



# مهندسی برق، مکانیک، کامپیوتر و فناوری اطلاعات

دانشگاه شیراز

۱۲ مهر ماه ۱۳۹۷

October 4, 2018

www.emce.ir

ثبت شده در پایگاه استنادی علوم جهان اسلام (ISC)



## مترجم زبان فارسی به زبان اشاره فارسی برای ناشنوایان

مرضیه دیوانی<sup>۱</sup>، محبوبه شمس<sup>۲</sup>، عبدالرضا رسولی کناری<sup>۳</sup>، محمد لطافت ابریکوه<sup>۴</sup>

<sup>۱</sup>دانشگاه صنعتی قم، divani.m@qut.ac.ir

<sup>۲</sup>دانشگاه صنعتی قم، shamsi@qut.ac.ir

<sup>۳</sup>دانشگاه صنعتی قم، rasouli@qut.ac.ir

<sup>۴</sup>دانشگاه صنعتی قم، letafat.m@qut.ac.ir

### چکیده

عامل مهم در ارتباط میان انسان ها، توانمندی در برقراری ارتباط از روش های مختلف می باشد. استقلال و سلامت افراد ناشنوا تا حد زیادی وابسته به ارتباطات است. پشتیبانی و حمایت نشدن از جانب جامعه شنوا، افراد ناشنوا را به سمت منزوی شدن می کشاند. بنابراین، در نظر گرفتن نیازهای ناشنوایان در تکنولوژی های اطلاعاتی و ارتباطی جدید، امر بسیار مهمی است. به منظور غلبه کردن بر موانع و فاصله هایی که میان افراد شنوا و افراد دچار آسیب شنوایی وجود دارد، سیستم مترجمی را برای ترجمه ی زبان فارسی به زبان اشاره ی فارسی پیشنهاد می دهیم. این سیستم، متن به زبان فارسی را دریافت کرده و طی فرآیندی آن را به زبان اشاره ی فارسی تبدیل می نماید. برای انجام این مهم با چالش های متعددی چون کمبود حرکات اشاره به نسبت تعداد کلمات موجود در زبان فارسی، مشکلات غیرساخت یافته بودن زبان فارسی و همچنین مشکلات حافظه در نمایش خروجی مواجه هستیم؛ بدین منظور معماری سیستم های مشابه قابل استفاده نمی باشد. در این مقاله، معماری جدیدی پیشنهاد شده است که در آن از پردازش های زبان فارسی، استفاده از ساختار درختی برای تبدیل متن دریافتی به زبان اشاره ی فارسی و در نهایت استفاده از انیمیشن برای نمایش زبان اشاره استفاده شده است. معماری پیشنهادی کارایی، سرعت و حافظه ی ذخیره سازی را بهبود می بخشد.

**واژگان کلیدی:** ناشنوایان، زبان اشاره فارسی، پردازش زبان طبیعی، سیستم مترجم.

### مقدمه

جریان اصلی جامعه به طور مرسوم جامعه کم شنوایی و ناشنوایی را کنار گذاشته است. ناتوانی افراد کم شنوا در برقراری ارتباط صمیمانه، همچنین نبود آموزش مرسوم به اندازه کافی با توجه به نیازهای جامعه در این کنارگذاری تأثیر به سزایی دارند. بنابراین آن ها خود را در موقعیت تقریباً نامساعدی در تمامی حوزه های زندگی یافته اند. مطالعات نشان می دهد که ۹۰٪ افراد کم شنوا و ناشنوا در جهان هرگز به مدرسه نرفته اند، بنابراین از سواد کمتری برخوردار هستند. از این رو آن ها نیاز به توسعه ابزاری دارند که نه تنها به آن ها در برقراری ارتباط با جریان اصلی جامعه کمک کند، بلکه آن ها را قادر به یادگیری سازد. زبان مهم ترین وسیله ی ارتباطی در جوامع اجتماعی است و شاید بتوان گفت بدون زبان یا وسیله ای که به رد و بدل کردن پیام کمک کند جوامع انسانی شکل نمی گرفت. در واقع برای در کنار هم ماندن نیاز است تا ارتباط برقرار شود و خواسته ها، احساسات، ترس ها و شاید آرزوها به دیگری منتقل شود. پس با این نگاه می توان ارتباطی ناگسستنی بین زبان (به معنای وسیله ی برقراری ارتباط) و شکل گیری جامعه قائل شد (خالقی طرقي، ۱۳۸۹). زبان عبارت است از تمام

نمادها، علائم و وسائلی از هر نوع، که به وسیله‌ی حواس مختلف، بین افراد جامعه رابطه برقرار می‌کند. با توجه به این تعریف غیر از زبان گفتاری، زبان اشاره و... نیز برای تفهیم و ایجاد رابطه بین افراد جامعه وجود دارد (روح‌الامینی، ۱۳۸۲).

