

SLIM
Internet-Draft
Intended status: Standards Track
Expires: December 12, 2017

G. Hellstrom
Omnitor
June 10, 2017

Human Language Modality Grouping Semantics in Session Description
Protocol
draft-hellstrom-language-grouping-00

Abstract

In a real-time communication session, there may be a need or preference for receiving the same content in two or more simultaneous modalities. There may also be a need to prioritize which media to use for language communications. This document defines the semantics for grouping media in the Session Description Protocol (SDP) containing human languages in order to assign their relative priority and also grouping media that are desired to be received together. The semantics defined in this document are to be used with the SDP Grouping Framework. Applications are for example a possibility to indicate when a sign language is most desired, but written language an acceptable lower rated alternative, offered for the other party to select between. Applications are also for indication of need for both spoken and written content in a real-time call (captioning) together, or provision of both spoken language original and sign language interpretation of the original language together in a real-time session. These indications are specified for the sending and receiving direction separately.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on December 12, 2017.

Internet-Draft

Modality Grouping

June 2017

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	2
2.	Terminology	4
3.	Requirements for Modality Grouping	4
3.1.	Simultaneous Use of Different Modalities	4
3.2.	Preference for Language in Different Modality	4
4.	Modality Grouping	4
4.1.	Simultaneous Use of Different Modalities	4
4.2.	Preference for Language in Different Modality	5
5.	SDP Offer/Answer Considerations	5
6.	Examples	6
6.1.	Desire by caller to receive both spoken and written language form of media	6
6.2.	High preference by caller to receive sign language and lower preference for text.	7
7.	Acknowledgements	8
8.	IANA Considerations	9
9.	Security Considerations	9
10.	References	9
10.1.	Normative References	9
10.2.	Informative References	10
	Author's Address	10

[1.](#) Introduction

In certain applications it is of interest to indicate a need for, or the availability of, transformed version of the contents of a media

stream in another media, while still also providing the original.

The application of this indication may for example be for rapid subtitling of speech either manually or automatically. It may also be sign language interpretation of speech, or spoken interpretation

of sign language when both the original and the interpretation is delivered to the user.

This specification defines an indication that language contents in one modality is desired simultaneously with a different modality. The mechanism used is based on the Session Description Protocol (SDP) Grouping Framework [[RFC5888](#)] and used with SDP [[RFC4566](#)]. The same indication is used for indication of preparedness to send language contents in one modality simultaneously with same content in a different modality.

When starting a conversation in a media-rich environment, the users may have very specific preferences for using one modality (spoken, written or signed) over other possible but less preferred modalities. In traditional call establishment, it is the answering part who is expected to start the conversation by a greeting. In the media-rich environment, the modality and language of this greeting sets the expectations for what modality and language to mainly use in the session. Deviation from this initial expectation is usually possible during the session by mutual agreement between the participants, but may be time consuming and cause uncertainty.

A way for the parties to not only indicate alternative languages and modalities for the communication directions in the session, but also indicate preference for specific modalities per direction provides the opportunity to more exactly describe the desired language communication for a session, while still providing information about less preferred alternatives. This specification defines a mechanism for indicating modality preference based on the Session Description Protocol (SDP) Grouping Framework [[RFC5888](#)].

The expected application area is wide. By old tradition, the most common modality for real-time interaction is spoken communication. In some settings, e.g. where silence is required, it may be desirable to express a preference for using written communication, while still leaving a possibility open for traditional spoken communication by an

indication on lower preference level. For persons having full ability to both use sign language and spoken language, but not wanting to force the other party to bring in a sign language interpreter in the call, it may be of importance to be able to indicate the sign language capability on a lower preference level and the spoken language capability on a higher level. Some persons with disabilities may strongly prefer to conduct a written conversation, while still wanting to express that a spoken conversation is possible as a last resort. Many other situations exist in the media-rich communication environment when the media preference indication is of value for a smooth initiation of a real-time session.

The mechanisms for specifying simultaneous use of language in different modalities and preference between modalities may be combined with a mechanism for specifying language use in media.

[2. Terminology](#)

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [[RFC2119](#)].

