

# Verbalization for Business Rules and Two Flavors of Verbalization for Fact Examples

Maurice Nijssen and Inge Lemmens

PNA Group, Geerstraat 105,  
6411 NP Heerlen,  
The Netherlands  
{maurice.nijssen, inge.lemmens}@pna-group.nl

**Abstract.** In the literature of NIAM, ORM, CogNIAM, OWL, Business Rules and SBVR [1, 5, 7, 16, 18] one increasingly encounters the modeling process of verbalization. Most fact based conceptual analysts are aware that process models need to be extended with fact schemas including concept definitions as well as concrete examples of input and output, satisfying the conceptual schema. Not adding this extension to process models regularly leads to misinterpretation and low productivity. Could there be a misunderstanding with respect to the process of verbalization as used in the various fact orientation approaches? In this paper we demonstrate that there are three *quite different* verbalization processes that have so far been referred to by the process name ‘verbalization’, resulting in quite different output. We will argue that all three types of verbalization are useful. To avoid further misunderstanding we propose to call these Verbalization for Business Rules and Verbalization for Fact Examples with and without using a fact type form (fact pattern), respectively, or more in the style of SBVR Structured English: (1) Verbalization with keywords, (2) Verbalization using a fact type form without keywords and (3) Verbalization without using a fact type form and without keywords. Each of these has a specific aim and each is useful in conceptual modeling.

## 1 Introduction

In this paper we will demonstrate that there are three useful types of verbalization processes. Such verbalization modeling processes are an essential part of domain-specific and generic (meta) conceptual schema modeling as well as business communication. The importance of processes for modeling has been stressed in the literature [3, 6, 8, 9, 11, 14, 15]. Although the process ‘verbalization’ has been used extensively in the last 35 years in the fact orientation community [11], it turns out now that there are three different types of verbalization. The first we will call Verbalization for Business Rules and the second and third Verbalization for Fact Examples, *with* and *without* using a fact type form. “Object-Role Modeling (ORM) is a fact-oriented approach for modeling, transforming and querying information in terms of the underlying facts of interest, where facts and rules may be verbalized in language readily understandable by non-technical users of the business domain” [2].

## 2 Is There a Problem with Respect to Verbalization in the ORM Community?

As we will be making extensive use of the CogNIAM knowledge triangle, the triangle is introduced in figure 1:

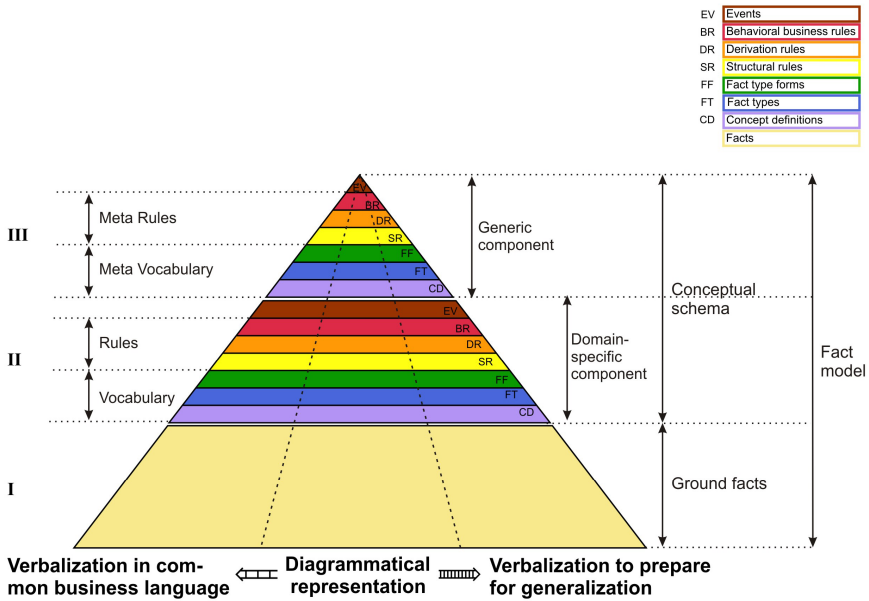


Fig. 1. Knowledge Triangle

In figure 1, ‘Verbalization in common business language’ should be interpreted as a synonym for ‘Verbalization for Business Rules’; ‘Verbalization to prepare for generalization’ is a synonym for ‘Verbalization for Fact Examples’.

In the ORM community the word ‘verbalization’ is used interchangeably for three very different purposes. Do we have a clear case of an undetected homonym? As this is hardly recognized, ambiguity and misunderstanding may occur.

Let us give an example from the SBVR specification, annex E.2.2.1.9 [13]:

<rental> incurs <late return charge>  
 <rental> incurs <late penalty charge>

These are two specific fact type forms (each representing a fact type) at the domain-specific level (level II in the knowledge triangle (figure 1)). In the CogNIAM knowledge triangle these are specified in the middle vertical lane, the lane used by conceptual modelers (see figure 1). A rule for the associated fact types could be:

Each rental incurs at most one late return charge.

This rule is included in the left part of the middle level (level II), aimed at business persons that prefer to read rules in unambiguous Structured English. The step is going from the middle part of the middle level (level II) to the left part of the middle level (level II). On page 11 of Halpin and Morgan's book [9] the word *verbalize* is used to denote this process.

