21), the Board herby adopts the ICANN Five-Year Operating Plan for FY2016 – FY2020.

Rationale for Resolution 2015.04.26.21

As a new element of ICANN's planning process, the ICANN Five-Year Operating Plan for FY16-FY20 complements the Five-Year Strategic Plan, will guide ICANN's activities for the next five years, and will inform ICANN's annual operating plans and budgets.

With the focus to provide the public with more insight and advance ICANN's accountability and transparency, the Five-Year Operating Plan sets forth details for each Strategic Objective and Goal – portfolios of ICANN activities, key operational success factors (outcomes), key performance indicators (measurements), key dependencies, and phasing over the five years (at the Goal level); and is completed by a five-year financial model, which describes the principles and approach to ensure financial accountability and sustainability in achieving the ICANN mission.

The ICANN Five-Year Operating Plan for FY16-FY20 is the result of a collaborative and bottom-up multistakeholder process, which included extensive public input. Public comments were sought from 11 November 2014 to 4 January 2015. Also, the Community discussions at ICANN 52, Singapore, involving ICANN's Supporting Organizations, Stakeholder Groups, Constituencies, and Advisory Committees, have further refined the ICANN Five-Year Operating Plan for FY16-FY20.

Adopting the ICANN Five-Year Operating Plan will be

advantageous to all stakeholders and the entire ICANN community. This decision itself will have no specific fiscal impact that is not, or will not be, anticipated through the annual Operating Plan and Budgets going forward for the next five years. Further, this action will have no direct impact on the security and stability of the domain name system.

This is an Organization Administrative Function that has already been subject to lengthy public comment, as note above.

g. Structural Improvements Committee Chair

Whereas, Ray Plzak is a member of the Board and current Chair of the Structural Improvements Committee (SIC).

Whereas, Mr. Plzak's current term on the Board expires at the conclusion of the Annual General Meeting in October 2015, and Mr. Plzak is not seeking another term.

Whereas, Rinalia Abdul Rahim is a current member of the Board and member of the SIC.

Whereas, to facilitate the smooth transition of leadership of the SIC at the expiration of Mr. Plzak's term, the BGC recommended that the Board immediately appoint Rinalia Abdul Rahim as the Chair of the SIC and retain Ray as a member of the SIC.

Resolved (2015.04.26.22), the Board appoints Rinalia Abdul Rahim as the Chair of the Structural Improvements Committee (SIC) and retains Ray Plzak a member of the SIC effective immediately.

Rationale for Resolution 2015.04.26.22

The Board is committed to facilitating a smooth transition in the leadership of the Structural Improvements Committee (SIC) when Ray Plzak's term on the Board expires at the conclusion of the Annual General Meeting

in October 2015. In light of the upcoming expiration of his term on the Board, Mr. Plzak suggested that he step down now in order to provide for a transition period to a new SIC Chair while he is still on the Board. As the Board Governance Committee (BGC) is tasked with recommending committee assignments, the BGC discussed Mr. Plzak's proposal and has recommended that the Board appoint Rinalia Abdul Rahim as the new SIC Chair effectively immediately.

The Board agrees with Mr. Plzak and the BGC that it is appropriate to appoint Ms. Abdul Rahim as the Chair of the SIC, effectively immediately, and retain Mr. Plzak as a member of the SIC until the end of his term.

This action will have no financial impact on the organization and no direct impact on the security, stability, or resiliency of the domain name system.

This is an Organizational Administrative function that does not require public comment.

h. Funding Digital Services platforms and code-base review

Whereas, staff has compiled a complete list of all digital services offered by ICANN to its served communities.

Whereas, ICANN offers a total of 85 such digital services, some 50 of which are services that have been partially or wholly developed by ICANN staff, or under ICANN staff supervision, leaving a code-base for maintenance under ICANN staff control.

Whereas, the Board Risk Committee has reviewed preliminary findings as presented by the Chief Innovation and Information Officer (CIIO) during ICANN52 in Singapore.

Whereas, the Board Risk Committee has reviewed the CIIO's short- and longer-term treatment of IT security

matters on 17 April 2015 and agrees with the CIIO's recommendations that there is an immediate need to assess the software code-base managed by ICANN staff that has not already been assessed.

Whereas, the individual assessments may not individually reach the threshold of US$500,000 requiring Board approval, however because collectively they may reach that threshold, the Board Risk Committee further referred this matter to the Board Finance Committee.

Whereas, the Board Finance Committee has recommended that the Board delegate to the President and CEO, or his designee(s), the authority to perform all necessary contracting and disbursements to address the immediate need of assessing the software code-base managed by ICANN staff.

Whereas, there are sufficient funds in the FY15 contingency fund to cover the costs of this project.

Resolved (2015.04.26.23), the Board authorizes the President and CEO, or his designee(s), to perform all necessary contracting and disbursements to obtain a comprehensive review and security vulnerability assessment of all software platforms in use at ICANN for delivering digital services, including contracting with external service providers, acquiring needful tools, expenditure disbursement and undertaking remediation measures as appropriate.

Resolved (2015.04.26.24), the Board directs the President and CEO, or his designee(s), to provide regular updates to the Board Risk Committee on the progress of the long-term plan to ensure systems design and systems architecture are integrated into standard ICANN processes, and that security considerations occupy an essential role in corporate decision making.

Rationale for Resolutions

2015.04.26.23-2015.04.26.24

As part of ICANN's digital services health-check, during the first quarter of FY152014, ICANN's IT organization initiated an RFP process to select a suitable external third-party with a reputation and the needful skills to assess all the services and the underlying technologies ICANN has deployed. Following the RFP process, ICANN selected and engaged the services of a globally-recognized leader in undertaking such assignments.

The selected contractor performed a thorough analysis of the ICANN portfolio of digital services. ICANN staff decided to leverage the SANS Institute 20-factor Critical Security Controls framework (see http://www.sans.org/critical-security-controls/controls. The contractor produced a report during the first quarter of FY15 to identify those framework-factors that met or exceeded the "Green" standard, while also identifying those framework-factors that could use further attention.

The report particularly highlighted one factor – Application Software Security – for deeper analysis.

Concurrently, staff inventoried all the digital services it offers the ICANN community. That number stands at 85 today. Staff catalogued the number of software platforms (development environment plus database or content management system), which have been leveraged to develop these services over the last 15+ years. Staff also determined that ICANN delivers digital services leveraging 10+ software platforms for the benefit of its served communities.

Following the SANS Institute framework-based assessment, ICANN IT staff initiated a 16-projects portfolio, focused on improving ICANN's defences in those IT infrastructure areas meriting further attention.

Staff analysed the nature of data captured, manipulated, stored and delivered by these services. The analyses looked at data integrity, data sensitivity and data privacy,

among other factors. The result of this analysis showed a concentration of high-sensitivity data in services that serve ICANN's Contracted Parties community.

Staff retained the services of a deep-specialty firm with expertise in the software package and platform utilized by ICANN to specifically assess digital services deployed for the benefit of the New gTLD program. This specialty firm produced a report in late February of 2015, identifying areas that merited further attention.

Staff has determined that all other (~10) software platforms merit similar assessments. In attempting to estimate the costs of this project, staff approached three large firms with extensive ranges of skill sets and knowledge on numerous software platforms. Staff then also made cost inquiries at smaller, niche or subject matter expert firms that have concentrated expertise on just one or a few software platforms. The estimates received from the larger firms were significantly higher than those from the niche firms, even though both size firms have relatively equal expertise on any given software platform for which the niche firms have concentrated expertise. Accordingly, staff appropriately determined to recommend using numerous, smaller niche firms, rather than one larger firm for this project. This will have the added benefit of allowing multiple assessments to be performed in parallel.

The Board reviewed staff's recommendation for assessing potential software-driven vulnerabilities in the code-base of services leveraging these platforms, and the determination that the proposal met the standard for such assessments. The process for selection of subject matter expert firms for such assessments does not call for public consultation, as the assessment of the code-base is the primary consideration and the expenditure with any given vendor is not expected to reach the level requiring a public bidding process as set out in ICANN's Procurement Guidelines (see https://www.icann.org/en/system/files/files/procurement-

guidelines-21feb10-en.pdf [PDF, 1.03 MB]). However, the collective amount anticipated to be spent in this effort across firms is anticipated to be above the contracting and disbursement limit for which ICANN management alone can approve.

It should be noted that this project is just the first step in a comprehensive approach. ICANN acknowledges that we have experienced some security issues, resulting from various causes in the recent past, and the Board and staff are committed to taking the steps necessary to help ensure such issues, or any other issues, do not arise in the future. To that end, the Board has directed the President and CEO, or his designee(s), to dedicate additional attention and resources to all IT facilitates to ensure that they achieve and/or maintain the level of security that is appropriate and warranted given ICANN's mandate and to report periodically back to the Board on continued progress.

There will be a financial impact on ICANN in engaging in such an assessment but it is already covered in the budget under the contingency fund.

This is an Organizational Administrative function that does not require public comment.

i. ## Investment management – Adjustments to the account structure

Whereas, the Board approved in previous years the new gTLD Investment Policy and the creation of three different investment accounts to hold and manage the funds resulting from new gTLD application fees collected.

Whereas, the new gTLD Investment Policy requires that, when the aggregate amount of remaining new gTLD funds reaches $150 million, those remaining funds be managed by two investment firms instead of three.

Whereas, the new gTLD remaining funds amount to

US$171 million as of 31 March 2015.

Whereas, auction proceeds have been collected for a total (net of auction costs) of approximately US$59 million.