[3. Requirements for Modality Grouping](#)

[3.1. Simultaneous Use of Different Modalities](#)

The grouping semantics for indication of simultaneous use of different media for different transforms of the same language shall have the ability to indicate:

- o That the language contents of one media is desired or offered also simultaneously in a transformed form in other media grouped with the first.
- o The direction of language communication that the indication is valid for; sending or receiving as seen from the party providing the SDP.

[3.2. Preference for Language in Different Modality](#)

The grouping semantics for indication of relative preference between

use for language communication in different media and modalities shall have the ability to indicate:

- o The order of preference for receiving or providing language contents in the media included in a grouping.
- o The direction of language communication that the indication is valid for; sending or receiving as seen from the party providing the SDP.

4. Modality Grouping

4.1. Simultaneous Use of Different Modalities

The "Human Language Simultaneous Send" (HLSS) and "Human Language Simultaneous Receive" (HLSR) grouping semantics and the SDP "group" attribute defined in [[RFC5888](#)] are used to associate media in which it is indicated that different transforms of the same content is

either desired to be received by a party or offered for sending by a party.

The "a=group:HLSS" semantics SHOULD be used to indicate media grouping for preparedness for sending of same language contents in different transforms in all media included in the group.

The "a=group:HLSR" semantics SHOULD be used to indicate media grouping for preference for reception of language contents in different transforms in all media in the group.

The HLSS and HLSR semantics MAY be used together with mechanisms for detailing language use in media. One such mechanism is [[I-D.ietf-slim-negotiating-human-language](#)].

4.2. Preference for Language in Different Modality

The "Human Language Preferred Send" (HLPS) and "Human Language Preferred Receive" (HLSR) grouping semantics and the SDP "group" attribute defined in [[RFC5888](#)] are used to associate media among which it is indicated an order of preference for using the media for language contents. The order of preference is that the media

identity first in the group has highest preference and the following have lower preference in the same order as they appear in the group definition.

The "a=group:HLPS" semantics SHOULD be used to indicate media grouping for preparedness for sending of language contents with preference in the same order as the media identities appear in the group with the first having highest preference. .

The "a=group:HLSR" semantics SHOULD be used to indicate media grouping for preference for reception of language contents with preference in the same order as the media identities appear in the group with the first having highest preference.

The HLSS and HLSR semantics MAY be used together with mechanisms for detailing language use in media. One such mechanism is [\[I-D.ietf-slim-negotiating-human-language\]](#).

5. SDP Offer/Answer Considerations

The following SDP offer/answer considerations according to [\[RFC3264\]](#) apply.

An application that understands the received HLSR, HLSS, HLPR or HLPS grouping semantics SHOULD make efforts to satisfy the preferences expressed by the grouping semantic.

HLSR, HLSS, HLPR or HLPS grouping semantics corresponding to what the application prefers to receive and what the application is prepared to send, best matching the received preference indications and its own capabilities SHOULD be included in the answer.

The offering party SHOULD analyze the answer and make best effort to transmit language contents in media according to the answer.

The grouping semantics defined in this document are only informing about language contents disposition in media and SHOULD not be taken as reasons to enable or reject media streams.

Media not included in any HLPR or HLPS grouping are assumed to be assigned lower preference for being used for language communication than the ones included in HLPR or HLPS grouping.

If the HLSR, HLSS, HLPR or HLPS grouping semantics are used without any further language specifications, video media SHOULD be assumed to be used for sign language.

Note that grouping of "m" lines MUST always be requested by the offerer, but never by the answerer. Since SIP provides a two-way SDP exchange, an answerer that requested grouping would not know whether the "group" attribute was accepted by the offerer or not. An answerer that wants to group media lines issues another offer after having responded to the first one (in a re-INVITE, for instance).

6. Examples

6.1. Desire by caller to receive both spoken and written language form of media

```
v=0

o=Laura 289083124 289083124 IN IP4 two.example.com

c=IN IP4 233.252.0.1/127

t=0 0

a=group:HLSR 1 2

m=audio 30000 RTP/AVP 0

a=mid:1

m=text 30002 RTP/AVP 98 97
```

```
a=mid:2

m=video 30004 RTP/AVP 34

a=mid:3
```

Note that also the video media needs to include a 'mid' attribute even when it is not included in any grouping for the grouping to be

valid.