## زبان اشاره

بسیاری از افراد کم‌شنوا در سراسر جهان از زبان اشاره استفاده می‌کنند. زبان اشاره یک زبان دیداری است که از شکل‌های دست، بیان چهره‌ای ژست‌ها و زبان بدنی جهت برقراری ارتباط استفاده می‌کند. زبان اشاره به خودی خود زبان کاملی است که واژه‌های خاص خود، دستور زبان و ساختاری متفاوت با زبان گفتاری دارد. این زبان در هر کشوری متفاوت است و لهجه‌ی خاص خود را نیز دارد (موللی، ۱۳۹۰). گاهی اشارات به حدی روشن و واضح هستند که هر بیننده‌ای به سهولت آن‌ها را درمی‌یابد؛ مانند اشارات ویژه‌ی خوردن، آشامیدن، خوابیدن (بهداری و همکاران، ۱۳۷۵). یا برای مثال در زیمبابوه، اشاره‌ی نان شبیه بریدن نان است. موز شبیه پوست کندن موز است، اشاره‌ی مادر شبیه شیر دادن، اشاره‌ی پدر شبیه ریش است و اشارات دیگر شبیه این اشاره که اغلب حس آن کلمه را به مخاطب منتقل می‌کند. بنابراین در یک جامعه، افراد اشارات را بر حسب چگونگی کاربرد آن کلمه در آن جامعه به کار می‌برند. البته کودکان و خانواده‌ها، علائم را بر حسب نیازهایشان گسترش می‌دهند (موللی، ۱۳۸۹).

## زبان اشاره‌ی فارسی

زبان اشاره‌ی فارسی زبان اول ناشنوایان ایران و همچون دیگر زبان‌های اشاره‌ی دنیا زبانی طبیعی است. زبان اشاره‌ی فارسی اگرچه تحت تأثیر زبان فارسی است؛ اما مستقل از آن است و هم از نظر واژگان و هم از نظر دستور زبان با زبان فارسی تفاوت دارد. همچنین این زبان، از نظر ماهیتی، برخلاف سایر زبان‌های ایرانی که ماهیتی آوایی - شنیداری دارند، دارای ماهیتی بصری - تصویری است.

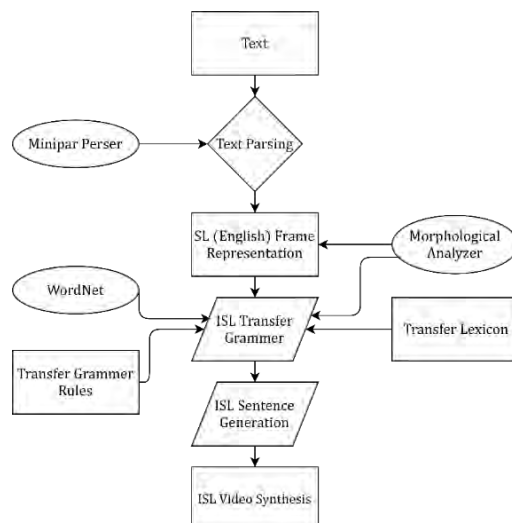
## پیشینه پژوهش

داسگوپتا<sup>۱</sup> و باسو<sup>۲</sup>، در سال ۲۰۰۸، به منظور تبدیل اتوماتیک متن انگلیسی به زبان اشاره‌ی هندی، چارچوبی ارائه کردند که این روش، از قوانین گرامری انتقال برای زبان اشاره‌ی هندی به منظور تولید جمله بهره می‌برد. در این سیستم ابتدا جمله‌ی ورودی، با استفاده از پارسر وابستگی<sup>۳</sup>، پیش‌پردازش و تجزیه و تحلیل می‌شود تا به وسیله‌ی آن اطلاعات نحوی و عملکردی بیشتری از جمله شناسایی شود. این اطلاعات که به صورت یک ساختار میانی مستقل از زبان ارائه می‌شوند، به عنوان ورودی ماژول تولید قرار می‌گیرند. دو عملیات اصلی در ماژول تولید، انجام می‌شود که عبارتند از: الف) انتخاب واژگان و ب) تطبیق ترتیب کلمات. در مرحله‌ی انتخاب واژگان، توسط یک لغت-نامه‌ی دوزبانه، کلمات معادلی که مناسب زبان اشاره‌ی هندی هستند، انتخاب و جایگزین کلمات جمله می‌شوند. آن‌گاه در مرحله‌ی تطبیق ترتیب کلمات، انتقال ساختار جمله‌ی ورودی به زبان اشاره‌ی هندی انجام می‌گیرد. در نهایت، خروجی مرحله‌ی قبل در قالب ویدئوی از پیش ضبط شده، نمایش داده می‌شود (Dasgupta and Basu, 2008). معماری این سیستم در شکل ۱۰ نمایش داده شده است.