Resolved (2015.04.26.25), the Board authorizes the President and CEO, or his designee(s), to take all actions necessary to consolidate the new gTLD remaining funds with two of the three existing investment managers.

Resolved (2015.04.26.26), the Board authorizes the President and CEO, or his designee(s), to take all actions necessary to invest the proceeds generated through the last resort auctions in the New gTLD Program in a segregated investment management account.

Rationale for Resolutions 2015.04.26.25-2015.04.26.26

By the end of June 2012, and pursuant to the New gTLD Investment Policy (available at https://www.icann.org/resources/pages/investment-policy-new-gtld-2013-01-07-en), the application fees received in the first application round in the New gTLD Program have been invested in investment accounts at three different investment firms. The Board-approved New gTLD Investment Policy includes a provision requiring that once the remaining funds under management reach $150 million, only two investment managers should be used. The current level of remaining new gTLD funds is US$171 million (as of 31 March 2015), and therefore is approaching the US$150 million threshold.

Separately, net auction proceeds gathered through the last resort auctions within the New gTLD Program of US$59 million have been collected over the past eight months and kept in a separate bank account. These funds need to be invested until the mechanism for

disposition of the auction funds is determined.

As a result, the Board Finance Committee has approved a staff recommendation that that: (1) the remaining new gTLD funds are consolidated into two investment managers, as required by the of the New gTLD Investment Policy; and (2) the third investment manager (i.e., the investment manager that will no longer have New gTLD application funds under management) will be requested to create a new investment account, dedicated to managing auction proceeds received through the New gTLD Program.

This decision is in line with prior Board actions on the management of application fees collected within the New gTLD Program. This decision has no impact on the security, stability and resiliency of the Internet DNS.

This is an Organizational Administrative Function that does not require public comment at this stage.

j. AOB

No resolutions taken.

Published on 28 April 2015

crowell moring

Annex 2.

KU LEUVEN

Faculty of Arts
Department of Linguistics
Contact Information Redacted

Flip Petillion
Partner
Contact Information Redacted

By mail land email: Contact Information Redacted

10 March 2014

Dear Sir,

I refer to your request to (1) review and comment on ICANN's process regarding strings that may be confusingly similar to a degree that it prevents them from being delegated; and (2) review ICANN's decision to put .hotels and .hoteis in a contention set in light of:

- ICANN's framework for assessing string similarity;
- the String Similarity Panel's decision that .hotels and .hoteis were confusingly similar; and
- the other gTLD applications and the existing DNS.

String comparison is a subject that I regularly examine on a professional basis. As explained below, I have complemented, as a linguist, my analysis and conclusions with a psycholinguistic analysis, provided by Dr. Emmanuel Keuleers.

In Attachments 1 and 2, I have added detailed information on our respective professional backgrounds and qualifications.

Below, I will describe:

1. the context of your request;
2. the multi-disciplinary methodology I have adopted;
3. our linguistic and psycholinguistic findings ; and
4. my conclusions.

Executive summary

ICANN's methodology for assessing string similarity does not permit any valid conclusions on string similarity. Various aspects of the methodology demonstrate that *construct validity*, *i.e.*, the need for a methodology to measure what it actually claims to measure, has been largely overlooked. There is no reference to any attempt to link the methodology to empirical evidence obtained by *experimental testing of string confusion in human readers*. Moreover, given that the ICANN approach is based on human and algorithmic assessment of string similarity, it is surprising that the reliability of the process is undocumented. For a method to be reliable, it is not only necessary to have different raters. There must also be a high agreement between those raters and there must be a high correlation between the raters and the algorithmic assessment. ICANN's methodology does not seem to have involved multiple raters, let alone in agreement with each other and having a high correlation with the algorithmic assessment. Finally, ICANN did not take into account linguistic evidence taken from real world language data, although such data is crucial to any valid evaluation of whether or not a particular pattern is sufficiently frequent and productive to avoid string confusion.

In contrast, to examine the string similarity between *hotels* and *hoteis*, we have applied a dual approach combining psycholinguistic experiments with corpus-linguistic analyses. The essence of this approach is that, through experiment and a psycholinguistic literature review, we first analyzed to what extent the visual similarity between the characters *l* and *i* may cause problems of word identification in words differing only in these characters (as compared to words differing only in other characters). Once we had analyzed what happens in the mind of the language user, we then analyzed real life language data to assess to what extent the language user is frequently exposed to words differing only in the characters *l* and *i*. Where this is the case, we have explained why visual word recognition does not cause any difficulties in this case.

This approach led us to make the following findings:

- In normal human reading, character misidentification is very uncommon. In order to induce errors in identification, experiments studying character identification must degrade the perceptibility of characters.
- A psycholinguistic study based on existing behavioral data on word recognition shows that behavioral evidence does not support the claim that readers are less sensitive to the difference between *l* and *i* than to the differences between other characters. The evidence does not support the view that visual similarity between the *l* and *i* characters is a cause of word confusion in ordinary circumstances.
- A second psycholinguistic study based on newly collected behavioral data from more than 1,600 participants in a worldwide online vocabulary study shows that the character

string *hoteis* is not confused with the character string *hotels* or *vice versa* in ordinary visual word recognition.

- A corpus-linguistic study involving a quantitative analysis of large English data sets shows that the English language includes a substantial number of words that differ from each other only in the alternation of *l* and *i*.
- A second corpus-linguistic study shows that the alternation of *l* and *i* is not only frequently occuring; it also occurs in common words. This explains why the alternation of *l* and *i* as observed in *hotels/hoteis* does not confuse the language user: he is frequently and repeatedly exposed to this pattern.
- Finally, a quantitative analysis of a large English-Portuguese parallel corpus shows that interlingual English-Portuguese orthographic neighbors presenting the alternation of *l* and *i* in plural noun endings is a reasonably frequent phenomenon.

This approach showed that the word pair *hotels/hoteis* does not cause user confusion. The patterns underlying the word pair *hotels/hoteis* explain why string confusion should not be expected and why behavioral data shows that the average, reasonable Internet user does not in fact become confused between the *hotels* and *hoteis* strings.

I. Context

The company Booking.com B.V. (Booking.com) has applied to ICANN to operate the generic top-level domain (gTLD) .hotels in the Internet root zone. The company Despegar Online SRL has applied to ICANN to operate the gTLD .hoteis in the Internet root zone.

ICANN has informed Booking.com that it is of the opinion that the gTLD strings .hotels and .hoteis are confusingly similar. ICANN informed Booking.com that:

> *"After careful consideration and extensive review performed against criteria in Section 2.2.1.1. of the Applicant Guidebook, the String Similarity Panel has found the applied-for string (.hotels) is visually similarly to another applied-for string (.hoteis), creating a probability of user confusion."*

The two strings were put into a contention set by ICANN. This prevents both strings from being delegated.

On 22 March 2013, I advised that the language user is able to visually distinguish the words *hotels* and *hoteis*. My advice was based on a study of an English corpus of very frequent words.

On 7 June 2013 and 9 January 2014, ICANN published additional information on the process that was used by the String Similarity Panel in assessing the confusing similarity of applied-for gTLD strings.

I have analyzed ICANN's methodology and decision in view of this new information.

II. Relevant Principles and Methodology

An analysis of the methodology that ICANN used to evaluate the similarity between letter strings allows me to conclude that ICANN's approach does not take the necessary account of behavioral and linguistic evidence (A.). Below, I explain why there is a need for a psycholinguistic and linguistic approach. I also explain the linguistic and psycholinguistic hypotheses underlying my analysis (B.). I clarify why this approach requires the expertise of a multidisciplinary team. Finally, I specify how this requirement has been taken into account in this report (C.)

A. ICANN's standard and methodology for assessing confusing string similarity

I have examined and analyzed the following material from ICANN:

- ICANN's Applicant Guidebook, containing a description of the review methodology and ICANN's standard for assessing string confusion by the so-called String Similarity Panel;
- the qualifications of the String Similarity Panel and the expected review methodology as set forth in ICANN's expression of interest document available at http://archive.icann.org/en/topics/new-gtlds/eoi-string-sim-31jul09-en.pdf;
- the process description of the String Similarity Panel as posted by ICANN on 7 June 2013 as well as a letter of 18 December 2013 from the Panel's Manager to ICANN. This letter was said to provide "*a summary of the process, quality control mechanisms and some considerations surrounding non-exact contention sets for string similarity evaluation as requested by ICANN*" and;
- the list of applied-for gTLD strings as made available on https://gtldresult.icann.org/application-result/applicationstatus/viewstatus.

These documents show that ICANN's standard for assessing confusing string similarity is as follows:

> "***Standard for String Confusion*** – *String confusion exists where a string so nearly resembles another visually that it is likely to deceive or cause confusion. For the likelihood of confusion to exist, it must be probable, not merely possible that confusion will arise in the mind of the average, reasonable Internet user. Mere association, in the sense that the string brings another string to mind, is insufficient to find a likelihood of confusion.*"

ICANN's methodology was to use a proprietary algorithm (called SWORD) to evaluate the similarity between letter strings and supplemented the outcome of this algorithm with the judgment of a panel, the String Similarity Panel. While the details of this algorithm and process have not been made public, its results suggest that the algorithm uses a similarity metric based exclusively on visual character overlap. The final determination of similarity was left entirely up to the judgement of ICANN's String Similarity Panel. The background of the Panel's members and the methodology that was used are unclear. Also unclear is the basis on which the likelihood of confusion for the average, reasonable Internet user was analysed.

