ᲡᲐᲥᲐᲠᲗᲕᲔᲚᲝᲡ ᲛᲔᲪᲜᲘᲔᲠᲔᲑᲐᲗᲐ ᲔᲠᲝᲕᲜᲣᲚᲘ ᲐᲙᲐᲓᲔᲛᲘᲘᲡ ᲛᲝᲐᲛᲑᲔ, ტ. 12, №3, 2018
BULLETIN OF THE GEORGIAN NATIONAL ACADEMY OF SCIENCES, vol. 12, no.3, 2018

Linguistics

Theory of Neutral Signs (TNS)

Tamar Makharoblidze

Ilia State University, Tbilisi, Georgia

(Presented by Academy Member Thomas Gamkrelidze)

ABSTRACT. Research on technological solutions recognition for Sign Language (SL) already has its history. As for the last of two to three dozens of years, a number of virtual studios worldwide carried out various experiments. Despite such an intense interest in the issue of a sign-spoken language translator (SSLT), such a device has not yet been developed. The sign classification and the new theory-Theory of Neutral Signs (TNS) for the abovementioned task, taking into consideration the latest trends, new technologies, algorithms and approaches, are presented in the paper. The biggest problem for elaborating a good engine of SL machine translation is a lack of sign separators or spaces. In sign language (SL) texts, it is hard to understand where the beginning or ending of a proper sign is. Studying the nature of neutral signs (NS) will help us effectively perform segmentations of phrases in chunks. It will allow us to set up a "silence" threshold and detect sign activity, like speech to text processing, in order to recognize and describe the meaningful signs. Crucially, NS becomes a part of a neuro-net, and knowing its structure allows us to make segmentation more accurate, more precisely defining necessary information. *© 2018 Bull. Georg. Natl. Acad. Sci.*

Keywords: Sign languages, sign recognizing, GESL, computer linguistics

Research on technological solutions for SL recognition already has its history, as during the last two to three dozen years, a number of virtual studios worldwide carried out various experiments [1-12]. Despite such an intense interest in the issue of a Sign-Spoken language translator (SSLT), such a device has not yet been developed. The present paper discusses the sign classification and the new theory for the above mentioned task, taking into consideration the latest trends, new technologies, algorithms and approaches.

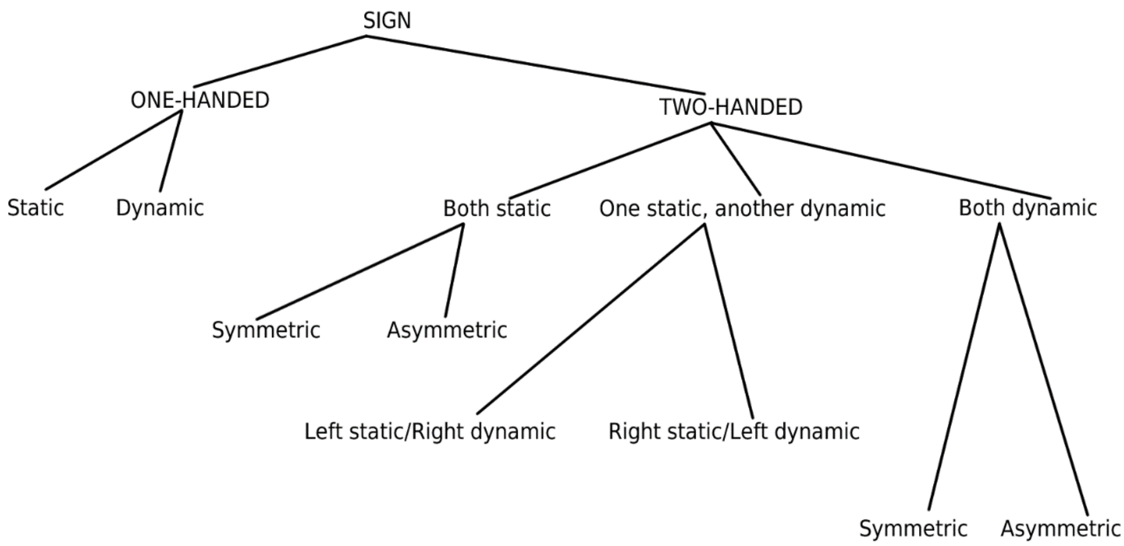
Sign classification. Signs can be static or dynamic, one or two-handed. Two-handed signs may be symmetric or asymmetric. In addition, among two-handed signs, either both hands are producing dynamic or static signs, or one hand produces a static sign while the other one does a dynamic sign. For sign classification we used the combination approach:

Dynamic gradation (with space and time parameters) – The signs are statistic or/and dynamic. Dynamic signs may have one, two or more movement phases;

Composition of a sign / sign structure – the signs may have one, two, three or four (very rarely five) elements or independent signs with (sometimes totally different) meanings. Signs may be as follows: A=a; A=a+b, A=a+b+c, etc.;

For our description, one-handed and two-handed signs can be described in the same way, although there can be a significant difference between the sign producers and their moving/sign producing kinetics.

