ARTO Categories and the ARTO String Notation

Chik Wing Cheung Aaron

2024-03-22

What are the ARTO categories

Signs exhibit patterned iconicity, in which distinct patterns are consistently used by gesturers and signers; the approach which result in these patterns are "iconic strategies", or "representational strategies"(Hwang, 2017). Previous studies have classified iconic strategies in various ways (Müller, 2013; Hwang, 2017).

As part of our HKSL project, we have developed the **ARTO categories**. The ARTO categories are a classification system of iconic strategies employed by morphemes. It describes how morphemes (and its constituent hands) represent their referents.

There are a total of five ARTO categories. Three of the five categories are adapted from Ortega et al. (2019), which in turn are adapted from Müller (2013). They are:

- ***Acting*** (**A**): Also known as *handling* in Hwang et al. (2017). Represents *human motion*, where a human is using or handling the object with their hands. The hands depict *themselves*.
    - Example: KEY in HKSL.
- ***Representing*** (**R**): Also known as *instrument* in Hwang et al. (2017). The hand *embodies an object* as a whole, where the hands do not represent the motion of the hands, but the body or the motion of the object. The hands depict the *object*.
    - Example: DRAW in HKSL: the dominant hand represents the brush, and the other hand represents the canvas.
- ***Tracing*** (**T**): Also known as *drawing*. The hands trace the outline of the referent.
    - Example: ROUND (圓), SQUARE (方) in HKSL.

Apart form these categories, we also added:

- ***Orthography*** (**O**): The hand embodies or traces a writing symbol (orthography).
    - Example: EAST (東) in HKSL.
- ***None*** (**N**): There is no iconicity strategy employed, or none of the categories apply.

– Example: AGE (年紀) in HKSL.

Two more categories exist in Müller (2013) and Hwang et al. (2017) which are unused in our classification system:

- ~~**Personification (P)**~~: Hwang et al. (2017). "The hands depict the size and shape of specific parts of non-human animate entities whose bodies are represented by signers'or gesturers'bodies".
    – Examples: BEAR in JSL; CHICKEN in HKSL.
    – Signs that should fit into *Personification* should also belong to *Representing (R)*.
- ~~**Molding (M)**~~: Müller (2013). "The hands mold or shape a transient scultupre, such as a picture frame or a bowl".
    – Examples: BALL in HKSL.
    – Signs that should fit into *Molding* should also belong to *Representing (R)*.

ARTO is submorphemic and complex

In Ortega et al. (2019), the iconic strategies are said to be employed by the entire sign (the lexeme), and each sign can be said to adopt one strategy only. Unlike Ortega et al. (2019), the ARTO category is a submorphemic feature, and each sign may be assigned one or more than one ARTO categories (including the N category).

Why are the ARTO categories submorphemic? In essence, this is because they are further divisible on the morphemic level, but having semantic meaning such that they must be above phonological levels. Take WHATSAPP (verb) in HKSL for example. The sign has one morpheme, which consists of two hands that are dependent on each other (one hand alone does not make the morpheme). The dominant hand is the letter "W"spelt out, therefore it belongs to the *Orthography (O)* category. The other hand represents the phone, and belongs to the *Representing (R)* category. Since the two different categories apply to different hands within the same morpheme, the ARTO categories must be below morphemic level.

As with many efforts of putting natural data into well-defined boxes, there are nuances in iconic strategies that make assigning ARTO categories quite difficult. One common issue is fusion of morphemes, resulting in one hand being assigned multiple ARTO categories. The sign UGLY in HKSL is an example of multiple ARTO categories in one morpheme. The sign is one-handed, but the hand carries two categories: it is *tracing (T)* because the hand outlines the face; it is also *none (N)* because of the "bad"handshape used. For this morpheme, we say that it belongs to both *tracing* and *none*. While we are not entirely convinced that multiple categories should be permitted, having them where needed comes with the advantage of identifying fusional morphology where it exists.

ARTO String Notation

To make the data collection process more efficient, I have developed the ARTO String Notation to efficiently capture the ARTO categories for all morphemes in a sign. This was developed as a superset and replacement of the earlier tentative "Fourslots Notation". With this Notation, the ARTO Categories of all morphemes in a sign can be represented by an ARTO Code.

There are two types of symbols: the ARTO Symbols and the reserved symbols.

ARTO Symbols

Category	Symbol
Acting	A
Representing	R
Tracing	T
Orthography	O
None	N

Reserved symbols

Function	Symbol
Morpheme delimiter	,
Empty placeholer	0
~~Duplicate flag~~ (unused)	~~1~~
~~Symmetrical flag~~ (compatibility use only)	~~2~~

Syntax

The overall structure of the ARTO Code in the String Notation can be expressed in this Regular Expression:

```
/^(?:(?:[ARTON]+[arton]{0,})(?:, ?)?)+$/g
```

Basics The ARTO Category of hand(s) in a morpheme can be represented by their corresponding symbol.