<sup>۱</sup> Dasgupta

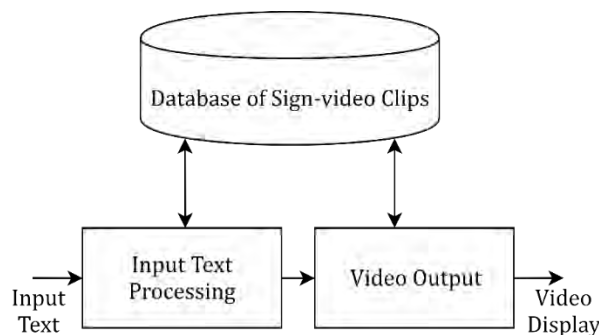
<sup>۲</sup> Basu

<sup>۳</sup> Dependency Parser



شکل ۱۰ معماری سیستم مترجم متن انگلیسی به زبان اشاری هندی (Dasgupta and Basu, 2008)

به منظور ارزیابی این سیستم از کارشناسان زبان‌شناسی برای درک مفهوم خروجی استفاده شده است (Dasgupta and Basu, 2008). سرکار<sup>۴</sup> و همکارانش، در سال ۲۰۰۹، به منظور کاهش موانع ارتباطی میان افراد کم‌شنوا و افراد عادی جامعه، نرم افزار مترجمی را توسعه دادند که متن ورودی را به زبان بنگال دریافت کرده و به طور خودکار به زبان اشاری بنگالی (BaSL)<sup>۵</sup> که توسط انجمن بنگالی در سراسر بنگال غربی ایجاد شده است، تبدیل می‌کند. فرهنگ لغت، که مبنای کار این مترجم است، در حدود هزار کلمه را شامل می‌شود که این کلمات، عمدتاً ساده هستند و از کتاب‌های درسی کودکان جمع‌آوری شده‌اند. روش کار این مترجم این گونه است که کلمه یا جمله‌ی ساده در بخش مربوطه وارد شده و پردازش‌ها و جابجایی‌های لازم در ساختار کلمه یا جمله صورت می‌گیرد و در نهایت اشاری معادل آن کلمه یا جمله نمایش داده می‌شود (Sarkar et al, 2009). معماری این مترجم در شکل ۲ آورده شده است.



شکل ۲ معماری سیستم مترجم متن به زبان اشاری بنگالادی (Sarkar et al, 2009)

معماری این مترجم شامل سه ماژول مختلف است که عبارتند از:

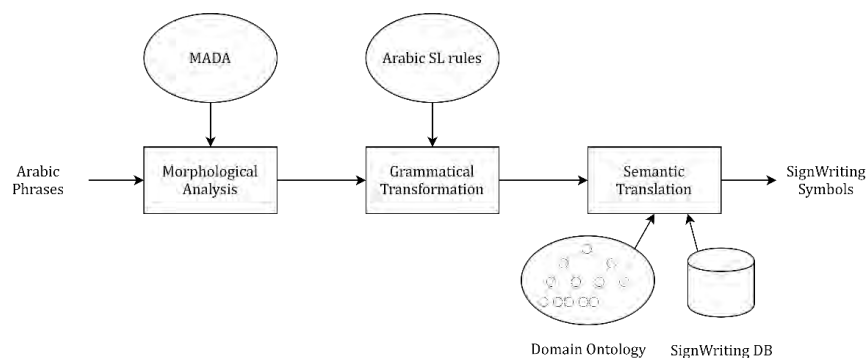
- ۱- ماژول پردازش متن ورودی
- ۲- پایگاه‌داده‌ای حاوی کلیپ‌های ویدئویی زبان اشاره
- ۳- ماژول خروجی ویدئویی

<sup>۴</sup> Sarkar

<sup>۵</sup> Bangla Sign Language

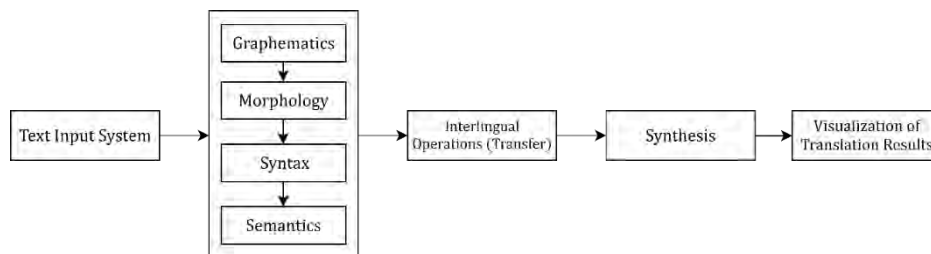
این سیستم برای نمایش خروجی برنامه (زبان اشاره‌ی بنگال) از ویدئوهای از قبل تهیه شده استفاده می‌کند. این مترجم به دنبال استانداردسازی زبان اشاره‌ی بنگالی در تمام مناطق بنگلادش می‌باشد (Sarkar et al, 2009).