The aim of the outlined methodology is clearly to predict string confusion by human readers. Several aspects of the methodology show that *construct validity*, *i.e.*, the need for a methodology to measure what it actually claims to measure, was largely overlooked:

- First, there is no reference to any attempt to link the methodology to empirical evidence obtained by *experimental testing of string confusion in human readers*.
- Second, most of the evidence on human string confusion relates to experimental psychology literature. Given that the identity of the panel members is unknown, there is no clear basis on which to question the expertise of the panel members. However, it is remarkable that the panel's composition does not include behavioral scientists given its task of evaluating the behavior of the average, reasonable Internet user.
- Third, since the approach taken by ICANN uses human and algorithmic assessment of string similarity, it is surprising that the reliability of the process is undocumented. For the method to be reliable, it is not only necessary to have different raters, but also to have a high level of agreement between these raters and a high correlation between raters and the algorithmic assessment. ICANN's methodology does not appear to have involved different raters, let alone agreement between those raters or a high correlation between them and the algorithmic assessment.
- Finally, ICANN's approach does not take account of linguistic evidence taken from real world language data, despite the fact that such data is crucial to evaluate whether or not a certain pattern is sufficiently frequent and productive not to lead to string confusion.

This leads me to conclude that **no conclusion is possible based solely on the ICANN methodology** and that another approach is needed. In the next chapter we clarify what that approach should be.

B. Need for a psycholinguistic and linguistic approach

Analyzing the likelihood of confusion of the average, reasonable Internet user when confronted with similar strings can only be done on the basis of real word data. This data can be obtained and analyzed through different methods. The data and analysis of one specific method will serve as a rater. If the data and analysis from various relevant methods are in agreement with each other and show a high degree of correlation, we can reach a valid conclusion about the likelihood of confusion of the average reasonable Internet user.

1. Our **approach** is a **dual one**. First, we make use of a **psycholinguistic approach** to determine whether or not, and to what extent, humans can distinguish character strings which differ only in the *l* and *i* characters. To that end, we use large behavioral datasets of visual word recognition performance. On the basis of this analysis, we gather behavioral evidence to determine the extent to which accurate recognition occurs under normal viewing conditions.

If this evidence shows that the similarity between character strings which differ only in the *l* and *i* characters does not hamper recognition, we go on to analyze whether a **corpus-linguistic analysis** explains this behavior. For this analysis, we use a linguistic corpus, *i.e.* a large and structured set of texts (nowadays usually electronically stored and processed), whose purpose is

to give a complete picture of a language. These structured set of texts are used to perform statistical analysis and hypothesis testing, checking occurrences or validating linguistic rules within a specific language territory (an intra-lingual corpus) or between different languages (an inter-lingual corpus). The analysis of a linguistic corpus will allow us to evaluate the frequency of the occurrence of the pattern underlying the word pair *hotels* and *hoteis* and to determine the extent to which the similar character strings belong to the regular lexicon of the language under study.

2. The **background** to this dual approach stems from **fundamental insights about the way reading proceeds and about the way language functions**.

Reading. Skilled human readers routinely identify all words in a text. Even the best computer-based recognition systems never achieve human accuracy. The human visual system is very different from computer vision systems. Also, the visual word recognition system is different from typical visual object perception. Analysis of brain activity during reading shows that the processing of written text is a highly specialized task that strongly activates the middle portion of the left *fusiform gyrus*, a specific region of the brain which is not strongly involved in the processing of other forms of visual information (Cohen et al., 2000, McCandliss, Cohen & Dehaene, 2003).

The process of becoming a skilled reader has much in common with learning to recognize faces. When we first see the faces of identical twins we may be confused, but through repeated exposure we learn to reliably distinguish between them. Equally, we are able to recognize that a familiar face belongs to the same person with or without glasses, with or without a beard or a mustache, with or without make-up, and so on.

Discriminating between characters is the basis of the reading process and involves a similar learning process. We learn that small visual differences, for instance between *l* and *i*, are always important, while, at the same time, we learn to discard large but non-meaningful differences, such as between uppercase *A* and lowercase *a*. After a few years of exposure to characters, nearly all humans become experts at this task. The exposure to characters continues and expertise grows throughout our lifetime. This is why Finkbeiner and Coltheart (2009) write that "*letters are highly overlearned visual patterns*".

Mueller and Weidemann (2012) give an overview of the research on human character recognition since the problem was first experimentally studied by Catell (1886). The study by Geyer (1977), which presented characters in lowercase Tactype Futura demi 5424, a very simple non-ornamental font, is interesting in our case because domain names are usually displayed as lowercase characters in a simple sans-serif type.

Geyer (1977) begins by noting that, in ordinary circumstances, human character identification is nearly flawless. He writes:

> "*One problem in the development of alphabetic confusion matrix data is limiting correct performance.* **Under normal viewing conditions, correct recognition is highly probable** *and a resultant confusion matrix is uninteresting.*"

To induce errors in character recognition, Geyer found, like all researchers in the field, that he had to create adverse circumstances for his participants. He adjusted the brightness and duration of the presentation of the characters such that participants were only able to achieve correct identification in half of the trials. To achieve this characters were presented very dimly and very quickly. Geyer calculated that, in these conditions, there was a 50% chance of correctly identifying *l* and a 34% chance of correctly identifying *i*. In addition, there was a 12% chance that an *l* would be identified as an *i* and an 18% chance that an *i* would be identified as an *l*. We must place these results in perspective: there were six characters with a lower chance of correct identification than *i* (*s*, *z*, *e*, *c*, *t*, and *f*) and four additional characters with a lower to equal chance of correct identification than *l* (*r*, *q*, *o*, and *u*). Moreover, there were 14 pairs more likely to confusion than *l/i* (*a/o*, *f/j*, *f/l*, *e/o*, *b/h*, *o/a*, *p/n*, *q/a*, *s/n*, *t/i*, *t/l*, *v/w*, *y/v*, and *z/x*).

Similar results can be found throughout literature. However, it is important to understand that these results present character identification at the absolute limits of visual perception. Even in these circumstances, many character pairs are more often confused than *l/i*.

So, while character identification can be affected under very adverse circumstances, such results do not tell us much about the confusion of characters within words in ordinary circumstances. In particular, we would like to know whether, and to what extent, humans can distinguish character strings which differ only in the *l* and *i* character. We will study this problem using large behavioral datasets of visual word recognition performance.

Language. Language is *structured* and *productive*. By *structured* we mean that there are basic elements which are combined to produce more complex elements. By *productive* we mean that a limited number of complex elements can be combined according to rules for the production of almost limitless coherent meaningful utterances.

Every language consists of a fixed set of phonemes (sounds) and graphemes (letters) that can be combined without limitation. This linguistic reality poses no problems to the language user, who is used to being confronted with words that differ from each other in only a single character. This does not prevent the language user from visually distinguishing these words so as to see them as different meaningful entities. Therefore, **string similarity is an inherent feature of all natural languages**.

In order to observe to what extent language users are able to distinguish between two character strings, one can **analyze to what extent the similar character strings belong to the regular lexicon of the language that is examined**. To do so, the most effective methodology is corpus linguistics. Corpus linguistics is a method of linguistic analysis that uses a collection of natural or "real word" texts known as corpus. Corpus linguistics offers a unique insight into the dynamic of language that has made it one of the most widely used linguistic methodologies (Baayen 2008, Johnson 2011, Linquist 2009, McEnery & Hardie 2011).

3. We will now clarify the **psycholinguistic and linguistic hypotheses** on which this approach is based.

For the **psycholinguistic analysis**, the task we will investigate is human lexical decision. In this task, participants have to press a button indicating whether a character string is known to them (WORD response) or whether it is unknown to them (NONWORD response). After the participant's answer is registered, another string is presented. Each string is presented only once to each participant.

The basic hypothesis behind this psycholinguistic approach is that if the alternation between *l* and *i* is NOT confusing, then the non-Portuguese language user, having an understanding of English, Dutch, French, German, etc., should be able to identify hotels as a WORD and hoteis as a NONWORD.

If the Internet user is unable to make the distinction between *l* and *i* in a character string only differing through *l/i* alternation, then the number of WORD responses to a nonword (*i.e.* the nonword stimulus) will approach the WORD responses to a word (*i.e.* the word stimulus). This would imply that the nonword stimulus is mistaken for the word stimulus. For instance, if the character string *hoteis* is mistaken for *hotels*, then the proportion of WORD responses to *hoteis* should approach the proportion of WORD responses to *hotels*, assuming that most English speaking participants will not know the Portuguese word *hotéis* (or at least would not classify it as a word in the test context). (

In other words, if readers of this text are able to distinguish the character strings *hotels* and *hoteis* in the paragraph above (and in the current paragraph), then the hypothesis of confusion should be rejected.

An improved version of the hypothesis can be made in probabilistic terms: If nonwords differing from words only by substitution of the character *i* for the character *l* evoke significantly more WORD responses than nonwords differing from words by substitution of *l* for any other character, then it is likely that the substitution of *l* by *i* evokes errors in word identification.

We can frame this hypothesis even more carefully in terms of the signal detection paradigm. In this paradigm, *sensitivity* indicates how readily a particular difference can be detected. This sensitivity can be measured parametrically by using the traditional d' measure, or non-parametrically by using the A index (Zhang & Mueller, 2005). The A index makes less assumptions about the structure of the underlying data and is more fit for our current report. A varies between 0 and 1, with a higher value indicating better sensitivity. In other words, if a character string is routinely mistaken for another character string, then sensitivity will approach 0. If a difference is always detected, sensitivity will approach 1.

