

A Multilingual Analysis of the Notion of Instrumentality

Asanee Kawtrakul, Mukda Suktarachan (Kasetsart univ. Bangkok, Thailand),
Bali Ranaivo-Malancon, Pek Kuan Ng, (Univ. Sains Malaysia, Penang, Malaysia),
Achla Raina (IIT Kanpur, India),
Sudeshna Sarkar (IIT Kharagpur, India),
Alda Mari (Enst-Cnrs, Paris, France),
Sina Zarriess (Universität Potsdam, Germany),
Elixabete Murguia (Univ. Deusto, Bilbao, Spain),
Patrick Saint-Dizier (Irit-Cnrs, Toulouse, France)

Abstract

Instruments are expressed in language by various means: prepositions, postpositions, affixes including case marks, non-finite verbs, etc. We consider here 12 languages from five families in order to be able to identify the different meaning components that structure instrumentality.

1 Credits

This work has been made possible partly via the STIC-Asia cooperation framework.

2 Introduction

It is difficult to give a comprehensive definition of what instrumentality is. In WordNet it is defined as 'an artifact, or a set of artifacts, that are instrumental (i.e. behave as instruments) in accomplishing some end', i.e. reaching a certain goal. In this definition, the triple relation agent-instrument-goal (as in: *John cuts the bread with a knife*, where *John* is agent, *knife* is instrument that does the cutting, and *bread cut* is the goal), is left vague in what concerns the exact involvement of the agent and the instrument in the action, and the control the agent has on the instrument and on the action (Mari and St-Dizier 01).

If almost anything can be an instrument, we can nevertheless formulate a few criteria that we will try to elaborate in this paper. First, an instrument is basically non volitional. When humans play the role of instruments, they are obviously volitional, but the action is controlled by another agent who acts as an 'initiator agent', taking the initiative of the action. Instruments cannot be easily associated with traditional thematic roles, in fact this is not of much interest, because this is too superficial a

notion and also because instruments are generally modifiers, not arguments.

In this paper, we address instrumentality as conveyed by prepositions or equivalent means (e.g. postpositions, affixes). Our aim in this study is twofold: (1) to identify the conceptual facets of instrumentality so that a conceptual semantics can be defined in the spirit of (Talmy, 01, 03), (Wierzbicka 92, 96) and (2) to elaborate an accurate enough model for answering questions about instruments within a cooperative question-answering system. Instead of focusing on a specific language to elaborate all possible forms of instrumentality, we found it more adequate to develop a multilingual approach, considering languages from various families.

We consider for Europe: German, Spanish, French, Italian; for India: Kashmiri, Urdu, Hindi and Bengali; for the far-east: Thai and Malay; for Northern Africa and the Middle East: Arabic, and Berber dialects (in the group of Amazigh languages). In this paper, we use the first letter of each language to identify it: Thai, Malay, Hindi, Urdu, Kashmiri, Bengali, German, Spanish, French, Italian, Arabic, BeRber). We have also more or less adequately transcribed characters into latin characters. The upper case A is equivalent to aa.

3 An overview of preposition structures

3.1 European languages

German, Spanish, Italian and French, like most European languages have prepositions that introduce instrumental PPs. The most current prepositions are:

- German: mit, mit Hilfe von, mittels, durch, anhand, kraft,

dank, per.

- French: avec, par, au moyen de, grâce à, à l'aide de, à travers.

- Spanish: con, en, por, a través de, mediante, por medio de, a base de, con la ayuda de, gracias a.

- Italian: con, per mezzo di, tramite, per, grazie a, con l'aiuto di.

