To: CLARIN Board of Directors
From: CLARIN Standards Committee

Proposal for a stable basis for namespaces within the clarin.eu domain

The CLARIN Standards Committee wishes to open and maintain a facility for regulating the use of XML namespaces based on the "clarin.eu" domain. XML namespaces act as identifiers/discriminators for XML vocabulary (tag and attribute names) and are most commonly known in the form of prefixes, as in "<tei:body>", which is the <body> element coming from vocabulary defined by the TEI Consortium vs. e.g. "<html:body>", which may be used to denote a <body> element as defined by the W3C Consortium. Such prefixes are merely a shorthand way of tying the given tag or attribute name to longer identifiers that usually (although not necessarily) take the form of a URL, and as such, are very often treated as web addresses. The addressing function should be treated as mere convenience that is very often used to document the given kind of vocabulary in a corresponding web document.

The CSC would like to regulate the use of XML namespaces that take the form of URLs beginning with "http://clarin.eu/" (the fact that the address "http://clarin.eu" resolves to "https://www.clarin.eu/" is of no consequence in this regard -- we concentrate on the initial form of the identifier) in the following manner:

All namespaces whose usage is regulated by the CSC begin with the sequence:

```
http://clarin.eu/standards/
```

and the above sequence itself (or any sequence shorter than that) is not used as a namespace string. Whatever comes after this initial sequence will be referred to below as the "suffix". Thus, we postulate the following, where "+" signals string concatenation:

```
CLARIN Namespace: base + suffix

base: http://clarin.eu/standards/
```

The suffix can have internal structure, e.g. "projectX/spec1/ns", where "projectX" is e.g. an abbreviated name of the project in question, "spec1" indicates one of the XML vocabulary sets defined by that project, and "ns" is just a mnemonic signalling that we are looking at a namespace, even though it looks like a web address. Naturally, "projectX/ns" is also a possible suffix. Numbers after "ns" may be used in cases where the vocabulary in question undergoes evolution while the names of the project and specification do not change. The use of "ns" and potential numbers following it is proposed as a useful convention rather than

a formal requirement. The first part of the suffix, however (full or abbreviated project name) *is* a requirement.

The CSC wishes to regulate the usage of CLARIN namespaces through the adoption of the principle that each namespace whose basis is "http://clarin.eu" must resolve to a public document providing, minimally,

- a description of the vocabulary in question (the description may be brief and use further web references, but should be informative enough to provide at least a general idea of what the vocabulary is used for and to indicate (roughly) the potential range of documents making use of the namespace)
- the name of the project or organization submitting the request for a CLARIN namespace (henceforth "requesting agency")
- the date of the decision by the CSC to assign the namespace in question to the requesting agency.

The content of the proposed document associated with the requested namespace should be provided by the requesting agency together with the request for the individual namespace string. Each such case will be subject to a discussion and a possible vote by the Committee. The Committee may request additional information from the requesting agency. A positive decision will be signalled by the presence of the submitted text (or a version thereof) at a web address to which the requested namespace resolves.

We would like to request permission from the BoD to implement this idea together with the technical infrastructure section responsible for the clarin.eu domain and for serving the relevant web pages.