Classification of signs schematically looks as follows:



The signs may be simple or compound. Compound signs may have two or more (up to five as maximum) meaningful signs in the strict sequence.

MSa+NSab+MSb=MSc

MSb+ NSba+MSa ≠ MSc

For example, in GESL, the sign for “agricultural” is the combination of three MS: “village”, “variety” and “function”.

Fig. 1. The sign for “agricultural” in GESL

The types of signs in signing process, and Theory of Neutral Signs (TNS). To elaborate SSLT from SL into spoken languages is more difficult compared with the reverse version — translation from spoken language into SL. Usually, SL texts are performed smoothly and there are no spaces between meaningful signs (MS). The biggest problem for elaborating a good engine of SL machine translating is a lack of sign separators or spaces. In SL texts, it is hard to understand where is the beginning or ending of a proper sign. To overcome this obstacle, we offer a new theory: The theory of neutral signs (TNS).

There are two types of manual signs:

Signs with meaning – MS (meaningful sign). These are the signs with lexical content (like words) or with morpho-semantic meanings (such as particles or morphemes of different grammar categories), and

Signs without any meaning, which serve as a connection for the manual positions of two neighboring meaningful signs (MSs). It is a neutral sign (NS). NS could also be called a garbage sign. NSs are inter-signs between MSs.

MS can be static or dynamic, one- or two-handed, simple or compound with two or three (and rarely more) signs in a specific sequence. The compound signs can be described as A+B(+C+D)=S.

NS is a dynamic sign between MSs (static or dynamic). Unlike MS, NS is always dynamic. Every MS has three steps of sign production:

The first step is preparation or excursion - MSe;

The second step is a top MSt - the moment of sign exposition; and

The last third step is post-production or recursion (or disposition) - MSr.

The first and third steps are usually mixed with the parts of neighboring signs. At the beginning of the signing process, there is a neutral sign beginner — NSb, and it brings the hand(s) from zero position to MSe. (Zero position is the position hands hanging down and may be slightly bent in the elbows.) NSf is the final neutral sign in the signed text, bringing the hand(s) to zero position from MSr.

In an SL text sequence, the signing dynamics of the two signs is Sa+Sab+Sb. In real signing time, there are the three signs, where Sab is NS between these two MSs (Sa and Sb). This type of NS is a middle or intermediate. It connects two MSs having the mixed characteristics from the ending part of the first (MSr) and the beginning part of the second sign — MSe. Thus, there are three types of NS:

NS connecting (Sab, MSr+MSe);

NSb – the first, beginning sign, and

NSf – last, finishing sign.

In SL, the phrase / sentence “I paint” looks as follows:

MR(I)e+MR(I)t+MR(I)r+MS(paint)e+MS(paint)t+MS(paint)r

MR(I)r+MS(paint)e=NS(I+paint)

However, this description is still incomplete, as MR(I)e and MS(paint)r will be bordering with the other signs in a longer sequence, creating specific NSs, or if this is a separate text, then before MS(I)e, there will be NS(I)b and MS(paint)r followed by NS(paint)f. This SL text will be described as NS(I)b+MR(I)e+MR(I)t+NS(I+paint)+MS(paint)t+MS(paint)r+NS(paint)f.

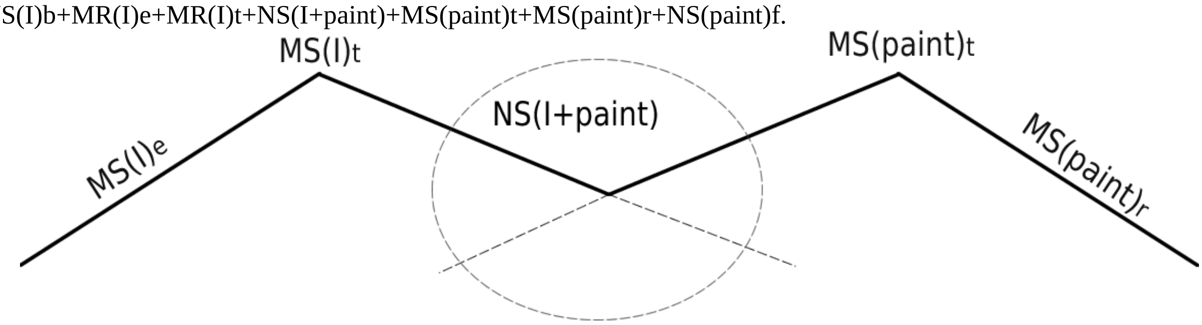


Fig 2. SL text fragment scheme

One minute of GESL fluent text has approximately 55-70 MS with 40-45 seconds duration, and NS duration is approximately 10-17 seconds. According to our experimental calculations, NS is approximately 10-20% of GESL fluent texts. Studying the nature of NS will help us effectively perform segmentations of phrases in chunks. It will allow us to set up a “silence” threshold and detect sign activity, like speech to text processing, in order to recognize and describe MS-s. Crucially, NS becomes a part of a neuro-net, and knowing its structure allows us to make segmentation more accurate, more precisely defining the necessary information.