المسعود<sup>۶</sup> و الخلیفه<sup>۷</sup>، در سال ۲۰۱۱، سیستمی را به منظور ترجمه‌ی متن عربی به زبان اشاره‌ی عربی در حوزه‌ی فقه نماز ارائه کردند. سیستم پیشنهادی آن‌ها، با اعمال قوانین ترجمه‌ی زبان اشاره‌ی عربی و همچنین با استفاده از آنتولوژی حوزه‌ی در نظر گرفته شده، متن عربی را به زبان اشاره‌ی عربی ترجمه می‌کند. این سیستم دارای مجموعه‌ای از پردازش‌هاست که شامل تجزیه و تحلیل مورفولوژیک (تجزیه و تحلیل تکواژشناسی)، تحول دستوری و ترجمه‌ی معنایی می‌شود. این سیستم، در ابتدا متن عربی را به عنوان ورودی دریافت کرده و هر جمله را به بخش تجزیه و تحلیل مورفولوژیک و ابهام‌زدایی، ارسال می‌کند. آن‌گاه ماژول تحول دستوری، نتایج حاصل از قسمت قبل را به عنوان ورودی دریافت می‌کند و قوانین زبان اشاره‌ی عربی را بر روی هر یک از کلمات با توجه به این‌که دارای چه نقشی هستند، اعمال می‌کند. ماژول ترجمه‌ی معنایی، نتیجه‌ی پردازش پیشین را دریافت می‌کند و به جستجوی هر یک از کلمات در درون آنتولوژی حوزه‌ی مرتبط، می‌پردازد تا بتواند کد مربوط به اشاره‌ی آن کلمه را بیابد. اگر یک کلمه، دارای اشاره‌ی نظیر با خود نبود، آن‌گاه این کلمه با یکی از مترادف‌هایش که دارای یک اشاره در درون پایگاه داده‌ی Signwriting است، جایگزین می‌شود. سپس، هر کد اشاره، با نماد اشاره‌ی متناظرش که در درون پایگاه داده‌ی نوشتار اشاره ذخیره شده است، تعویض و جایگزین می‌شود. اگر کلمه‌ای دارای اشاره‌ی نظیر با خود در درون آنتولوژی حوزه‌ی در نظر گرفته شده نبود، آن کلمه به صورت هجی‌انگشتی، به نمایش درخواهد آمد (Almasoud and Khalifa, 2009). معماری این سیستم در شکل ۳ نمایش داده شده است.



شکل ۳ معماری سیستم مترجم زبان عربی به زبان اشاره‌ی عربی (Almasoud and Khalifa, 2009)

گریف<sup>۸</sup> و همکارانش، در سال ۲۰۱۳، سیستم تبدیل و ترجمه‌ی زبان روسی به زبان اشاره‌ی روسی (RSA)<sup>۹</sup> را برای افرادی که دارای اختلالات شدید شنوایی هستند، معرفی نمود. در آن زمان گام‌های اولیه‌ی این سیستم مورد بررسی قرار گرفت. ساختار کلی سیستم مورد نظر آن‌ها که ترجمه‌ی کامپیوتری زبان روسی به زبان اشاره‌ی روسی را انجام می‌داد، همانند شکل ۴ تعریف شده بود:



شکل ۴ معماری سیستم تبدیل زبان روسی به زبان اشاره‌ی روسی (Grif et al, 2012)

<sup>۶</sup> Almasoud

<sup>۷</sup> Al-Khalifa

<sup>۸</sup> Grif

<sup>۹</sup> Russian Sign language

## ماژول‌های اصلی این ساختار عبارتند از:

- تفسیر متن ورودی
- اعمال تغییرات بین زبانی به منظور دستیابی به زبان مقصد
- ترکیب زبان اشاره
- مشاهده‌ی نتیجه‌ی ترجمه، به زبان اشاره و به صورت ویدئو

در مرحله‌ی تفسیر متن ورودی، مواردی چون نویسه‌شناسی، تجزیه و تحلیل مورفولوژیک، تحلیل نحوی و نیز تجزیه و تحلیل معنایی همان‌طور که در شکل آمده است، مورد بررسی قرار می‌گیرد.

در مرحله‌ی نویسه‌شناسی، عناصر متن ورودی از هم جداسازی می‌شود. در مرحله‌ی آنالیز مورفولوژیک، از روش‌های مبتنی بر دیکشنری گرامری زبان روسی که حاوی کلمات کلیدی این زبان است بهره می‌برد.

در تحلیل نحوی، ارتباط بین کلمات تعریف و بررسی می‌شود که در این تحلیل از تحلیل معنایی نیز استفاده می‌شود. روش اصلی تجزیه و تحلیل نحوی و معنایی آن‌ها، درخت وابستگی متن ورودی می‌باشد. در ماژول بعدی با استفاده از درخت وابستگی مرحله‌ی قبل، متن ورودی به زبان اشاره‌ی روسی تبدیل می‌شود.

