

MMUSIC Working Group
Internet-Draft
Intended status: Standards Track
Expires: August 23, 2013

R. Gellens
Qualcomm Technologies, Inc.
February 21, 2013

Negotiating Human Language Using SDP
draft-gellens-negotiating-human-language-01

Abstract

Users have various human (natural) language needs, abilities, and preferences regarding spoken, written, and signed languages. When establishing interactive communication "calls" there needs to be a way to communicate and ideally match (i.e., negotiate) the caller's language needs, abilities, and preferences with the capabilities of the called party. This is especially important with emergency calling, where a call can be routed to a PSAP or call taker capable of communicating with the user, or a translator or relay operator can be bridged into the call during setup, but this applies to non-emergency calls as well (as an example, when calling an airline reservation desk).

This document describes the need and expected use, and discusses the solution using either an existing or new SDP attribute.

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 August 23, 2013.

Copyright Notice

Copyright (c) 2013 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.	Expected Use	4
4.	Desired Semantics	4
5.	Proposed Solution	4
5.1.	Possibility: Re-Use existing 'lang' attribute	5
5.2.	Possibility: Define new 'humintlang' attribute	6
6.	Silly States	6
7.	IANA Considerations	7
8.	Security Considerations	7
9.	Changes from Previous Versions	7
10.	References	7
10.1.	Normative References	7
10.2.	Informational References	7
	Author's Address	8

[1.](#) Introduction

When setting up interactive communication sessions, human (natural) language negotiation is needed in some cases. When the caller and callee are known to each other or where context implies language, such language negotiation may not be needed. In other cases, there is a need for the caller to indicate language preferences, abilities, or needs, including specific spoken, signed, or written languages. This need exists when setting up SIP or other sessions (including emergency and non-emergency calling). For various reasons, including the ability to establish multiple streams each using a different media (e.g., voice, text, video), it makes sense to use a per-stream negotiation mechanism, using SDP.

This approach has a number of benefits, including that it is generic and not limited to emergency calls. In some cases such a facility isn't needed, because the language is known from the context (such as when a caller places a call to a sign language relay center). But it seems clearly useful in many other cases. For example, it seems generally useful that someone calling a company call center be able to indicate if a specific sign and/or spoken language is needed. The

UE would need to set this, but could default to the language used for the interface with the user.

Including the user's human (natural) language requirements in the session establishment negotiation is independent of the use of a relay service and is transparent to a voice service provider. For example, assume a user within the United States who speaks Spanish but not English places a voice call using an IMS device. It doesn't matter if the call is an emergency call or not (e.g., to an airline reservation desk). The language information is transparent to the IMS carrier, but is part of the session negotiation between the UE and the terminating entity. In the case of a call to e.g., an airline, the call can be automatically routed to a Spanish-speaking agent. In the case of an emergency call, the ESInet and the PSAP may choose to take the language into account when determining how to route and process the call (e.g., language and media needs may be considered within policy-based routing).

By treating language as another attribute that is negotiated along with other aspects of a media stream, it becomes possible to accommodate a wide range of users' needs and called party facilities. For example, some users may be able to speak several languages, but have a preference. Some called parties may support some of those languages internally but require the use of a translation service for others, or may have a limited number of call takers able to use certain languages. Another example would be a user who is able to speak but is deaf or hard-of-hearing and requires a voice stream plus a text stream (known as voice carry over). Making language a media attribute allows the standard session negotiation mechanism to handle this by providing the information and mechanism for the endpoints to make appropriate decisions.

Regarding relay services, in the case of an emergency call requiring sign language such as ASL, there are two common approaches: the caller initiates the call to a relay center, or the caller places the call to emergency services (e.g., 911 or 112). In the former case, the language need is ancillary and supplemental. In the latter case, the ESInet and/or PSAP may take the need for sign language into account and bridge in a relay center. In this case, the ESInet and PSAP have all the standard information available (such as location) but are able to bridge the relay sooner in the call processing.