It is important to note that, regardless of visual similarity, there may be a number of reasons why participants in the lexical decision task respond "WORD" to nonwords. For instance, participants may mistakenly press the wrong button, a nonword may actually be a word in their dialect or the nonword may be confused for another word by sound. Therefore, it is important to compare the sensitivity to a particular difference with a realistic baseline sensitivity level. Hence, the sensitivity to the *l/i* difference, such as in *hotels/hoteis*, must be compared with the sensitivity to the *l/~i* difference, where *~i* means any letter that is *not i*.

For the subsequent **corpuslinguistic analysis**, the basic idea behind this approach is that if a certain alternation is sufficiently frequent, and thus productive, in a language, there is no reason to accept that this alternation would be perceived by the language user as too similar to allow him to distinguish two strings differing only by that alternation.

As a matter of fact, the basic hypothesis can be defined as follows: **If the analysis of large computerized corpora shows us that a certain alternation occurs reasonably frequently in the language under study, it should be clear that strings differing only in this alternation are sufficiently distinct from each other to be recognized as different words by the language user.**

C. Multidisciplinary team

The approach we describe under B. is multi-disciplinary and requires psycholinguistic as well as linguistic expertise. This explains our decision to call upon a psycholinguist to perform the first part of the study. **Dr. Emmanuel Keuleers** kindly agreed to put his excellent psycholinguistic expertise at our disposal and to do the experiments needed.

As a linguist with strong expertise in corpus-linguistics and quantitative language research, I performed the second part of the study and I was also responsible for the integration of the two analyses in this report.

III. The analysis

In this section, we apply the twofold approach outlined under II. to the word pair *hotels* and *hoteis*. The first step is to collect behavioral data on the basis of psycholinguistic experiments (A.) to see to what extent the visual similarity between the characters *l* and *i* causes problems of word identification in words differing only in these characters. We then pass to the corpuslinguistic analysis (B.) of real life language data to see to what extent the language user is frequently exposed to words differing only in these characters.

A. Psycholinguistic analysis

Two studies

In our first study, we analyze existing behavioral data on word recognition coming from the British Lexicon Project to study. For a detailed presentation of this materials, see attachment 3. This analysis is focused on all words having the same length but differing in just one letter.

In the second study we collected new behavioral data focusing more specifically on the word pair *hotels/hoteis*.

These two studies, performed by Dr. Emmanuel Keuleers, are detailed in attachment 5.

Our findings

The first study shows us that in normal human reading, character misidentification is very uncommon. This is apparent from both the academic literature and experiments studying character identification. Experiments show that the perceptibility of characters must be degraded to induce errors in identification.

Research using behavioral data does not support the hypothesis that the visual similarity between the characters *l* and *i* causes increased misidentification.

This was confirmed in the second study by Dr. Emmanuel Keuleers focusing on the visual similarity between the character strings *hotels* and *hoteis*.

The results of both studies suggest very strongly that character differences play a negligible role in ordinary word recognition. In contrast, linguistic differences inform decisions in a predictable manner.

B. Corpuslinguistic analysis

Three studies

First, we examined to what extent the alternation *l/i* is frequent in English and generates word pairs differing only by this alternation. To do so, we performed a **quantitative analysis of two general English word lists** to analyze to what extent the alternation *l/i* occurs in English since word lists give general information about the words belonging to the lexicon of a language (our first study).

If the alternation under study is not restricted to a limited set of word pairs, one can even go further and study what the frequency of the words affected by the alternation is. The more frequent these words are, the less marginal the alternation within that language is and the more language users will be confronted with/used to the alternation (our second study). For this second study, we analyzed a **frequency list based on a large monolingual corpus**, giving us information on the usage of word pairs in real language usage situations.

Since *hotéis* is a Portuguese word, one could also think of the word pair *hotels/hoteis* as an inter-lingual orthographic neighbor or cognate. By inter-lingual cognates, we mean words of two different languages that have identical or similar spellings. In many cases, they share the same origin (etymon). Since the word pair *hotels/hoteis* belongs to English and Portuguese, it is relevant to analyze to what extent bilingual word pairs of this type occur regularly between these two languages (our third study). To perform this analysis, we **examined a parallel corpus of Portuguese and English texts.**This should allow us to find out how many inter-lingual cognates exist between Portuguese plural noun ending in *–eis/-ais/-ois/-uis* (used with or without accents) and English words ending in *–els/-als/-ols/-uls*.

These three studies, performed by myself, are detailed in attachment 6.

Our findings

Our first study showed that English has quite a substantial number of word pairs differing only in the alternation *l/i*. Our second study showed that alternation *l/i* itself is not only a frequent pattern, but also that it occurs in frequent words. Finally, our third study showed that that English and Portuguese present a substantial number of inter-lingual cognates, i.e. completely analogous word pairs with exactly the same number of letters, differing only through the alternation of *l* and *i*. Most of these word pairs are quite frequent and concern the *ais/ail* alternation. 8 of them offer the *eis/els* alternation. As a result, language users who speak both English and Portuguese are used to the alternation of *l* and *i* between English and Portuguese.
The outcomes of our corpuslinguistic studies are thus consistent with and confirm the empirical behavioral data showing that language users who understand English and/or Portuguese are not confused by the alternation of *l* and *i*.

IV. Conclusion

ICANN did not apply a valid scientific approach in addressing the issue of string similarity. Its decision to put *.hotels* and *.hoteis* in a contention set is based on an undocumented approach that did not take account of empirical behavioral data and linguistic evidence that is essential in assessing string similarity and the likelihood of confusion of the average, reasonable Internet user.

Our multidisciplinary approach, that is based on psycholinguistic experiments and corpuslinguistic analyses shows that there is no reason to believe that the word pair *hotels* and *hoteis* will be confusingly similar to the average, reasonable Internet user.
Empirical behavioral data on character confusion shows that several alternations (including the alternations *t/i*, *l/f*, *l/t*, *e/o* and *e/a*) are more likely to cause confusion than the alternation *l/i*. However, while ICANN put *.hotels* and *.hoteis* in a contention set, it did not consider word pairs that were differing from each other only by these more similar alternations to be confusingly similar (*parts/paris*;*maif/mail*; *srt/srl*; *vote/voto*; and *date/data*). The character alternations in these word pairs are at least as confusing as those in the *hotels/hoteis* word pair.

As a result, ICANN's conclusion that it is probable that confusion will arise in the mind of the average, reasonable Internet user when the *.hotels* and *.hoteis* strings are both delegated is both arbitrary, in that character pairs that are at least equally confusing as *l/i* are not considered confusingly similar, and contradicted by our analysis.

Yours sincerely,

Prof. Dr. Piet Desmet

ATTACHMENTS:

1 **Curriculum Vitae Prof. Dr. Piet Desmet**

2. **Curriculum Vitae Dr. Emmanuel Keuleers**

3. **List of Materials Reviewed**

4. **Relevant publications or sources referred to or cited in support**

5. **Psycholinguistic analysis performed by Dr. Emmanuel Keuleurs**

6. **Corpus-linguistic analysis performed by Prof. Dr. Piet Desmet**

ATTACHMENT 1 - CURRICULUM VITAE PROF. DR. PIET DESMET

Full Professor of French and Applied Linguistics
(KU Leuven & KU Leuven Kulak - Belgium)

Coordinator of the research team ITEC (*Interactive Technologies*)

Former Dean of the Faculty of Arts, KU Leuven Kulak

Contact information

KU Leuven Kulak
Faculty of Arts
Contact Information Redacted

Personal information

Contact Information Redacted

Degrees

MA in Romance Philology, K.U.Leuven, 1987, maxima cum laude
BA in Philosophy, K.U.Leuven, 1987, magna cum laude
Teacher training French as a Foreign Language, K.U.Leuven, 1987, maxima cum laude
Ph.D. in Romance Philology (French linguistics), 1994, maxima cum laude with congratulations of the jury

Career

Research Assistant in French Linguistics, Department of Linguistics, KU Leuven, 1987-1994
Post-doctoral Research Fellow, Department of Linguistics, KU Leuven, 1994-1996
Associate Professor ('Docent') of French and Applied Linguistics, KU Leuven, 1996-1999
Associate Professor ('Hoofddocent') of French and Applied Linguistics, KU Leuven, 2000-2002
Professor ('Hoogleraar) of French and Applied Linguistics, KU Leuven, 2003-2006
Full Professor ('Gewoon Hoogleraar') of French and Applied Linguistics, KU Leuven, 2006-

Research

- Areas : French & General Linguistics
  Applied Linguistics
  Corpus Linguistics
  Computer-Assisted Language Learning
  Foreign Language Learning & Second Language Acquisition

- Coordinator of the research team ITEC *(Interactive technologies)*

- Director of numerous research projects
  e.g. EU: Ling@tic (Interreg III), Cobalt (Interreg IV), GOBL (LLP), European Survey on Language competences (ESLC), etc.
  iMinds (Innovation in ICT): MAPLE (mobile learning), LLINGO (serious gaming), G@S (minigames), iRead+ (input enhancement), EDUTAB (tablets in education)
  Industrial research Fund: Adaptive learning environments
  Institute for Science & Technology: Baekeland (corrective feedback), e-testing, etc.
  KU Leuven: Idiomatic, Language Portal, Jurimatic, Lingu@flex, etc.
  Flemish government: SELOR (e-testing), OBPWO

- Director of eleven Ph.D. theses (in Applied Linguistics – 6 completed and 5 running) and of numerous MA-theses in CALL and Foreign Language Learning

- Author of about a hundred publications (books & articles in international peer-reviewed journals). For an overview, see https://lirias.kuleuven.be/cv?u=U0019561

- Presenter at numerous international conferences (CALICO, WORLDCALL, EUROCALL, UNTELE, EDMEDIA, etc.) and organizer of different international conferences

- Member of the Editorial Board of *ITL International Journal of Applied Linguistics*, of *Revue française de linguistique appliquée* and of *Le français dans le monde. Recherche & applications*.