در ماژول بعدی، ساختار جمله در زبان اشاره‌ی روسی مورد بررسی قرار می‌گیرد. در نهایت توسط ماژول تصویرساز، به صورت زبان اشاره‌ی روسی نمایش داده می‌شود (Grif et al, 2012).

در جدول ۱ خلاصه‌ای که سیستم‌های مترجم مشابه را مشاهده می‌کنید:

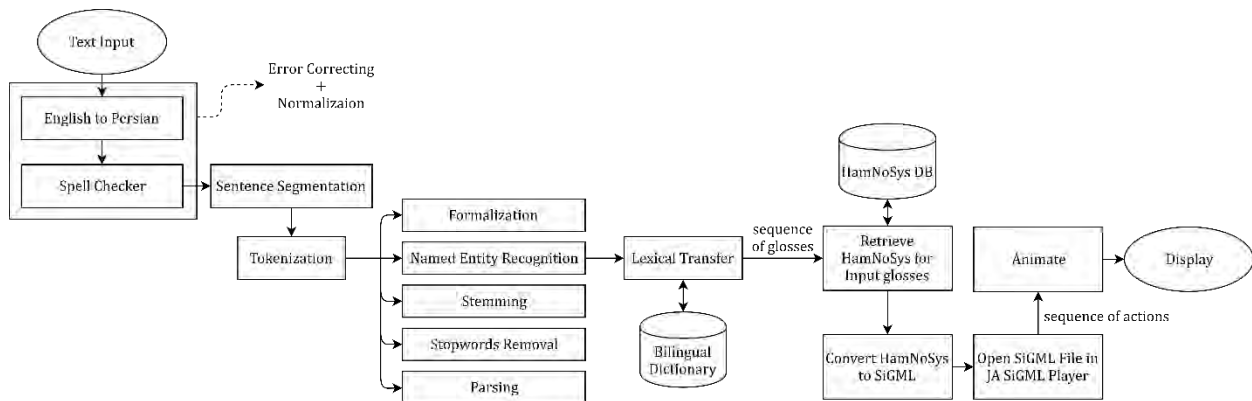
جدول ۱ جمع‌بندی سیستم‌های پیشین مترجم متن به زبان اشاره

ردیف	نام پروژه	سال	زبان مبدأ	زبان مقصد	ورودی	نوع نمایش
۱	An English to Indian Sign Language Machine Translation System	۲۰۰۸	زبان انگلیسی	زبان اشاره هندی	متن	ویدئو
۲	A Translator for Bangla Text to Sign Language	۲۰۰۹	زبان بنگال	زبان اشاره بنگالی	متن	ویدئو
۳	A Proposed Semantic Machine Translation	۲۰۱۱	زبان عربی	زبان اشاره عربی	متن	Signwriting
۴	Computer Sign Language Translation System for Hearing Impaired Users / Development of Computer Sign Language Translation Technology for Deaf People	۲۰۱۳	زبان روسی	زبان اشاره روسی	متن	ویدئو

## معماری پیشنهادی

رویکردهای مختلفی برای ترجمه‌ی ماشینی وجود دارد که انتخاب روش مناسب برای ترجمه، بر اساس نوع زبان مبدأ و مقصد، متفاوت است. در معماری پیشنهادی ما، رویکرد مبتنی بر انتقال که زیر مجموعه‌ای از ترجمه‌ی مبتنی بر قاعده است، پیش گرفته می‌شود. در این رویکرد بیش‌تر از ترجمه‌ی لغت به لغت به صورت مستقیم و با استفاده از یک دیکشنری دوزبانه استفاده می‌شود؛ لازم به ذکر است در این

روش، امکان برخی تغییرات کوچک یا برخی جابه‌جایی‌ها وجود دارد. این روش ترجمه، برای ترجمه‌ی زبان‌هایی مناسب است که بسیار شبیه به یکدیگر هستند و تفاوت عمده‌ی آن‌ها در لغات مورد استفاده‌ی آن‌ها است.



شکل ۴ معماری پیشنهادی سیستم مترجم مبتنی بر آواتار زبان فارسی به زبان اشاره‌ی فارسی

در شکل ۴، معماری پیشنهادی سیستم مترجم مبتنی بر آواتار زبان فارسی به زبان اشاره‌ی فارسی نمایش داده شده است. به منظور تحقق یافتن هدف موردنظر ما، یعنی ترجمه‌ی زبان فارسی به زبان اشاره‌ی فارسی، تعدادی از ماژول‌ها می‌بایست در کنار یکدیگر قرار گیرند و مجموعه‌ای از پردازش‌ها را به صورت مرحله به مرحله به انجام رسانند.