By making this facility part of the end-to-end negotiation, the question of which entity provides or engages the relay service becomes separate from the call processing mechanics; if the caller directs the call to a relay service then the human language facility provides extra information to the relay service but calls will still function without it; if the caller directs the call to emergency services, then the ESInet/PSAP are able to take the user's human

Gellens

Expires August 23, 2013

[Page 3]

language needs into account, e.g., by routing to a particular PSAP or call taker or bridging a relay service or translator.

The term "negotiation" is used here rather than "indication" because human language (spoken/written/signed) is something that can be negotiated in the same way as which forms of media (audio/text/video) or which codecs. For example, if we think of non-emergency calls, such as a user calling an airline reservation center, the user may have a set of languages he or she speaks, with perhaps preferences for one or a few, while the airline reservation center will support a fixed set of languages. Negotiation should select whichever language supported by the call center is most preferred by the user. Both sides should be aware of which language was negotiated. This is conceptually similar to the way other aspects of each media stream are negotiated using SDP (e.g., media type and codecs).

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 [RFC 2119](#) [[RFC2119](#)].

3. Expected Use

This facility is expected to be used by NENA and 3GPP. NENA is likely to reference it in NENA 08-01 (i3 Stage 3) in describing attributes of calls presented to an ESInet, and in that or other documents describing Policy-Based Routing capabilities within a Policy-Based Routing Function (PCRF). 3GPP is expected to reference this mechanism in general call handling and emergency call handling. Recent CRs introduced in SA1 have anticipated this functionality being provided within SDP.

4. Desired Semantics

The desired solution is a media attribute that may be used within an offer to indicate the preferred language of each media stream, and within an answer to indicate the accepted language. The semantics of including multiple values for a media stream within an offer is that the languages are listed in order of preference.

(While it is true that a conversation among multilingual people often involves multiple languages, it does not seem useful enough as a general facility to warrant complicating the desired semantics of the SDP attribute to allow negotiation of multiple simultaneous languages within an interactive media stream.)

5. Proposed Solution

An SDP attribute seems the natural choice to negotiate human

(natural) language of an interactive media stream. The attribute

value should be a language tag from the IANA registry [IANA-lang-tags]

5.1. Possibility: Re-Use existing 'lang' attribute

[RFC 4566](#) [[RFC4566](#)] specifies an attribute 'lang' which sounds similar to what is needed here, the difference being that it specifies that 'a=lang' is declarative with the semantics of multiple 'lang' attributes being that all of them are used, while we want a means to negotiate which one is used in each stream. This difference means that either the existing 'lang' attribute can't be used and we need to define a new attribute, or we finess/update the semantics of 'lang' such that for non-interactive streams, multiple 'lang' values means all are used, while for interactive streams, one is used (or possibly the author has misunderstood [RFC 4566](#)).

The text from [RFC 4566](#) [[RFC4566](#)] is:

a=lang:<language tag>

This can be a session-level attribute or a media-level attribute. As a session-level attribute, it specifies the default language for the session being described. As a media-level attribute, it specifies the language for that media, overriding any session-level language specified. Multiple lang attributes can be provided either at session or media level if the session description or media use multiple languages, in which case the order of the attributes indicates the order of importance of the various languages in the session or media from most important to least important.

The "lang" attribute value must be a single [[RFC3066](#)] language tag in US-ASCII [[RFC3066](#)]. It is not dependent on the charset attribute. A "lang" attribute SHOULD be specified when a session is of sufficient scope to cross geographic boundaries where the language of recipients cannot be assumed, or where the session is in a different language from the locally assumed norm.