- Two spin-off companies:
  BLCC (Blended language learning for professionals)
  Telraam -> Indie Education -> Televic education (educational technology)

- National Representative for Belgium within EUROCALL

IV. TEACHING

- Different BA and MA Courses in French and Applied Linguistics, notably an MA course on CALL

- Coordinator of the multimedia language lab at KU Leuven Kulak

V. MANDATES

- Member of the Board of KU Leuven Research & Development

- Vice-President of the *Alliance française West-Vlaanderen*

- Member of the Board of *WTV Zuid West-Vlaanderen* (regional broadcasting company)

Further information

Home page Dutch: http://wwwling.arts.kuleuven.ac.be/franling/pdesmet/
Home page French: http://wwwling.arts.kuleuven.ac.be/franling_f/pdesmet

10 most relevant publications (last 5 years)

1. Montero Perez, M., Van Den Noortgate, W., Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System*, 41 (3), 720-739.
2. Vandewaetere, M., Cornillie, F., Clarebout, G., Desmet, P. (2013). Adaptivity in Educational Games: Including Player and Gameplay Characteristics. *International Journal of Higher Education*, 2 (2), 106-114.
3. Wauters, K., Desmet, P., Van Den Noortgate, W. (2012). Item Difficulty Estimation: an Auspicious Collaboration Between Data and Judgment. *Computers and Education, 58*, 1183-1193.
4. Vanderbauwhede, G., Desmet, P., Lauwers, P. (2011). The shifting of the demonstrative determiner in French and Dutch in parallel corpora: from translation mechanisms to structural differences. *Meta: Journal des Traducteurs, 56 (2)*, 443-464.
5. Vandewaetere, M., Desmet, P., Clarebout, G. (2011). The value of learner characteristics in the development of computer-based adaptive learning environments. *Computers in Human Behavior, 27*, 118-130.
6. O'Regan, B., Rivens Mompean, A., Desmet, P. (2010). From spell, grammar and style checkers to writing aids for English and French as a foreign language : challenges and opportunities. *Revue Française de Linguistique Appliquée, 15* (2), 67-84.
7. Wauters, K., Desmet, P., Van Den Noortgate, W. (2010). Adaptive Item-Based Learning Environments Based on the Item Response Theory: Possibilities and Challenges. *Journal of Computer Assisted Learning, 26* (6), 549-562.
8. Soylu, A., De Causmaecker, P., Desmet, P. (2009). Context and Adaptivity in Pervasive Computing Environments: Links with Software Engineering and Ontological Engineering. *Journal of Software, 4* (9), 992-1013.
9. Vandewaetere, M., Desmet, P. (2009). Introducing psychometrical validation of questionnaires in CALL research: The case of measuring attitude towards CALL. *Computer assisted language learning: an international journal, 22* (4), 349-380.
10. Desmet, P. (2007). L'apport des TIC à la mise en place d'un dispositif d'apprentissage des langues centré sur l'apprenant. *ITL: review of applied linguistics, 154*, 91-110.

Supervision of PhD's

Completed

1. Montero Perez, Maribel, Desmet P. (sup) , Peters, E. (cosup.) (2013). *Watch and Learn?! Five studies into the use and effectiveness of captioned video for L2 listening comprehension and vocabulary acquisition,* 300 pp.
2. Soylu, Ahmet, De Causmaecker, P. (sup.), Desmet, P. (sup.), Duval, E. (sup.) (2012). *Exploiting metadata, ontologies and semantics to design/enhance new end-user experiences for adaptive pervasive computing environments.*, 222pp.

3. Wauters, K., Van Den Noortgate, W. (sup.), Desmet, P. (cosup.) (2012). *Adaptive item sequencing in item-based learning environments*, 249 pp.
4. Vandewaetere, M., Clarebout, G. (sup.), Desmet, P. (cosup.) (2011). *Learner control for adaptive learning: The importance of learners' perceptions*, 268 pp.
5. Vanderbauwhede, G., Desmet, P. (sup.), Lauwers, P. (cosup.) (2011). *Le déterminant démonstratif en français et en néerlandais à travers les corpus: théorie, description, acquisition*, 470 pp.
6. Verleyen, S., Desmet, P. (sup.), Swiggers, P. (cosup.) (2005). *Fonction, forme et variation: analyse métathéorique de trois modèles du changement phonique au XXe siècle (1929-1982).*

Running

1. Cavdar, Ilgün. *Plus-value didactique d'un environnement d'apprentissage électronique axé sur la compétence linguistique*
2. Cornillie, Frederik. *Effectiveness of corrective feedback for the development of second language grammatical competence in game- and task-based tutorial CALL: perceptions, usage and motivation*
3. Lagatie, Ruben. *Automatic generation of personalized feedback in electronic learning environments.*
4. Stockman, Caroline. *Technology acceptance and a teacher's interactional self-narrative.*
5. Vankeirsbilck, Pascale. *Repenser le subjonctif dans l'enseignement/apprentissage du FLE. Analyse théorique et expérimentale.*

Research funding granted in the last 5 years

iRead+: the Intelligent reading companion.
IBBT-ICON
Duration: 01.01.2012-31.12.2013
Budget: 445.000 euro
Promotor

Games at School (G@s)
IBBT-ICON
Duration: 01.01.2012-31.1.22013
Budget: 271.000 euro
Promotor

Games online for basic language learning (GOBL)
EU-LLP
Duration: 01.01.2012 – 31.12. 2014
Budget:125.000 euro
Promotor

Harnessing collective intelligence in e-learning environments
Industrial Research Fund K.U.Leuven (IOF) – Knowledge platform

Duration: 01.10.2007 - 30.09.2012
Budget: 600.000 euro
Promotor and general coordinator

Clarin
Flemish Ministry – EWI
Duration: 01.03.2010-30.09.29012
Budget: 45.000 euro
Co-promotor
Cobalt: COmmunicating and building Bridges thanks to the Acquisition of Languages through Technologies.
European project (Interreg IV A): EFRA-Funding and and cofunding by Min. of Flemish Community, Prov. West-Vlaanderen, VDAB, Forem, Indie Education, BLCC and Conseil Régional Nord-Pas-de-Calais
Duration : 01.03.2009 - 29.02.2012
Budget: 1.533.000 euro
Promotor and general coordinator

Language Learning in an Interactive Game Environment (LLINGO)
IBBT-ICON
Duration: 01.10.2009 - 30.009.2011
Budget: 450.000 euro
Promotor

Mobile, adaptive & personalized learning experience (MAPLE)
IBBT-ICON
Duration: 01.10.2009 – 30.09.2011
Budget: 528.000 euro
Promotor
Savoirs numériques 5962
Conseil regional – contract research
Duration: 01.06.2009 -31.12.2012
Budget: 100.000 euro
Promotor

European Survey on Language Competences (ESLC)
Fleminsh Ministry of Education
Duration: 24.03.2009-31.12.2012
Budget: 950.000 euro
Co-promotor

DPC. Dutch Paralell Corpus. A Multifunctional & Multilingual Corpus (Dutch-English, Dutch-French)
STEVIN - Nederlandse Taalunie
Duration: 01.05.2006 – 30.09.2009
Budget: 498.109 euro
Promotor

SELFIN- actualisering van de taalexamendatabank ATLAS
SELOR – contractonderzoek
Duration: 15.06.2007 – 14.06.2008
Budget: 175.000 euro
Promotor

SoE-project: Digitale (vak)didactiek voor toekomstige leerkrachten via USolv-IT en Franel. (= e-learning French and Dutch)
Associatie K.ULeuven, School of Education
Duration: 15.12.2007 - 14.12.2008
Budget: 70.000 euro
Promotor

The Language Portal. "Taalportaal. Een elektronische oefen- en testomgeving voor de bachelor taal- en letterkunde".
OI K.U.Leuven
Duration: 01.02.2009 - 31.01.2011
Budget: 120.000 euro
Copromotor
JURI-Flex - De ontwikkeling van complexe juridische en heuristische oefentypes.
OI K.U.Leuven
Duration: 01.10.2007 - 30.09.2010
Budget: 120.000 euro

Research on the levels and knowledge of French in primary school education
Flemish Ministry of Education
Duration: 01.10.2006 - 31.12.2011
Budget: 500.000 euro
Co-promotor

Experience in research with and management of large-scale research infrastructure, possibly as part of a consortium or in an international context

Piet Desmet coordinates or has coordinated (during the last five years) the following large-scale projects in a complex, often international context: iRead+, G@S, LLINGO, MAPLE, Cobalt, DPC, SELFIN,

Contribution to policy-relevant research

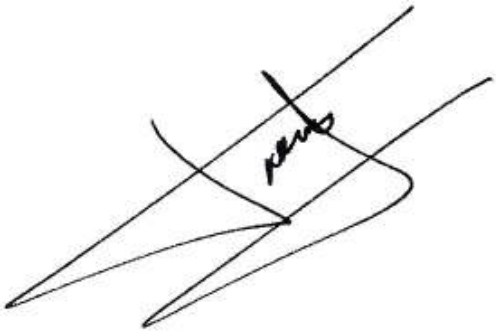
Different projects on language testing for the Flemish Ministry of Education and for the Belgian Ministry for Civil Affairs.