## پیش پردازش زبان فارسی

در روند هرگونه پردازش روی متن‌های زبان طبیعی، انجام یک‌سری عملیات پیش‌پردازش، امری اجتناب‌ناپذیر است. دقت این پیش‌پردازش‌ها، تأثیر به‌سزایی در نتایج اعمال الگوریتم‌ها در فازهای بعدی دارد. هرچقدر که دقت پیش‌پردازش بیشتر باشد، الگوریتم‌ها به نتایج واقعی خود نزدیک‌تر خواهند شد. از این‌رو، در اولین گام می‌بایست متون برای استفاده در گام‌های بعدی به شکلی استاندارد درآیند، که اصطلاحاً به آن نرمالسازی گفته می‌شود. در ابتدا بایستی همه‌ی کاراکترهای متن با جایگزینی با معادل استاندارد آن، یکسان‌سازی گردند. در پردازش رسم‌الخط زبان فارسی، با توجه به قرابتی که با رسم‌الخط عربی دارد، همواره در تعدادی از حروف مشکل وجود دارد که از جمله‌ی آن‌ها می‌توان به حروف «ک»، «ی»، همزه اشاره نمود. در اولین گام باید مشکلات مربوط به این حروف را برطرف ساخت. علاوه بر این، اصلاح و یکسان‌سازی کاراکتر نیم‌فاصله و فاصله در کاربردهای مختلف آن و همچنین حذف نویسه‌ی «\_» که برای کشش کاراکترهای چسبان مورد استفاده قرار می‌گیرد و مواردی مشابه برای یکسان‌سازی متون، از اقدامات لازم قبل از شروع فازهای بعدی می‌باشد. در این فاز مطابق با یک‌سری قاعده‌ی دقیق و مشخص، فاصله‌ها و نیم‌فاصله‌های موجود در متن برای علاماتی نظیر «ها» و «ی» غیر چسبان در انتهای لغات و همچنین پیشوندها و پسوندهای فعل‌ساز نظیر «می»، «ام»، «ایم» و موارد مشابه جهت استفاده در فازهای بعدی، اصلاح می‌گردند (نوری، ۱۳۹۲).

## جداکننده‌ی جملات (Sentence Splitter):

پس از پایان مرحله‌ی نرمال‌سازی متون، ابزار تشخیص‌دهنده‌ی جملات با استفاده از علامت‌های «.»، «؟»، «!»، «» و «>» و به کارگیری برخی دستورات گرامری زبان فارسی و در نظر گرفتن برخی لغات آغازکننده‌ی جملات، مرز جمله‌ها را برای استفاده در گام‌های بعدی تعیین می‌نماید. به بیان دیگر، به کمک این پردازش‌گر، می‌توان جملات متن را استخراج کرد. این ابزار بایستی تشخیص صحیح و دقیق جملات را در متن ورودی داشته باشد. برای ایجاد این ابزار، باید ابتدا تمامی کاراکترها، نمادها و احیاناً قواعد دستوری که باعث شکسته شدن جملات می‌شوند، شناسایی گردند. با توجه به پایه بودن جمله در بسیاری از پردازش‌های زبانی، خروجی دقیق این ابزار از

درجه‌ی اهمیت بالایی برخوردار است. شبه‌کد پیاده‌سازی این ماژول را در **Error! Reference source not found.** مشاهده می‌کنید:

```

Algorithm Sentence_Splitter
1      Start
2      Input (Normalized_Text)
3      Output (Tokenize_Sentence)
4      Begin
5          char [] delimiterChars = { '!', ':', ';', '!', '?', '€' }
6          string [] Tokenize_Sentence = Normalized_Text.Split (delimiterChars,
            StringSplitOptions.RemoveEmptyEntries)
7          return Tokenize_Sentence []
8      End

```

شکل ۵ شبه کد جداکننده‌ی جملات متن

### جداکننده‌ی کلمات (Tokenize):

پس از جداسازی جملات متن، نوبت به مازول جداکننده‌ی کلمات می‌رسد. این مازول، با استفاده از علامت‌های «فاصله»، «د»، «ر»، «»، «>»، «+»، «<»، «؟»، «/»، «@»، «\$»، «/»»، «^»، «&»، «\*»، «-»، «<»، «>»، «:»، «{»، «}»، «\t»، «\r»، «\n»، واژه‌ها را شناسایی می‌نماید. به عبارت دیگر به کمک آن، می‌توان کلمات متن را استخراج نمود.

شبه‌کد پیاده‌سازی این مازول در سیستم مترجم را در شکل ۶ مشاهده می‌کنید.

```

Algorithm Tokenize
1      Start
2      Input (Normalized_Text)
3      Output (Tokenize_Text)
4      Begin
5          char[] delimiterChars = { ' ', ',', '!', ':', ';', '+', '=', '?', '(', ')', '@', '$', '%', '^', '&', '*', '-', '(', ')', '!', '{', '}', '\t', '\r', '\n' }
6          string [] Tokenize_Text = Normalized_Text.Split (delimiterChars,
          StringSplitOptions.RemoveEmptyEntries)
7          return Tokenize_Text []
8      End