3.2 Arabic and Berber

Arabic essentially has the preposition *bi*, used as a prefix of the noun it heads:

Aktoub bi al kalami (I write with a pen).

bi can also be associated with a specific noun or deverbal form (e.g. 'by applying') to characterize in more depth the instrumental relation. This entails generic forms such as *bi-tarika*, *bi-istemali* (by means of, the first form reinforces the importance of the instrument, while the latter is more formal) and *bi-fadli* (thanks to). For example, we have a kind of nominal form *istikhdam* (= using or with the use of, constructed from the root 'use') in:

Fasser el massala bi istikhdam mithel (explain a problem 'with the use of' example). Arabic makes explicit the metonymy we have in European languages.

There are a few other prepositions such as *min khilal* (through) and *min a* to express the duality source + instrument of the argument (as in *drink with a bottle*, literally). In the spatial context, *au* can be used (to reach the top by this trail), and *ala* is used to express a channel of communication.

Berber is composed of a large number of dialects, some just spoken in small 'tribes', others in larger communities. It is basically dialectal, but a number of common elements can be identified. We consider here Berber from the Moroccan Rif, and the Algerian Kabyle. The main instrumental preposition is *əg*, but prefixes are also used *s-*, *ge-*, *th-*, *kh-* which affect the morphology of the noun they are attached to. The prefix *s-* is widely used, with some variants (*si*, *sei*, *so* in Kabyle), *əg* (also realized as *g'* or *ge* in some places), is appropriate only when the action is under full control of the agent. It is not employed when the instrument is abstract. In the Rif area, *s-* focusses on the instrument contributing to the action, it makes a kind of fusion between the

action verb and the instrumental noun. *ge* is only used in spatial contexts involving means of transportation (travelling by plane), *kh* is used only in spatial contexts involving paths.

3.3 The Indic language family

The four languages considered, Kashmiri (Raina, 02), Urdu, Hindi and Bengali, share some similarities due to their common origins but also contrasts that result from independent evolutions which are useful to our analysis. They all have suffixes and postpositions. Case suffixes (*vibakhti*) are added to nouns or pronouns, but case suffixes do not correspond strictly to thematic roles (*kAraka*). Postpositions may also be used instead of *vibaktis* or in conjunction with them, as a kind of re-inforcement. Basically, but with some nuances, Hindi, Urdu, Bengali and Kashmiri languages have the following kernels:

- direct instrument: *se* (H/U), *sity* (K), *-e/-te*, *diye* after null, *dwArA* after *-r* (B). Example: H: *raam ne chaabi se taala khola* (ram - erg. *key with lock open-past*= Ram opened the lock with a key).
- means instrument: *me* (H/U), *ke zariye* (U), *manz* (K), *-e/-te* (B). Example: U: *raam gaadi ke zariye daftar gayaa* (ram *car by means of office go-past* = Ram went to work by car).
- causal instrument: *ke kaaran* (H), *ki vajah se* (U), *kiny* (K), *kArANe*, *kripAya* (B). Example: K: *dil chi sayaaahat kiny amir* (Delhi be-present *tourism with rich* = Delhi is rich with tourism),
B: *dillI paryatan-er khAtire samriddha* or *dillI paryatan-er kripAya samriddha* (Delhi be-rich-er *thanks to tourism*)
- agentive instrument: *ke dwaraa* (H) *ke zariye* (U) *zariy*, *desi* (K), *diye* (B) after *-ke* when the nominal form is animate and specific, *diye* after null when the nominal form is animate and general (B). Example: H: *raam ne shyaam dwara apna kaam karvaaya* (ram erg. *shyam by self's work do-cause-past* = Ram got his work done by Shyam)
- action instrumentalised: *kar* (H/U), *kerith* (K). Example: U: *raam kuud kar ghar ke andar daakhil huaa* (ram *jump participle house*

into enter-past = Ram entered the house by jumping in it)

There are less important cases that capture notions like containment, which will be detailed in section 5. Postpositions may vary also depending on the semantic type of the NP they head.