The question is: Can the 'lang' attribute be used for our purposes? Using it to negotiate the language for a media seems at first glance to violate its semantics as defined in [RFC 4566](#) [[RFC4566](#)]. But there are existing examples of it being used in exactly the way we need. For example, [draft-saintandre-sip-xmpp-chat-04](#) [I-D.saintandre-sip-xmpp-chat] contains an example where the initial invitation contains two 'a=lang' entries for a media stream (for English and Italian) and the OK accepts one of them (Italian), which matches what we need:

Example: (F1) SIP user starts the session

Gellens

Expires August 23, 2013

[Page 5]

```
INVITE sip:juliet@example.com SIP/2.0
To: <sip:juliet@example.com>
From: <sip:romeo@example.net>;tag=576
Subject: Open chat with Romeo?
Call-ID: 742507no
Content-Type: application/sdp

c=IN IP4 s2x.example.net
m=message 7313 TCP/MSRP *
a=accept-types:text/plain
a=lang:en
a=lang:it
a=path:msrp://s2x.example.net:7313/ansp71weztas;tcp
```

Example: (F2) Gateway accepts session on Juliet's behalf

```
SIP/2.0 200 OK
To: <sip:juliet@example.com>;tag=534
From: <sip:romeo@example.net>;tag=576
Call-ID: 742507no
Content-Type: application/sdp

c=IN IP4 x2s.example.com
m=message 8763 TCP/MSRP *
a=accept-types:text/plain
a=lang:it
a=path:msrp://x2s.example.com:8763/lkjh37s2s20w2a;tcp
```

5.2. Possibility: Define new 'humintlang' attribute

Instead of re-using 'lang' we may define a new media-level attribute 'humintlang' (for "human interactive language") to negotiate which human language is used in each (interactive) media stream:

```
a=humintlang:<language tag>
```

This is a media-level attribute. In an offer, it specifies the desired language(s) for the media. Multiple humintlang attributes can be provided in an offer for a media stream, in which case the order of the attributes indicates the order of preference of the various languages from most preferred to least preferred. Within an answer it indicates the accepted language for the media.

The "humintlang" attribute value must be a single [RFC 3066](#) [RFC3066] language tag in US-ASCII [RFC3066]. It is not dependent on the charset attribute. A "humintlang" attribute SHOULD be specified when placing an emergency call (to avoid ambiguity) or in any other case where the language cannot be assumed from context.

6. Silly States

Gellens

Expires August 23, 2013

[Page 6]

It's possible to specify a "silly state" where the language specified does not make sense for the media type, such as specifying a signed language for an audio media stream.

An offer SHOULD NOT be created where the human language does not make sense for the media type. If such an offer is received, the receiver MAY ignore the language specified, or MAY attempt to interpret the intent (e.g., if American Sign Language is specified on an audio media stream, this might be interpreted as a desire to use English).

7. IANA Considerations

TBD.

8. Security Considerations

TBD

9. Changes from Previous Versions

Changes from -00 to -01:

- o Changed name of (possible) new attribute from 'humlang' to "humintlang"
- o Added discussion of silly state (language not appropriate for media type)
- o Added Voice Carry Over example
- o Added mention of multilingual people and multiple languages
- o Minor text clarifications

10. References

10.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC3066] Alvestrand, H., "Tags for the Identification of Languages", [RFC 3066](#), January 2001.
- [RFC4566] Handley, M., Jacobson, V. and C. Perkins, "SDP: Session Description Protocol", [RFC 4566](#), July 2006.

10.2. Informational References

[I-D.iab-privacy-considerations]

Cooper, A., Tschofenig, H., Aboba, B., Peterson, J., Morris, J., Hansen, M. and R. Smith, "Privacy Considerations for Internet Protocols", Internet-Draft [draft-iab-privacy-considerations-03](#), July 2012.

[I-D.saintandre-sip-xmpp-chat]

Saint-Andre, P., Gavita, E., Hossain, N. and S. Loreto, "Interworking between the Session Initiation Protocol (SIP) and the Extensible Messaging and Presence Protocol (XMPP): One-to-One Text Chat", Internet-Draft [draft-saintandre-sip-xmpp-chat-04](#), October 2012.

[IANA-lang-tags]

"IANA Language Subtag Registry", , <www.iana.org/assignments/language-subtag-registry>.

Author's Address

Randall Gellens
Qualcomm Technologies, Inc.
5775 Morehouse Drive
San Diego, CA 92121
US

Email: rg+ietf@qti.qualcomm.com