NS recognition methods. Thus, NS can be considered as a space between MSs, or in other words, NS is a sign separator. The question is how NS can be recognized by the engines. We revealed four methods to identify NS in SL texts:

1. Synergistic method for NS recognition – This method can work by analyzing a big number of SL texts of different SLs, having a big common SL textual base that will be the NS base at the same time. Such a data base can be filled only with common effort using open sources and world-wide collaboration of the area specialists;

2. The approximate parameters of NS can be defined depending on its neighboring signs; NS can be characterized by having less tension of the manual muscles and skin and being without any accompanied mimic; by its non-prosodic element; with freely and lightly curved/hanging, resting fingers; with transitional palm orientation and hand location depending on the proper neighborhood in the sign sequence. NS may look like a MS, or it can be MS in another SL, but the above mentioned general parameters (such as less tension and non-prosody) help to separate any type of NS from MS.

3. Thus, in SL, the number of NS is much bigger than the number of MS. Mathematically, it can be calculated in the proper SL as a number of possible element combinations.

4. The combined identification of NS could be performed with Leap-Motion device and Myo armbands. The pause converged with Leap-Motion’s minimal activity can be considered as a sign-separator in SL process if it is not a static sign. Meanwhile, there is a limited number of static signs in any SL and they can be described in the proper SL corpora, or data-base, or learned by the neuro-nets. The engine can identify static signs and distinguish them from NSs and pausing.

The critical approach to TNS reveals some doubts, such as why should we invest in recognizing NS instead of recognizing MS? In any SL, obviously there are much more NSs than MSs, and MS has more concrete parameters than NS. It will be easier and seemingly more effective to create a base of well-described MS.

The micro-corpora of GESL was created recording the signs from the GESL dictionary [13, 14] with a few deaf persons. This recorded dictionary was oriented to Leap-Motion data. We tried to use the sign-to-word recognizing method, and we noticed that increasing the number of signs reduces the quality of the sign recognizing process. In addition, most signs are not static, but rather dynamic, and the problems were deeper in the case of combined or composed dynamic signs.

Two types of dynamics. The signing process is open, and all elements are displayed via manual and mimic dynamics. This process is dynamic, nonlinear and unpredictable. In the SL process, I distinguish two types of dynamic: A. Extra-formed visual-kinetic dynamic of signing dynamics, and B. Intro-formed linguistic dynamic of signed speech.

These two types of dynamics are overlapped as the first type; the Extra-formed dynamic, as the displayed kinetics of the signing process, serves as the means for the second type of Intro-formed (inner or language structural) dynamic. In other words, in SL, Intro-formed dynamics are displayed by Extra-formed ones. Thus, these two types of dynamics always occur together in SL signing processes.

Conclusions and challenges. The present paper sheds light on the theoretical frames for SL recognizing systems. The main input for scholarly studies can be considered the following:

A new theory of neutral signs TNS, sign classification system, and sign-producing revealed phases.

The signing process is as unpredictable and nonlinear as it is in spoken languages. As was described above, NSe meets NSr producing a unique mixed NS. The sequence itself and the mixed NS are often unpredictable, as it is impossible to know what the signer will say/sign and what style or linguistic variations he/she will use in the signing process. Sign neighborhood is unpredictable. In addition, taking into consideration the individual signing (kinetic) manners depending on the mood and physical conditions of a signer, we can say that NS variations in the sequence are chaotic. Thus, it could be challenging to overview the sign recognizing problem under the light of Chaos Complexity theory.

I would like to express deep grattitude to my Deaf sourses Natia Japoshvili and GESL interpreter Tamar Jikidze, also to computer developer George Mirianashvili.
This work was supported by Shota Rustaveli National Science Foundation (SRNSF), 216702 “The Georgian Sign Language Verbal Morphology”.