Awards

1988 - Laureate of the Travel grants competition (Reisbeurzenwedstrijd) of the Belgian Department of Education, Administration of Higher Education and Scientific Research.

2006 - Access to Language Education Award of CALICO

2008- Chevalier dans l'Ordre des Palmes académiques – French Ministry of Education



ATTACHMENT 2 - CURRICULUM VITAE DR. EMMANUEL KEULEERS

Dr. EMMANUEL KEULEERS
Research Fellow
Center for Reading Research
Department of Experimental Psychology
Ghent University, Belgium

Contact information
Department of Experimental Psychology
Ghent University
Contact Information Redacted

Web http://crr.ugent.be/members/emmanuel-keuleers

Personal information

- ° 1974
- Belgian

Degrees

- PhD in Linguistics, University of Antwerp, 2008
- Master in Theoretical and Experimental Psychology, Ghent University, 2000

Career

- October 2008-Present: Research Fellow — Department of Experimental Psychology, Ghent University.
- January 2008-July 2008: Researcher — CNTS Language Technology Group, University of Antwerp.
- January 2004–December 2007: PhD Researcher — Center for Psycholinguistics, University of Antwerp.
- November 2001–December 2003: Research Assistant — Center for Psycholinguistics, University of Antwerp.

Research Areas

- Psycholinguistics
  - Visual Word Recognition
  - Computational Models
  - Morphology
  - Crowdsourcing

    - Megatudies
- Computational Linguistics
    - Word frequencies
    - Semantic Models
    - Corpus development
    - Memory-based learning

Membership of Professional Organizations

- Fellow of the Psychonomic Society
- Member of the Association for Psychological Science
- Member of the Belgian Association for Psychological Sciences

Teaching

- Different BA and MA courses in Psychology and Psycholinguistics.
- Regularly invited Lecturer at Research schools and workshops

Further information

Home page: http://crr.ugent.be/members/emmanuel-keuleers

10 most relevant publications (last 5 years)

1 van Heuven, W.J.B., Mandera, P., Keuleers, E., & Brysbaert, M. (2013). SUBTLEX-UK: A new and improved word frequency database for British English. The Quarterly Journal of Experimental Psychology.

2 Frost, R., & Keuleers, E. (2013). What Can We Learn From Monkeys About Orthographic Processing in Humans? A Reply to Ziegler et al. Psychological Science, 24(9), 1868–1869.

3 Kuperman, V., Drieghe, D., Keuleers, E., & Brysbaert, M. (2013). How strongly do word reading times and lexical decision times correlate? Combining data from eye movement corpora and megastudies. The Quarterly Journal of Experimental Psychology, 66(3), 563–580.

4 Keuleers, E., & Brysbaert, M. (2012). Detecting inherent bias in lexical decision experiments with the LD1NN algorithm. In G. Libben, G. Jarema, & C. Westbury (Eds.), Methodological and Analytic Frontiers in Lexical Research (pp. 231–248). John Benjamins Publishing..

5 Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. Behavior Research Methods, 44, 287-304. doi: 10.3758/s13428-011-0118-4. doi (open access).

6 Keuleers, E., Diependaele, K., & Brysbaert, M. (2010). Practice Effects in Large-Scale Visual Word Recognition Studies: A Lexical Decision Study on 14,000 Dutch Mono- and Disyllabic Words and Nonwords. Frontiers in Psychology, 1. doi (open access).

7 Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. Behavior Research Methods, 42(3), 627-633. Award from the Psychonomic Society for Best Article of 2010 in BRM. doi, preprint.

8 Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. Behavior Research Methods, 42(3), 643-650.

9 Keuleers, E. & Daelemans, W. (2007). Memory-based learning models of inflectional morphology: A methodological case study, Lingue e Linguaggio, 6(2), 151–174. doi, preprint.

10 Keuleers, E., Sandra, D., Daelemans, W., Gillis, S., Durieux, G., & Martens, E. (2007). Dutch plural inflection: The exception that proves the analogy. Cognitive Psychology, 54(4), 283–318.

Research funding granted in the last 5 years

- TRIBAL: Translation Recognition in Bilinguals Across the Lifespan. Awarded by Ministerio de Economía y Competitividad, MINECO, Gobierno de España. Members: Jon Andoni Duñabeitia, Emmanuel Keuleers, Stéphanie Massol, Aina Casaponsa, Ainhara Martí. (2012-2015) [PI: Jon Andoni Duñabeitia][45,000 EUR]

- Erasmus Mundus Basileus Staff Exchange Grant (2009), 3,000 EUR

Awards

- Award from the Psychonomic Society for Best Article of 2010 in Behavior Research Methods. Presented at the 51st Annual Meeting of the Psychonomic Society. St. Louis, November 18, 2010.

ATTACHMENT 3 - LIST OF MATERIALS REVIEWED

1. Materials for the first psycholinguistic study

Several studies have been conducted in which lexical decision data was collected for tens of thousand of words in different languages. In the English Lexicon Project, Balota et al. (2007) collected responses to 40,000 English words and the same number of nonwords, using over 900 participants, each responding to 3,500 trials. In a similar study, the French Lexicon project, Ferrand et al. (2009) collected responses to 38,000 French words and nonwords. Keuleers et al. (2010) collected responses for 39 Dutch speaking participants answering to 14,000 Dutch words and nonwords. Finally, in the British Lexicon Project, Keuleers et al. (2012) collected data for 28,000 English words and nonwords using 78 British English participants. These studies are widely accepted by the psycholinguistic community as reliable tools for the investigation of the visual word recognition system.

For the current analysis, we have used the **data from the British Lexicon Project**. In contrast to the English Lexicon project, the stimuli in the British Lexicon project were presented in lowercase. The characters *l* and *i* are visually more similar in lowercase than in uppercase, presenting a more stringent test of the effect of visual similarity between those characters on word recognition.

The data from the British Lexicon Project, on which we will base the current analysis, are publicly available in the Supplemental Data Archive of the Psychonomic Society, with document object identifier doi:10.3758/s13428-011-0118-4. The details of the experimental procedures, as well as an analysis of the reliability of the results, are given in Keuleers et al. (2012).

2. Materials for the three corpuslinguistic studies

Our **first corpus study** was based on **wordlists**. The two general English word lists we used are the following: EOWL and wordsEN.txt. Both are freely available, fully downloadable and thus open to quantitative analysis.

The word list wordsEN.txt is available from SIL, the international organization on language studies (originally known as Summer Institute of Linguistics) and a pioneer in the field of quantitative analysis of linguistic data.
The list can be downloaded from:
http://www-01.sil.org/linguistics/wordlists/english/

The second word list is the "English Open Word List" (EOWL), developed by Ken Loge, and available from: http://dreamsteep.com/projects/the-english-open-word-list.html

Both word lists contain more than 100,000 words each.

Since there is inevitably some overlap between the two lists, a unique word list was created from both lists, whereby case distinction was ignored.

The resulting file (wordlist-fuse.txt) contains **167,081 unique word forms**.

Word lists	Number of words
Summer Institute of Linguistics – wordsEN	109.582 words
English Open Word List	128.983 words
Fusion of both lists	167.81 **unique words**

For our **second corpus study** in which we study the frequency of word pairs based an the alternation of L/I we decided to use one of the most authoritative corpora for modern English, i. e. the ***British National Corpus*** **(BNC)**. Its creation involved the collaboration of two universities (the University of Oxford and Lancaster University), a consortium of three publishers (Oxford University Press as the lead collaborator, Longman and W.&R. Chambers) and the British Library. BNC is a representative corpus of English covering 100 million words of written and spoken English from a wide variety of sources of the late 20th century.

As a matter of fact, we used the **BNC frequency lists** developed by Adam Kilgarriff, an internationally renowned expert in corpus linguistics and quantitative lexical analysis. For a detailed description, see http://kilgarrifff.co.uk
The frequency lists are available at: ftp://ftp.itri.bton.ac.uk/bnc/

For our **third corpus study** based on a parallel corpus English-Portuguese, we selected the **English-Portuguese part of the JRC-ACQUIS- corpus**. This English-Portuguese sub corpus contains 600M words (300M words for each language). The Acquis Communautaire (AC) is the total body of European Union (EU) law applicable in the the EU Member States. This collection of legislative text changes continuously and currently comprises selected texts written between the 1950s and now. This corpus is coposed by the Joint Research Center of the European Commission.

To our knowledge, the Acquis Communautaire is the biggest parallel corpus in existence, if we take into consideration both its size and the large number of languages involved. The most outstanding advantage of the Acquis Communautaire - apart from being freely available - is the number of rare language pair combinations.

For more info, see: http://ipsc.jrc.ec.europa.eu

ATTACHMENT 4 - RELEVANT PUBLICATIONS OR SOURCES REFERRED TO OR CITED IN SUPPORT

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., … Treiman, R. (2007). The English lexicon project. *Behavior Research Methods, 39*(3), 445–459.

Baayen, R.H. (2008). *Analyzing linguistic data*. Cambridge, Cambridge University Press.

Cattell, J. M. (1886). Über die Trägheit der Netzhaut und des Sehcentrums. *Philosophische Studien, 3*, 94–127.