In Bengali, case is indicated in several ways. Vibhaktis are suffixes that are added to the stems to form surface forms of words. In Bengali the nominal stems take one of the following suffixes: null or shunya vibhakti, -e, -te, -r, etc. Case is also indicated in Bengali by the use of postpositions. We have, for example: Cut bread with a knife, realized as either:

1. *chhuri diye ruti kATa* (knife diye bread cut), or:
2. *chhuri-te ruti kATa* (knife-te bread cut)

Postpositions in Bengali are derived from certain inflected forms of nouns, and also certain verbs in participle form. When these words are used as postpositions they are often not considered in their original sense but define a specific type of relationship of the noun phrase with the finite verb phrase in the sentence. These words appear in a fixed form (indeclinables) as postpositions. When a postposition is used to denote the case, the nominal word preceding the postposition takes on a vibhakti that is determined by the particular postposition. 'diye' is used after null vibhakti, 'dwArA' after -r vibhakti, etc.

3.4 Thai and Malay

Thai and Malay, although spoken in neighbour countries, are substantially different.

Thai, from the Thai-Kadai family, has 6 prepositions (Silapasarn, 98) to denote instruments, the most common being *doi* and *duai* which are used for concrete instruments, means of transportation, instruments close to manners, etc. *kap* is used when the instrument is a part of the body, while *thang* is used for means of transportation only. *tam* characterizes control of the agent, and *chak* is restricted to the instruments that convey an idea of source. Examples:

khian - duai - din so (Write - with - pencil)

pai - pa ris - doi - khruang bin (Go - Paris - by - plane)

These semantic distinctions are, however, often violated in colloquial Thai.

Malay, from the Malayo-polynesian family, has three ways to introduce instruments: preposition + NP, affixes and compounding. Affixed words are

built from stems which are instrumental nouns, this allows for the construction of the equivalent of PPs, based on the prototypical use of the instrumental noun. The most common being: prefixes: *beR-* (from *kuda*, horse, *berkuda*, on horseback), *meN-* (from *kunci*, key, *mengunci*, lock with key), prefix + suffix: *meN-* + *-kan* (from *paku*, nail, *memakukan*, to fasten with nails), and with suffix *-i* (from *ubat*, medicine, *mengubati*, by means of medicine). Prepositions occur as the head of PPs, and in verb particle constructions. PPs may also be subject complements, avoiding the use of verbs (*dia di rumah*, she at home). Besides affixes, Malay has 6 prepositions that denote instrumentality: *dengan*, *melalui*, *mengikuti*, *menerusi*, *dengan menggunakan*, *secara*.

A simple example is:

berhubung - melalui - telefon (communicate - by - telephone).

4 The meaning components of instrumentality

Let us now consider the different meaning components that emerge from our multilingual analysis. The results presented below are still exploratory due to the complexity of the notion. The distinctions made (e.g. between concrete and abstract instruments) may seem arbitrary: they are just meant to structure the presentation.

4.1 Concrete instruments

All languages studied have at least one basic instrumental mark operating over concrete objects (T: *duai*, M: *dengan*, H: *se*, U: *se*, K: *sity*, B: *diye*, -e, -te, G: *mit*, S: *con*, F: *avec*, A: *bi*, BR: *èg*). Several refinements are identified, for specific types of NPs, or to denote a specific intention:

- the instrument is a recipient (S: *en*) or, more generally, conveys an idea of container (e.g. spoon) (B: *-e kare*), the idea being is that the container is used to carry the object along a certain trajectory,
- the instrument is a part of the body (e.g. hand): T: *kap*. In this case, the instrument is not strictly artifactual.
- the goal is difficult to reach, it requires some efforts from the agent (S: *a base de*),

- the focus can be emphasized by using dedicated marks (G: mit Hilfe (von), Mittels (more formal: *Das Gericht hat mittels einstweiliger Verfügung den Drogenhandel untersagt* (the court has with provisional ordinance the drug traffic prohibited))).