ენათმეცნიერება

ნეიტრალური ჟესტის თეორია

თ. მახარობლიძე

ილიას სახელმწიფო უნივერსიტეტი, თბილისი, საქართველო

(წარმოდგენილია აკადემიის წევრის თ. გამყრელიძის მიერ)

ჟესტის ამოცნობის პროგრამული ენჯინის შექმნას და მასთან დაკავშირებულ კვლევებს უკვე აქვს თავისი ისტორია, რამდენადაც ბოლო ორი-სამი ათწლეულის მანძილზე ეს საკითხი საკმაოდ ინტენსიურად შეისწავლებოდა. მიუხედავად საკითხის მიმართ ასეთი დიდი ინტერესისა, ჯერჯერობით არ არსებობს ისეთი კომპიუტერული პროგრამა, რომელიც შეძლებდა ჟესტური ენიდან სამეტყველო ენაზე ტექსტის თარგმნას რეალურ დროში. წარმოდგენილ სტატიაში განხილულია ჟესტის კლასიფიკაცია. უახლესი კვლევებისა და ბოლო მიდგომების გათვალისწინებით, ზემოთ აღნიშნული პრობლემის გადასაჭრელად წამოყენებულია ახალი თეორია – ნეიტრალური ჟესტის თეორია. ჟესტური ენებიდან მანქანური თარჯიმნის შექმნისთვის უდიდეს დაბრკოლებას წარმოადგენს ჟესტთაშორისი სივრცეების, ანუ გამყოფების, არქონა. ჟესტური მეტყველების პროცესში პრაქტიკულად გაურკვეველია, თუ სად იწყება ან სად მთავრდება ესა თუ ის ჟესტი. ნეიტრალური ჟესტის

ბუნების შესწავლა დაგვეხმარება ჟესტური მეტყველების სეგმენტაციის პროცესში და შესაძლებელს გახდის ჟესტის აქტივობის ანალიზს. შესაბამისად, შესაძლებელი გახდება მნიშვნელობის მქონე ჟესტების გამოცალკევება და აღწერა. მნიშვნელოვანია იმის აღნიშვნა, რომ ნეიტრალური ჟესტი ასევე ხდება ე. წ. ნეირო-ნეტის ნაწილი და ამ სტრუქტურის ცოდნა კი ჟესტური ინფორმაციის აღწერის შესაძლებლობას იძლევა.

REFERENCES

1. Starner Th., Weaver J. and Pentland A. (1998) Real-time American Sign Language recognition using desk and wearable computer based video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1371-1375.
2. Gao Wen and Chunli Wang. (2002) Sign language recognition. *Series in Machine Perception and Artificial Intelligence*, 48: 91-120.
3. Yin, Ying (2014) Real-time continuous gesture recognition for natural multimodal interaction. Theses of PHD dissertation. Massachusetts Institute of Technology.
4. Murata, Tomoya, and Jungpil Shin (2014) Hand gesture and character recognition based on kinect sensor. *International Journal of Distributed Sensor Networks*, **1**: pp.1-6
5. Marin Giulio, Fabio Dominio, and Pietro Zanuttigh. (2015) Hand gesture recognition with jointly calibrated Leap Motion and depth sensor. Multimedia Tools and Applications, pp. 1-25.
6. McCartney Robert, Jie Yuan, and Bischof Hans-Peter (2015) Gesture recognition with the Leap Motion Controller. Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).
7. Justino Edson JR, Flávio Bortolozzi and Robert Sabourin (2005) A comparison of SVM and HMM classifiers in the off-line signature verification. Pattern recognition letters, pp. 1377-1385
8. Chuan Ching-Hua, Eric Regina and Caroline Guardino. (2014) American Sign Language recognition using Leap Motion sensor." Machine Learning and Applications (ICMLA), 13th International Conference on. IEEE.
9. Potter Leigh Ellen, Jake Araullo and Lewis Carter (2013) The leap motion controller: a view on sign language. Proceedings of the 25th Australian computer-human interaction conference: augmentation, application, innovation, collaboration. ACM
10. Tobias Pistohl, Tonio Ball, Andreas Schulze-Bonhage, Ad Aertsen, Carsten Mehring. (2008) Prediction of arm movement trajectories from ECoG-recordings in humans. *Journal of Neuroscience Methods*, **l.** 167: pp.105-114
11. Bernhard H. P., and G. Kubin (1991) Speech production and chaos. XI International Congress Phonetic Sciences. Aix-en-Provence, France, 9th to 24th August 2008
12. Kantz Holger, and Thomas Schreiber (2004) Nonlinear time series analysis. Cambridge university press, 7. p. 309.
13. Makharoblidze T. (2012) Georgian Sign Language. Ministry of Education and Science, USIAD, Save Children International. Tbilisi.
14. Makharoblidze T. (2015) Georgian Sign Language Dictionary. Ilia State University; Shota Rustaveli National Scientific Foundation. Tbilisi. ISBN 978-9941-16-225-5: 1368

Received April, 2018