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M.-A., & Michel, F. (2000). The visual word form area Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain, 123*(2), 291–307. doi:10.1093/brain/123.2.291

Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., … Pallier, C. (2010). The French Lexicon Project: Lexical decision data for 38,840 French words and 38,840 pseudowords. *Behavior Research Methods, 42*(2), 488–496. doi:10.3758/BRM.42.2.488

Finkbeiner, M., & Coltheart, M. (2009). Letter recognition: From perception to representation. *Cognitive Neuropsychology, 26*(1), 1–6. doi:10.1080/02643290902905294

Johnson, K. (2011). *Quantitative Methods In Linguistics*. New York, Wiley.

Keuleers, E., Diependaele, K., & Brysbaert, M. (2010). Practice Effects in Large-Scale Visual Word Recognition Studies: A Lexical Decision Study on 14,000 Dutch Mono- and Disyllabic Words and Nonwords. *Frontiers in Psychology, 1*. doi:10.3389/fpsyg.2010.00174

Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods, 44(1)*, 287–304. doi:10.3758/s13428-011-0118-4

Linquist, H. (2009). *Corpus Linguistics and the Description of English*. Endinburgh, Edinburgh University Press.

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, 7(7), 293–299. doi:10.1016/S1364-6613(03)00134-7

McEnery, T., & Hardie, A. (2011). *Corpus Linguistics: Method, Theory and Practice*. Cambridge, Cambridge University Press.

Mueller, S. T., & Weidemann, C. T. (2012). Alphabetic letter identification: Effects of perceivability, similarity, and bias. *Acta Psychologica, 139*(1), 19–37. doi:10.1016/j.actpsy.2011.09.014

Zhang, J., & Mueller, S. T. (2005). A note on ROC analysis and non-parametric estimate of sensitivity. *Psychometrika, 70*(1), 203–212. doi:10.1007/s11336-003-1119-8

ATTACHMENT 5 – PSYCHOLINGUISTIC ANALYSIS BY DR. EMMANUEL KEULEERS

This document details the studies that I performed, using the methodology, as explained in Prof. Dr. Desmet's synthetic expert report to which this document is attached.

In my first study, I analyzed existing behavioral data on word recognition coming from the British Lexicon Project to study. This analysis is focused on all words having the same length but differing in just one letter.

In the second study, I collected new behavioral data focusing more specifically on the word pair *hotels/hoteis*.

Dr. Emmanuel Keuleers

First study: Analysis of British Lexicon Project

We first selected all words from the British Lexicon Project for which 90% of participants gave a WORD response. In other words, we selected only stimuli for which a word response would be likely if visual similarity would lead to increased confusion. We then paired each of these words to nonwords having the same length but differing in just one letter (formally: having a Hamming distance of 1). Each of these pairs can be said to belong to a *substitution group*. For instance, the pair *bald/baid* belongs to the substitution group *l/i*, while the pair *elite/elire* belongs to the substitution group *t/r*.

For statistical analysis we considered only substitution groups with at least 10 items. We found 44 such substitution groups containing a total of 922 word/nonword pairs. The substitution group *l/i* contained 25 word/nonword pairs. The substitution group *i/l* contained less than 10 items, which we considered insufficient for reliable analysis.

We extracted the average accuracy for the word and nonword of each pair. In signal detection terms, word accuracy corresponds to the *hit rate*, while (1-nonword accuracy) corresponds to the *false alarm* rate. These quantities were used for computation of the *A* index, as described in Zhang & Mueller (2005).

Word	Nonword	WORD responses to word	WORD responses to nonword	A index
bold	boid	1.00	0.00	1.00
bulbs	buibs	0.97	0.00	0.99
fly	fiy	1.00	0.00	1.00
polls	poils	0.95	0.00	0.99
rolls	rolis	1.00	0.00	1.00
sly	siy	0.97	0.00	0.99
tingles	tingies	0.95	0.00	0.99
bled	bied	0.95	0.01	0.98
blew	biew	0.95	0.01	0.98
balm	baim	1.00	0.03	0.99
half	haif	0.95	0.03	0.98
old	oid	1.00	0.03	0.99
calls	cails	1.00	0.05	0.99
spindles	spindies	0.92	0.05	0.96
wells	weils	0.97	0.05	0.98
halts	haits	0.90	0.06	0.95
gold	goid	0.97	0.07	0.97
sold	soid	0.97	0.07	0.97
smalls	smails	0.93	0.08	0.96
angles	angies	0.92	0.10	0.95
bald	baid	0.97	0.10	0.96
dolls	doils	0.90	0.10	0.94
doll	doil	1.00	0.18	0.96
puddles	puddies	1.00	0.28	0.93
handled	handied	1.00	0.38	0.9

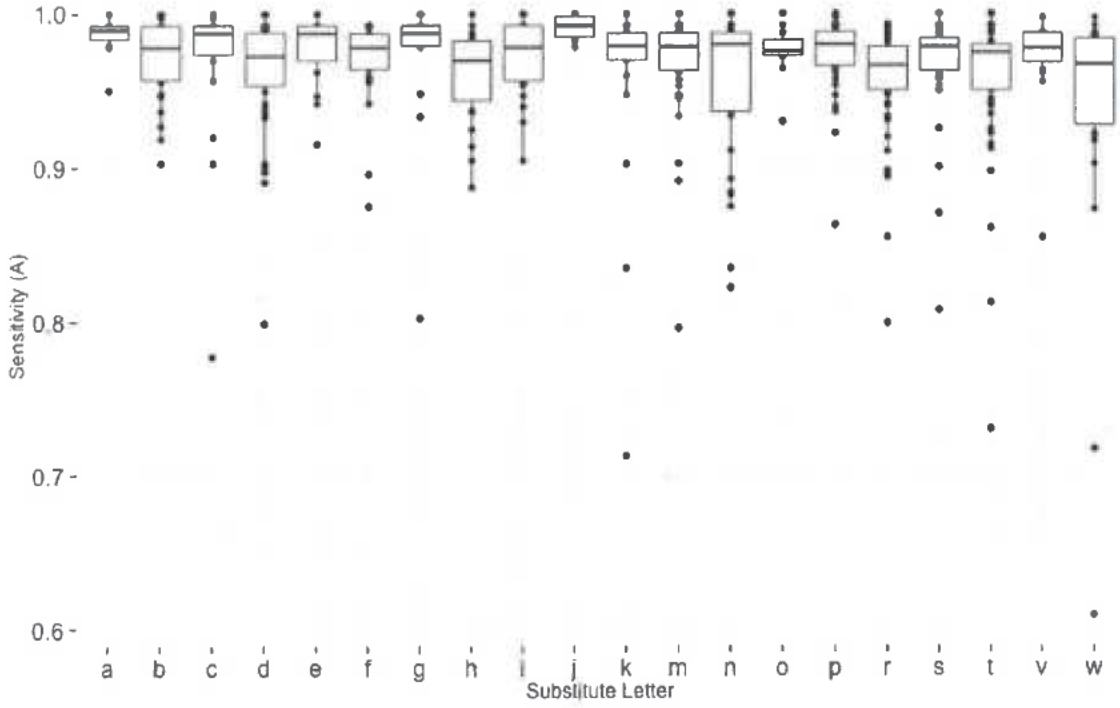


Figure 1 shows a box plot, comparing the sensitivity for all substitution groups for which the original letter was *l*. In general, the *A* indexes are very high for all substitutions. Median values for sensitivity, indicated by the horizontal lines in the boxes, were consistently over 0.95, indicating that none of the character substitutions caused a consistent misidentification of nonwords as words. A one-tailed *t*-test for samples with unequal variances was used to evaluate whether the mean sensitivity of items in the *l/i* substitution group (0.9728) was lower than the mean sensitivity of all items in other substitutions groups deriving from *l* (0.9648). The result of this test was non-significant ($t(31.237)=1.5589$). The non-parametric Mann-Whitney-Wilcoxon test yielded similar results, ($W=7631$, $p=0.8518$). Both tests point to the same result: Participants were not less sensitive to the difference between *l* and *i* characters than to pairs consisting of *l* and another character.

Table 1 shows the pairs included in the *l/i* substitution group. It is interesting to note that the variation in the number of WORD responses to nonwords does not appear to be a function of string similarity, which should be more or less equal for all forms. Instead, the results suggest that the proportion of WORD responses to a nonword reflects the degree to which participants can imagine that the string is actually a likely word. In other words, the variance in WORD responses to nonwords is most likely due to guessing.

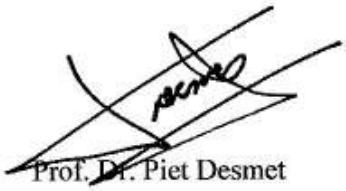
Table 1: Word/nonword pairs included in the l/i substitution group, together with the number of WORD responses to the word and the confusable nonword, and the sensitivity (A index).

Second study: Results on the word pair *hotels/hoteis* using an online vocabulary test

In addition to the results outlined above, we collected specific data using an online English vocabulary test. The design of this test, known as the *Ghent University Vocabulary Test*, is very similar to a lexical decision experiment, but it is presented in an educational game-like format. The test is taken daily by hundreds of internet users from all over the world, using many different devices. As such, the participants to the test represent a varied sample of internet users. From February 24th until February 25th 2014, we used the website to test specifically for the character strings *hoteis* and *hotels*. For each participant, we randomly picked one of these strings to add to the 100 items they would normally see during the test. The added strings were not used for scoring their vocabulary. Over the course of a day, we collected 853 responses to the string *hoteis* and 802 responses to the string *hotels*. For the string *hotels* participants gave 797 WORD responses (99.4%). For the string *hoteis* participants gave 47 WORD responses (5.5%). This is lower than the baseline proportion of WORD responses to nonwords during the same period (8.9%), which can be attributed to guessing. Since *hotéis* is a word in Portuguese, we were interested in what Portuguese speaking participants answered to this form. Interestingly, out of the 10 participants we had from Brazil during this period, the 5 who were given the string *hotels* all gave a WORD response and out of the 6 who were given *hoteis* 5 gave a WORD response. In contrast, out of the 50 participants we had from Germany during the same period, not one of the 23 participants who were given *hoteis* gave a WORD response while all but one of the remaining 27 participants gave a WORD response to *hotels*. The results suggest very strongly that character differences play a negligible role in ordinary word recognition, whereas linguistic differences predictably inform decisions.