The second major difficulty is prototypicality (Rosch, 78). When the instrument used is not very prototypical of the action, several languages reinforce the instrumental prepositions to, sort of, coerce the type of the noun so that it can become an acceptable instrument. We have examples in S: por medio de, B: sAhAjye, sahojoge, I: per mezzo di, F: au moyen de, par le biais de (biais= bias which directly expresses this idea), as in:

F: *Il a ouvert la porte au moyen d'un cric* (he opened the door by means of a jack).

At a conceptual level, it is quite difficult to characterize what is a prototypical instrument for a given action (characterized by subject-verb-object: *John opens the door*). Each event has its own prototypical instrument, making corpus studies extremely large, probably unfeasible. When searching on the web, we find an incredible variety of instruments to open a door, almost impossible to classify. Next, prototypicality is not a boolean notion: instruments are more or less prototypical. Since the instrument is very much dependent on the verb and on the object, we cannot foresee any form of incorporation in the verb that would give us indications. A direction could be to assume Qualia structures (Pustejovsky 91) associated with each potential instrument that describes the function of the object in the telic role. For example, key(X) would have open(X, door), with *door* being quite generic. This approach could work via a large lexical development for concrete nouns, it is much more risky when terms are abstract.

4.2 Abstract instruments

Abstract instruments (theorems, regulations, examples, etc.) are realized identically to concrete instruments, but with some typical marks such as: T: tam, H: dwAra, K: zariyi, B: dwArA, M. mengikut. At this stage, it is difficult to explain why marks are different from concrete instruments. An hypothesis could be that abstract instruments are closer to causes (see 5.5), or to more formal situations for which specific terms were developed (e.g. for G: kraft).

There are additional marks dedicated to particular fields: B: sahajoge, and A: min khilal when instruments are of type 'example' (explain with an example). U: -ke zariye, S: por medio de and G: Anhand, Kraft are more formal, stronger for Kraft and apply particularly to areas like juridical or psychological domains.

People and organizations can be seen as appropriate intermediaries for reaching a goal. They may be conceived as metaphorical instruments. Investigations show that people can get controlled much in the same way as concrete objects:

F: *Elle a informé Paul de son départ par Pauline* (She informed Paul of her leaving 'by' Pauline).

If we now consider: S: *Juan envió este paquete por correo* (John sent this parcel 'by' post)

Since post is the by-default medium to send packages, por is the only choice. Using more precise services, like FedEx, is considered to be an alternative way, in that case F: par, avec S: por, con are both acceptable.

4.3 Metaphorical instruments

Both concrete and abstract objects can be used metaphorically as instruments. Examples abound in the literature and on the Web. In 5.6 we examine the path metaphor which is very productive. Besides this case, we have a number of metaphors, such as: *write with your heart*, *fight with your head*, etc. These are not essentially different from metaphors observed in other situations (Lakoff and Johnson 99).

4.4 The overlap instrument-manner

In a number of cases, it is not very easy to make a distinction between instrument and manner. It seems there is a continuum between these two notions or even some form of overlap, where the object is both an instrument and a manner at various degrees, which may depend on context. A variety of marks contribute to characterize this overlap, manners at stake being quite diverse, but we will not go into the study of manners. Specific marks dealing with the manner/instrument ambiguity are: T: doi, G: durch (which is also used for metaphorical spatial uses), M: dengan menggunakan, S: en, con, a as in S: *escribir en/con rojo* (write in red),

T: *khian - duai - muek - daeng* (write - with - red - ink)

BR: *tète s-éfessen* (She-eats with-hands).

4.5 Causality

It is clear that, a priori, instruments can be viewed at various degrees as causes of an event. There is a kind of overlap between these two notions. Instruments are not volitional, so they are under the partial or full control of an agent (humans playing the role of instruments are also controlled by an agent). Typically I: a *causa di*, F: a *cause de*, S: a *causa de* signal that the instrument has brought about an event:

I: *Il castello e distrutto a causa di un violento incendio.* (The castle has been destroyed 'because of' a violent fire.)