An answer can confirm that both desired media will contain the same language contents.

```
v=0  
  
o=Laura 289083124 289083124 IN IP4 two.example.com  
  
c=IN IP4 233.252.0.1/127  
  
t=0 0  
  
a=group:HLSS 1 2  
  
m=audio 30000 RTP/AVP 0  
  
a=mid:1  
  
m=text 30002 RTP/AVP 98 97  
  
a=mid:2  
  
m=video 30004 RTP/AVP 34  
  
a=mid:3
```

- [6.2.](#) High preference by caller to receive sign language and lower preference for text.

```
v=0  
  
o=Laura 289083124 289083124 IN IP4 two.example.com  
  
c=IN IP4 233.252.0.1/127  
  
t=0 0  
  
a=group:HLPR 3 2
```

```
m=audio 30000 RTP/AVP 0
```


a=mid:1

m=text 30002 RTP/AVP 98 97

a=mid:2

m=video 30004 RTP/AVP 34

a=mid:3

Note that also the audio media needs to include a 'mid' attribute even when it is not included in any grouping for the grouping to be valid.

An answer can confirm that sign language will be sent in the video media.

v=0

o=Laura 289083124 289083124 IN IP4 two.example.com

c=IN IP4 233.252.0.1/127

t=0 0

a=group:HLPS 3

m=audio 30000 RTP/AVP 0

a=mid:1

m=text 30002 RTP/AVP 98 97

a=mid:2

m=video 30004 RTP/AVP 34

a=mid:3

7. Acknowledgements

-

8. IANA Considerations

This document registers the following semantics with IANA in the "Semantics for the "group" SDP Attribute" registry under SDP Parameters:

Semantics Token Reference

Human Language Simultaneous Send HLSS TBD: THIS DOCUMENT

Human Language Simultaneous Receive HLSS TBD: THIS DOCUMENT

Human Language Preference Send HLPS TBD: THIS DOCUMENT

Human Language Preference Receive HLPR TBD: THIS DOCUMENT

9. Security Considerations

Modality preference information may belong to the kind of sensitive user information that some users do not want to be presented to anyone. Measures for protection against unauthorized access to the modality preference information should therefore be prepared and activated when so required. Intended callees should be regarded to be authorized to access the callers modality preference information. The modality preference information should be treated with similar security and privacy measures as other user information such as addresses and language preferences.

10. References

10.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), DOI 10.17487/RFC2119, March 1997, <<http://www.rfc-editor.org/info/rfc2119>>.
- [RFC3264] Rosenberg, J. and H. Schulzrinne, "An Offer/Answer Model with Session Description Protocol (SDP)", [RFC 3264](#), DOI 10.17487/RFC3264, June 2002, <<http://www.rfc-editor.org/info/rfc3264>>.
- [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", [RFC 4566](#), DOI 10.17487/RFC4566,

July 2006, <<http://www.rfc-editor.org/info/rfc4566>>.

Hellstrom

Expires December 12, 2017

[Page 9]

Internet-Draft

Modality Grouping

June 2017

[RFC5888] Camarillo, G. and H. Schulzrinne, "The Session Description Protocol (SDP) Grouping Framework", [RFC 5888](#), DOI 10.17487/RFC5888, June 2010, <<http://www.rfc-editor.org/info/rfc5888>>.

10.2. Informative References

[I-D.ietf-slim-negotiating-human-language]
Gellens, R., "Negotiating Human Language in Real-Time Communications", [draft-ietf-slim-negotiating-human-language-10](#) (work in progress), May 2017.

Author's Address

Gunnar Hellstrom
Omnitor
Hammarby Fabriksvag 23
Stockholm SE-120 30
Sweden

Phone: +46 708 204 288
Email: gunnar.hellstrom@omnitor.se

Hellstrom

Expires December 12, 2017

[Page 10]