RUN 跑 has one morpheme that acts out walking. Therefore the ARTO Code is:

```
A
```

Multimorphemic   Morphemes are delimited by commas (`,`).

STARBUCKS 星巴克 has three morphemes. Morpheme 1 represents 'awakening' ; morpheme 2 acts out coffee; morpheme 3 traces the outline of the hair of the Starbucks Siren. The ARTO Code would look like:

`R,A,T`

Other hand   The ARTO Symbols are case-sensitive. *Uppercase* denote the *dominant* hand, the *lowercase* denote the *other* hand.

For example, BANANA 香蕉 has one morpheme. The dominant hand acts out the peeling of the banana; the other hand represents the banana. The ARTO Code for BANANA woud look like:

`Ar`

If each hand has its own iconicity strategy (like BANANA), each Symbol must be written out, even if both hands may have the same ARTO Category. For example, the dominant hand of VOTE 投票 represents the ballot, and the other hand represents the ballot box. The ARTO Code for VOTE would be:

`Rr`

Since the two hands have their own referents, it is incorrect to write: ~~R~~.

If the iconicity strategy is effective only with two hands together, the ARTO Code is marked on the morpheme. The ARTO Code for MOVIE 電影, where two hands form the shape of a film projector, would therefore be:

`R`

In this case, ~~Rr~~ is incorrect because each hand only represent one part of a "film projector".

In very rare occasions, the morpheme consists only of the other hand. In that case, write only the lowercase letter, for example: `r,A` (first morpheme uses the other hand only).

Symmetrical signs   Symmetrical signs (where one hand is the exact mirror image of the other) are written like a one-handed sign. For example, START 開始, where two hands form an opening motion, should be:

`R`

In the absence of handshape data, `2` is used to denote symmetry. If handshape data is coded elsewhere, `2` should not be coded.

`R2`

Note that signs must have hands *mirroring* each other to be said as *symmetrical*. A counterexample is COMPETITION 比賽, where individual hand movements where two hands alternately move up and down, and is thus *not symmetrical*.

Another counterexample is LIVE 生活, where two hands move togetheer, i.e. the movement of the other hand is not the *mirror* image of the dominant hand, and is also *not symmetrical.*

Hands in a symmetrical sign can overlap, as long as they are *mirroring.* An example MISUNDERSTANDING 誤會, where forearms cross but the hands mirror each other, and thus is *symmetrical.* However, FREEWAY 公路 (cars whizzling on a highway bridge) is *not symmetrical* because despite the same handshape and movement and the crossing, the hands are at a different elevation and are thus not mirror images.

Morphemes with multiple ARTO Categories   Some signs use iconicity strategies that can arguably be "in between"two or more categories; they would belong to more than one ARTO category. In that case, write all applicable categories in the same morpheme in the correct lettercase. There is theoretically no limit on the number of categories for each morpheme.

For example, one variant of URINATE 小便 (extended from TOILET 廁所) can classified as both *orthography* and *representing* for its only morpheme. The ARTO Code for that is:

`OR`

Note that the order of the letters do not matter, so it is acceptable to write the ARTO Code as `RO`. However, the standard way is to list in alphabetical order.

Multiple ARTO Categories for the other hand would look like this:

`Rrt`

Obsolete notational systems

Fourslots Notation

Deprecated –the Fourslots Notation encodes ARTO data with exactly four columns. This is deprecated because the N category is not coded and it is difficult to parse by machine.

For example, BADMINTON 羽毛球 in the ARTO String Notation is `Tr,A`. In Fourslots, it is:

Acting	Representing	Tracing	Orthography
1_2	O_1_1	D_1_1	0

References

Hwang, S.-o., Tomita, N., Morgan, H., Ergin, R., İlkbaşaran, D., Seegers, S., ⋯Padden, C. (2017). Of the body and the hands: patterned iconic-

ity for semantic categories. Language and Cognition, 9(4), 573–602. doi:10.1017/langcog.2016.28

Müller, C. (2014). Gestural modes of representation as techniques of depiction. In C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill & J. Bressem (Ed.) *Body–language–communication: An international handbook on multimodality in human interaction*, Volume 2 (pp. 1687-1702). Berlin, München, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110302028.1687

Ortega, G., Schiefner, A., & Özyürek, A. (2019). Hearing non-signers use their gestures to predict iconic form-meaning mappings at first exposure to signs. *Cognition, 191*, 103996.