Causality (e.g. Talmy, 01) being a complex notion, it is not surprising that instruments, viewed as intermediaries at various degrees, share some features with causes. For example in *cut the bread with a knife*, the cause of the bread being cut is the action of the agent, but also the use of a prototypical property of the knife: the knife does the cutting. In (Talmy 01), the instrument is embedded into the causing event:

(caused event) RESULTS FROM (causing event)
where the causing event has the structure:

Instrument ACT ON object, where object is bound or related in some way to the object in the caused event.

As analyzed in (Mari and Saint-Dizier, 01), instrumentality is the convergence of several factors:

- the degree of involvement of the instrument in the action, therefore, the fact that the instrument causes the action or is just a means managed by the agent who is the main cause,
- the type of control the agent has on the instrument for the action at stake, from full control to lack of control,
- the control the agent has over the action as a whole.

Indic languages and Thai are particular explicit on these matters. They have specific marks for two major cases:

1. agentive instrument, action not controlled by the agent: H: *ke dwAra*, U: *ke zariye*
K: *zariy*, *desi*, T: *doi*,
2. causal instrument that does most of the action, under the control of the agent: H *ke kAran*, U: *ki vajah se*, K. *kiny*, T: *duai*,

Berber allows *èg* only when the agent controls the instrument. The other cases are expressed by non prepositional forms.

4.6 Instruments and paths

Another productive situation is the use of spatial metaphors to express instrumentality. The use of F: *par* and other marks (e.g. in B., U.), show that there is a close link between instrumentality and path descriptions (spatial as well as temporal paths). This is a kind of metaphorical use of paths viewed as instruments (as can be seen in (Lakoff et al. 99): 'action is motion, goals are paths, actors are travellers'). Using an instrument parallels the use of paths in the domain of space.

Marks denoting paths or sources are of much interest. Some have really restricted uses, whereas others are more flexible. We observe the following main components:

- paths: T: *tam*, A: *min khilal*, S: *por*, a *través de*(*por correo*, by post), *de*, G: *durch*, F: *à travers*, note the distinctions, e.g. in M: *melalui* (metaphorical paths: M: *berhubung melalui telefon* (communicate by telephone)), *menerusi* (channel of transmission), H, U: *me*, *se*, T: *thang*. In B, -e and -te denote paths where the agent that does the action has no control, whereas *diye* and *dhare* involve at least a partial control from the agent. In M, metaphorical passages require *melalui*.
- sources: F: *à*, A: *min a*, T: *chak* (for concrete and abstract sources). Example: A: *Achroubou mina Karoura* (I am drinking with bottle), which is also a kind of manner.

The duality path/instrument is particularly visible in, e.g.:

K: *raam vot tshochi vati kiny gari* (ram reach-past short route via home = Ram reached home by the short route).

Another interesting phenomenon occurs when an argument is both an instrument and a path, as in *look at the moon in a telescope*. Telescope is indeed the instrument used and also the path through which one looks, or which the light traverses. This double facet of the argument is visible in surface realizations, where the preposition used is ambiguous between instrument (first preposition) and path (second one) readings: G: *mit*, *durch*, S: *con*, *por*, M: *dengan*, *melalui*, F: *avec*, *dans*. When one wants

to strongly stress the path interpretation, then a more path-oriented preposition is used, e.g. S: a travès de.

4.7 Means of transportation as instruments

Means of transportation (trains, spoons, boxes, envelopes, etc.), sometimes viewed as containers, and mediums of transportation (by air) receive a special treatment in a number of languages: T: doi, thang, M: menerusi, melalui (for metaphorical mediums and passages), H: me, U: ke zariye, K: manz, zariy, B: kare, -e kare, -ya kare, A: ala, BR: ge-, kh-, G: per, S: por, en, F: par. We have, for example: U: *raam gaadi ke zariye daftar gayaa* (ram car by means of office go-past = ram went to post office by car)

T: *pai - pa ris - doi - khrueng bin* (Go to - Paris - by - plane)

B: *Nouko-ya kare phuketa jAo* or *Nouko-ya phuketa jAo* (boat-e kare phuket go or boat-e phuket go = go by boat to Phuket)

A distinction is made between the medium and the means as for: M: secara, which is used for means of communication such as email or letters. If the agent has effective control over the means, then, for example, S uses con.