```

شکل ۶ شبه کد جداکننده‌ی کلمات جمله

## تبدیل واژگانی (Lexical Transfer)

همان‌طور که می‌دانیم زبان اشاره‌ی فارسی و زبان فارسی، معادل یکدیگر نیستند؛ به عبارت دیگر ممکن است یک کلمه از زبان فارسی دارای هیچ اشاره‌ای در زبان اشاره‌ی فارسی نباشد یا یک اشاره در زبان اشاره‌ی فارسی، معادل چندین کلمه در زبان فارسی باشد. حال با توجه به این مسئله، معادل اشاره‌ی هر یک از کلمات موجود در جمله را با استفاده از یک دیکشنری دوزبانه‌ی زبان فارسی به زبان اشاره‌ی فارسی، در جمله قرار می‌دهیم. سیستم نشان‌گذاری هامبورگ، سیستمی است که جهت به نگارش درآوردن چگونگی انجام اشارات در زبان اشاره، مورد استفاده قرار می‌گیرد. می‌توان گفت این سیستم، مشابه الفبای فونتیک است که برای زبان‌های گفتاری مورد استفاده قرار می‌گیرد. سیستم نشان‌گذاری HamNoSys، می‌تواند تمامی اشارات موجود در همه‌ی زبان‌های اشاره را توصیف نماید. به علاوه، این سیستم

وابسته به قراردادهای زبان اشاره نیست که در کشورهای مختلف، انواع مختلفی داشته باشد؛ از این رو، می‌تواند به عنوان سیستم کاملی در سطح بین‌المللی مورد استفاده قرار گیرد (Kaur and Kumar, 2016).

## تبدیل HamNoSys به زبان SiGML

در سیستم مترجم زبان فارسی به زبان اشاره فارسی، برای به نمایش در آوردن متن ورودی به زبان اشاره توسط شخصیت انیمیشنی، از تبدیل نمادهای HamNoSys اشارات به زبان SiGML استفاده می‌کنیم.

### زبان SiGML<sup>۱۰</sup>

زبان SiGML، یک زبان نشان‌گذاری برای حرکات اشاره است. این زبان، هر از نمادهای HamNoSys را به صورت تگ‌هایی از XML توصیف می‌کند. زبان SiGML، از نشان‌گذاری HamNoSys مربوط به اشارات زبان اشاره ساخته می‌شود که قابلیت اجرا شدن توسط شخصیت سه‌بعدی در برنامه‌ی JA SiGML را دارد (Kaur and Kumar, 2016).

### نحوه‌ی استفاده از سیستم نشان‌گذاری HamNoSys و زبان SiGML

به هنگام درج یک کلمه درون سیستم مترجم، به صورت همزمان، نحوه‌ی انجام اشاره‌ی آن کلمه نیز با استفاده از نشان‌گذاری HamNoSys ذخیره می‌شود تا به هنگام ترجمه، اشارات مربوط به هر کلمه قابل دستیابی باشند. نمادهای نشان‌گذاری HamNoSys مربوط به هر کلمه، به صورت رشته‌ای از کدهای اسکی هر نماد به همراه کلمه نگه‌داری و ذخیره می‌شود. سپس آن را به زبان SiGML تبدیل مینماید. در شکل ۱۳ شبه کد تبدیل نمادهای HamNoSys کلمه‌ی ورودی به زبان SiGML را مشاهده می‌کنید در نهایت کدهای SiGML بدست آمده توسط خروجی گرافیکی مورد نظر که یک آواتار طراحی شده میباید نمایش داده می‌شود.

### بحث و نتیجه‌گیری

برای استفاده از یک سیستم باید از کارایی آن سیستم اطمینان حاصل شود. برای ارزیابی کارایی یک سیستم با توجه به معیارها و ملاک‌های سیستم مورد نظر و منطقه پیاده‌سازی سیستم روش‌های مختلفی وجود دارد. به طور کلی برای بررسی کارایی یک سیستم میتوان به معیارهای متنوعی از جمله سرعت، دقت، زمان و هزینه راه اندازی سیستم را مورد توجه قرارداد. در سیستم پیشنهادی به معیارهای سرعت و دقت و همچنین هزینه توجه شده است. **معیار سرعت**، به منظور پاسخ‌گویی سیستم مترجم و ترجمه‌ی متن ورودی برای ناشنویان، در کم‌ترین زمان؛ **معیار دقت**، به منظور انتقال حداکثری صحیح مفاهیم به ناشنویان و **معیار هزینه**، به دلیل هزینه‌ی بالای استفاده از مترجم انسانی برای انتقال مفاهیم به ناشنویان مورد توجه قرار گرفته است.