ATTACHMENT 6: CORPUS-LINGUISTIC ANALYSIS BY PROF. DR. PIET DESMET

This document details the corpus-linguistic analysis that I performed, using the methodology explained in the expert's report to which this document is attached.



Prof. Dr. Piet Desmet

First study: analysis of English wordlists - number of word pairs differing only in the alternation *l/i*

In order to establish the number of word pairs differing only in the alternation *l/i* – which we will call "minimal word pairs with *l/i*-alternation", a list of regex (regular expression) patterns was created. By regular expression we mean a sequence of characters that forms a search pattern. In this case, the characters under study (i.e. *l* and *i*) were neutralized to a dot. This allows us to identify the number of similar patterns.
The following **procedure** was used to collect the patterns. First, all occurrences of *l* and *i* were replaced by dots, hence neutralizing their meaning. A frequency list of unique word patterns was created, and only the minimal word pairs were stored in the file 'wordlists-patterns-li.txt'. A minimal word pair is a pattern matching at least two word forms after neutralization of *l* and *i*. This can be illustrated with the following example: *candies* and *candles* only differ in the fifth character position, indicating *i* and *l* respectively. The two words match the regular expression pattern /cand[li]es/, in which "[li]" represents one character, which is either *l* and *i*. Since we only compare the distinction *l* and *i*, we can just as well use a dot instead, as represented by the following record:

```
==== 2 cand.es
candies
candles
```

The record for each pattern consists of a header, showing the pattern, preceded by the number of words matching the pattern. The header is followed by a list of the possible words matching the pattern. On the basis of the word pattern file wordlists-patterns-li.txt all word forms matching the selected pattern are searched for in the fused word list and stored in wordlist-containing-li.txt. For each minimal word pair, a record was created.

This resulted in a list of **390 minimal word pairs with *l/i*-alternation, covering 788 word forms in total.**

Note that the above example shows a typical pattern of a minimal pair. However, the character opposition can occur at different positions in the same word, as illustrated by the following example, where "compiled" and "complied" match the pattern /comp[il][il]ed/:

== 2 comp..ed
compiled
complied

This shows that the confusion is not limited to minimal pairs in the strict sense of the term.

Some **typical examples** are the following records:

== 2 a..
ail
all
== 2 ba..s
bails
balls
== 2 cand.es
candies
candles
== 2 ce...ng
ceiling
celling
== 2 ent.t.es
entities
entitles
== 2 fa..
fail
fall
== 2 f..ed
filed
flied
== 2 f..er
filer
flier
== 2 hand.er
handier
handler
== 2 mudd.ed
muddied
muddled
== 2 padd.es
paddies
paddles
== 2 ro...ng
roiling
rolling
== 2 sp..t
spilt
split
== 2 s..ver
silver
sliver
== 2 ta..er
tailer
taller

== 2 to..
toil
toll

Note that in a few cases, the pattern matches three words, as in the following example which includes the proper noun Mali:

```
=== 3 ma..
Mail
mali
mall
```

This data shows that English has a substantial number of word pairs differing only in the alternation *l/i*.

Second study: British National Corpus (BNC) -) frequency of the words affected by the alternation *l/i*

One could argue that the selected word lists also contain very low frequency words. *E.g.* the word "eider" (referring to a species of duck) in the following word pair is not that frequently used:

```
=== 2 e.der
cider
elder
```

Therefore, we also decided to analyze the BNC frequency lists. This should allow us to examine to what extent the minimal word pairs with *l/i*-alternation concern only very low frequency words or also involve words with a higher frequency.

We selected the BNC frequency list for words having a frequency of occurrence of 5 or more, meaning that the word occurs at least 5 times in the corpus. Below this frequency it is quite hard to do statements about word frequency with sufficient precision.

The BNC frequency list contains 4 columns per word form. Each column corresponds to the following fields: frequency, word, part-of-speech (*i.e.*, verb, noun, etc.), number of files the word occurs in.

In order to test the patterns *l/i*, we first created a list containing the unique word forms from BNC. In other words, we selected the second column of the BNC frequency file and stored the unique word forms in the file bnc-uwf.txt, which resulted in a list of 131,236 unique word forms. The number of unique word forms for BNC is smaller than those in the fused word lists. This can be explained, because we selected the frequency lists of BNC containing only words having a minimal frequency of 5 for the whole corpus.

We then created a pattern list from bnc-uwf.txt, which resulted in **514 patterns differing only in the *l/i* alternation.** These were stored in wordlists-patterns-li-bnc.txt. Each pattern is preceded by the number of word forms matching the regex (regular expression) pattern. In total, in BNC there are **1,045 forms** matching the pattern *l/i*.

However, we should next verify to what extent the identified patterns correspond to English words as listed in a dictionary. For this operation, we took *Merriam Websters* as a reference (http://www.merriam-webster.com/), which also includes proper names and abbreviations. This matching operation resulted in a final list of **123 word pairs differing only in the *l/i* alternation.**

On the basis of the pattern list file wordlists-patterns-li-bnc.txt, all word forms in BNC, together with their frequency were selected and stored in the file wordlist-containing-li-bnc-freq.txt.

This list allows us to evaluate to what extent the word pairs with the *l/i* -alternation are actually used in reality. This information was mapped as set out in the examples below, where each word form is preceded by the BNC frequency.

The following examples show that the proper nouns Bali and Mali are less frequently used than the other words of the set:

== ba..
7563 ball
1023 bail
159 bali
== ma..
3405 mail
293 mall
166 mali

The following two examples show that *ladles* and *alms* are less frequently used than the alternative words of the same pattern:

== .ad.es
3281 ladies
12 ladles
== a.ms
4207 aims
86 alms

The following examples give clear information on the frequency, when compared to the same set in the first vocabulary lists:

== comp..ed
855 compiled
368 complied
== comp..es
105 complies
20 compiles
== e.der
1354 elder
56 eider
== ent.t.es
642 entities
119 entitles

The following records show that *fall* in its different forms is more often used than *fail*, but that both words are regularly used in different forms:

```
=== fa..
11119 fall
3370 fail
=== fa...ng
4745 falling
2264 failing
=== fa..s
3093 falls
1861 fails
```

These examples show that not only the alternation *l/i* itself is a frequent pattern, but that it also occurs in frequent words.

Third study: Parallel corpora English-Portuguese – frequency of inter-lingual English-Portuguese orthographic neighbors

In conducting our inter-lingual study, we followed the following procedure. We first selected all words ending in /[aeou]ls/ in an English sentence. Then, we selected all Portuguese words ending in /[aeou]is/ in the corresponding parallel sentence(s). These Portuguese words were then transformed to their English format ("i" maps to "l" and accents are discarded), and checked whether a match was found between the Portuguese and English word list. On the basis of this filter, **53 type pairs** (types: different word pairs) were found, representing 18,668 token pairs. The following list shows the 53 pairs sorted on reversed frequency, the most frequent pair on top. There are three columns used: rank number, frequency and type pair:

Rank number	Frequency	Type pair
1	9966	animais animals
2	5135	cereais cereals
3	2077	materiais materials
4	290	metais metals
5	221	minerais minerals
6	196	terminais terminals
7	122	hospitais hospitals
8	104	totais totals
9	94	cartéis cartels
10	66	individuais individuals
11	63	originais originals
12	51	manuais manuals
13	49	hotéis hotels
14	41	capitais capitals
15	35	tribunais tribunals
16	19	portais portals
17	18	canais canals
18	16	festivais festivals
19	14	fundamentais fundamentals
20	13	ideais ideals

Rank number	Frequency	Type pair
21	7	decimais decimals
22	6	animais animals
23	5	diagonais diagonals
24	5	quintais quintals
25	4	lintéis lintels
26	3	arsenais arsenals
27	3	decibéis decibels
28	3	finais finals
29	3	nanomateriais nanomaterials
30	3	ornamentais ornamentals
31	3	materiais materials
32	3	subtotais subtotals
33	3	transsexuais transsexuals
34	2	industriais industrials
35	2	marginais marginals
36	2	matériais materials
37	2	pastéis pastels
38	2	principais principals
39	2	cereais cereals
40	2	verticais verticals
41	1	aerosóis aerosols
42	1	corais corals
43	1	culturais culturals
44	1	dióis diols
45	1	géis gels
46	1	hoteis hotels
47	1	memoriais memorials
48	1	mongóis mongols
49	1	motéis motels
50	1	mutuais mutuals
51	1	pixéis pixels
52	1	radicais radicals
53	1	rivais rivals

It can be observed that English and Portuguese present a significant number of inter-lingual cognates, i.e. completely analogous word pairs with exactly the same number of letters, differing only through the alternation of *l* and *i*. Most of these word pairs are quite frequent and concern the *ais/ail* alternation. 8 of them offer the *eis/els* alternation.

As a result, language users who speak both English and Portuguese are used to the alternation of *l* and *i* between English and Portuguese. This confirms the empirical data that they are not confused by this alternation.