4.8 Language levels

Some marks are proper to formal discourse: G: Mittels, Kraft, Anhand, H: dwAra.

4.9 Positive or negative orientation

The languages we studied also abound in positive-oriented marks that express in a certain way the idea of 'thanks to': T: khop khun (+kah for feminine and krup for masculine), H: ke kAran, U: ki wajah, K: kiny, B: -er khAtire, (-er) kripAya, G: dank, S: gracias a, F: grâce à.

There are also several negative-oriented marks such as the following prepositional compounds: F: de la faute de, I: per colpa di, S: por colpa de (by the fault of), where the term 'fault' conveys a negative orientation.

4.10 Metonymies

In most languages, the prototypical action denoted by the instrument is implicit, it is analyzed as a metonymy: object for action. Action is inferred from the instrument and the verb in the given context. In a number of situations, A and M need to make explicit the action. For particular cases, gerundive forms may be preferred to PPs (but not

to be confused with manners, e.g. 'by swimming'), so that the verb that lexicalizes the action is present.

For example, in M, 'by the trail' in *to reach the top of a mountain by the trail* requires to make explicit how the trail is used: *dengan mengikuti denai itu* (litt.: with follow trail DET (same cases in Arabic and German)). Another case is: G: *Mit Flugzeugen lässt sich Geld verdienen*, (With planes you can money earn), where a concrete object replaces the whole procedure.

The metonymy could be reconstructed, for simple cases, by the Generative Lexicon (Pustejovsky 86), whose role is precisely to make explicit prototypical functions of objects via their telic role, as advocated above.

References

- Dowty, D., 1989, *On the Semantic Content of the Notion of Thematic Role*, in G. Chierchia, B. Partee, R. Turner (eds), *Properties, Types and meaning*, Kluwer Academic.
- Dowty, D., 1991, *Thematic Proto-roles and Argument Selection*, *Language*, vol. 67-3.
- Lakoff G., Johnson M., 1999. *Philosophy in the Flesh*. Basic books, NY, USA.
- Mari, A., Saint-Dizier, P., 2001, A Conceptual Semantics for Prepositions Denoting Instrumentality, in proc. 1st workshop on prepositions, Toulouse, and in *Syntax and semantics of prepositions*, P. Saint-Dizier (ed), Kluwer academic, 2006.
- Pustejovsky, J., 1991, *The Generative Lexicon*, *Computational Linguistics*, vol. 17, MIT Press.
- Raina, Achla M., 2002. *The Verb Second Phenomenon*, O.N. Koul and K Wali (eds.), *Topics in Kashmiri Linguistics*. Creative Books, New Delhi, India.
- Rosch, E., 1978. Principles of Categorization. In E. Rosch and B.B. Lloyd (eds.), *Cognition and Categorization*. Hillsdale : Lawrence Erlbaum Associates Publishers.
- Sinlapasarn, Upakitt. 1998. *Thai Grammar*. Thai Watthana Panich, Bangkok, Thailand.
- Talmy L., 2001, 2003. *Towards a Cognitive Semantics*, vol. 1 and 2. MIT Press.
- Wierzbicka, A., 1996. *Semantics primes and universals*. Oxford: Oxford University Press.
- Wierzbicka, A., 1992, *Semantic Primitives and Semantic Fields*, in A. Lehrer and E.F. Kittay (eds.), *Frames, Fields and Contrasts*. Hillsdale: Lawrence Erlbaum Associates, pp. 208-227.