ما در معماری پیشنهادی، برای نگه‌داری و ذخیره‌ی کلمات و نشان‌گذاری HamNoSys مربوط به هر کلمه، از ساختمان داده‌ی درخت پیشوندی بهره می‌بریم. جستجو در درخت پیشوندی، بسیار آسان و سریع است و از مرتبه‌ی طول کلمه می‌باشد. همچنین درخت، زمانی که شامل تعداد زیادی از رشته‌های کوچک باشد، فضای بسیار کمی را برای ذخیره اشغال می‌کند. برای بررسی دقت ترجمه، به ارزیابی ماژول تبدیل در سیستم مترجم متن فارسی به زبان اشاره فارسی پرداختیم. این ارزیابی نشان داد استفاده از ماژول تبدیل کلمات به معادل خود در دنیای ناشنویان کارایی سیستم را افزایش داده است. در نهایت، ما برای نمایش خروجی برنامه از انیمیشن استفاده کرده‌ایم که این کار باعث کاهش سربار حافظه می‌شود، زیرا اکثر سیستم‌های مشابه برای نمایش از ویدئوهای از قبل تهیه شده استفاده می‌کردند که این کار با بالا رفتن تعداد جملات و اشارات، باعث سربار حافظه شده و همچنین هزینه‌ی بالایی برای تهیه‌ی فضای ذخیره‌سازی،

<sup>۱۰</sup> Sign Gesture Markup Language



ترکیب ویدئوهای مربوط به کلمات درون متن ورودی و نیز ویرایش اطلاعات درون ویدئو صرف می‌شود. به منظور ویرایش ویدئو، پس از ضبط آن (برای اصلاح یا به‌روزرسانی آن) نیازمند متخصصی در زمینه‌ی ویرایش ویدئو، تقلیدی از حالت اصلی، و حضور همان اشاره‌کننده هستیم. که البته آسان‌ترین راه، همان ضبط کردن مجدد ویدئو است. اما اگر برای به تصویر در آوردن زبان اشاره، از آواتار یا یک شخصیت انیمیشنی استفاده کنیم، این رویکرد، با مشکلات ذکر شده در قسمت قبل مواجه نخواهد بود. با ترکیب سیستم به همراه آواتار یا شخصیت انیمیشنی، محتوا به صورت متن دریافت می‌شود؛ و به صورت بلادرنگ به فرم اشاره تبدیل می‌گردد، بدون آن‌که نیازی به در نظر گرفتن فضایی جهت ذخیره‌سازی ویدئو داشته باشد.

## منابع:

- بهادری، ایران، پیروزی، مرتضی، تهرانی زاده، حبیب، قاسمی شاد، الهه، لوح موسوی، محسن، محمودی، رضا، ۱۳۷۵، مجموعه اشارات ناشنویان، تهران، سازمان بهزیستی کشور
- خالقی طرقي، گلناز، ۱۳۸۹، ناشنویان و زبان اشاره، محدودیت‌ها و امکانات، پایان‌نامه کارشناسی ارشد، دانشگاه تهران
- روح‌الامینی، محمود، ۱۳۹۶، چاپ هفدهم، گرد شهر با چراغ، تهران، انتشارات عطار
- موللی، گیتا (مترجم)، برقراری ارتباط، جزوه راهنمای کارشناسان کودکان دچار اختلال ارتباطی، ۱۳۸۹، تهران، انتشارات مؤسسه فرهنگی هنری فرجام جام جم
- موللی، گیتا، ۱۳۹۰، کودک من کم شنواست، چاپ اول، تهران، انتشارات دانشگاه علوم بهزیستی و توانبخشی
- نوری، نجمه، ایجاد یک برچسب زن معنایی با استفاده از فریم نت برای متون فارسی، ۱۳۹۲، پایان‌نامه کارشناسی ارشد، دانشگاه فردوسی مشهد
- همتائی، حسن، مطیع نصرآبادی، علی، هاشمی گلپایگانی، سید محمدرضا، بازشناسی تصویری الفبای دستی فارسی ناشنویان با استفاده از شبکه عصبی و ماشین بردار پشتیبان، پنجمین کنفرانس ماشین بینایی و پردازش تصویر، ۱۳۸۷
- Almasoud, Ameera M. Al-Khalifa, Hend S. (2011). A proposed semantic machine translation system for translating Arabic text to Arabic sign language. Proceedings of the Second Kuwait Conference on e-Services and e-Systems (2011). 79-84
- Dasgupta, Tirthankar. And Basu, Anupam. (2008). An English to Indian Sign Language Machine Translation System
- Grif, Mikhail G. Korolkova, Olga O. Demyanenko, Yelena A. (2012). Computer Sign Language Translation System for Hearing Impaired Users. Strategic Technology (IFOST), 2012 7th International Forum on. 1-4
- Kaur, Khushdeep. Kumar, Parteek. (2016). HamNoSys to SiGML conversion system for sign language automation. Procedia Computer Science. Vol. 89, 794-803.
- Sarkar, Biswajit. Datta, Kaushik. Datta, CD. Sarkar, Debranjana. Dutta, Shashanka J. Roy, Indranil Das. Paul, Amalesh. Molla, Joshim Uddin. Paul, Anirban. (2009). A Translator for Bangla Text to Sign Language. Annual IEEE India Conference (2009). 1-4