On page 9 of the excellent book of Halpin and Morgan's book the word *verbalize* means taking the middle part of the lowest level (level I) as input, use the fact type form of the middle level (level II) and produce the natural language sentences of the right part of the lowest level (level I).

If we would use the above two fact type forms of E. 2.2.1.9, and apply the above described verbalization process, we would obtain sentences for the right part of the lowest level (level I) of the following format:

- (1) 1122334455 incurs 150
- (2) 1122334455 incurs 170
- (3) 1234567890 incurs 180

What is the interpretation of 150 in sentence 1? Is it a late return charge, or a late penalty charge? Hence there is ambiguity.

### 3 How to Solve This Ambiguity?

This ambiguity can be solved by distinguishing two kinds of verbalization and the associated two kinds of fact type forms. The two kinds of fact type forms are needed if we want to eliminate the ambiguity. Both have been used in ORM and SBVR interchangeably.

Hence for the verbalization with the aim to express a (business) rule in unambiguous Structured English (Verbalization for Business Rules) the two fact type forms given above suffice. However, for the Verbalization for Fact Examples using a fact type form, not being rules, *another* fact type form pattern is recommended. In this case it would be:

**rental** <rental> *incurs late return charge of* <late return charge> Euros  
**rental** <rental> *incurs late penalty charge of* <late penalty charge> Euros

The words in bold type are the qualifications of the nearest role they are bound to. In case of ambiguity, the association is explicitly stored.

If we would use these last two fact type forms – let us call them *qualified* fact type forms – we would obtain the following three natural language sentences:

- (11) Rental 1122334455 incurs late return charge of 150 Euros.
- (12) Rental 1122334455 incurs late penalty charge of 170 Euros.
- (13) Rental 1234567890 incurs late penalty charge of 180 Euros.

This ambiguity problem is one of the problems that can be eliminated with the systematic use of the CogNIAM knowledge triangle and associated related concepts.

Please note that the initial step of verbalization where an end user talks to a colleague does *not* make use of a fact type form, as it is not yet available in the development phase at the time this kind of verbalization is used. However, when the fact type form is derived in the Conceptual Schema Design Procedure [9], this fact type form can be used in the communication of ground facts. This Verbalization for Fact Examples using a fact type form (qualified) is illustrated in Figure 2.

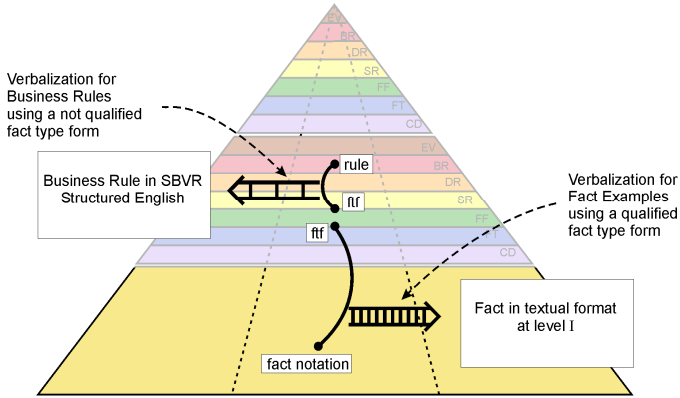


Fig. 2. Two different kinds of verbalization

#### 4 Verbalization for Business Rules (Using a Fact Type Form)

For the process “Verbalization for Business Rules”, the verbalizer takes the fact type form (not qualified) and the rule expressed in diagrammatic format and adds one or more keywords to get a rule in SBVR Structured English. Example:  
 Fact type form:

rental *incurs* late return charge

Verbalization towards rule:

Each rental *incurs* at most one late return charge.

Above form is nearly always used in SBVR.

#### 5 Verbalization for Fact Examples Using a Fact Type Form

For the process “Verbalization for Fact Examples (using a fact type form)” the verbalizer takes the given representation of facts (at level I), uses the fact type forms (of level II) and produces fact examples at the ground fact level (level I). Example:  
 Fact type form:

rental <rental> *incurs* late return charge of <late return charge> Euros

Values to be used:

1122334455; 150

Verbalization towards natural language sentences:

Rental 1122334455 incurs late return charge of 150 Euros.

In the paper Carver and Halpin recently presented at the EMMSAD conference [4], the authors asked the question with respect to normalization:

*What went wrong? A research-historical, psychological excursus*

*When such a fairly obvious error in a standard, accepted theory goes undetected for three decades, one cannot help but ask what went wrong. [...] Thus, even more interesting than the question what went wrong originally, is the question why no one detected it for so long. [...] The idea that their assumption was flat wrong, was too radical a thought to occur to anyone. [...] A contributing factor to this oversight, however, seems to have been the aforementioned, and mathematicians' natural tendency to focus on the syntax – to the neglect of semantics, in this case.*

Does the ORM conference community have the same tendency?

## 6 What Caused the Misunderstanding?

We could demonstrate that two kinds of verbalization produce the same result at the level of the knowledge triangle where there are only ground facts [13], i.e. facts that have no grammar function. However, if we apply the same types of verbalization at the domain-specific component of the conceptual schema or the generic component, then the results are *very different*.

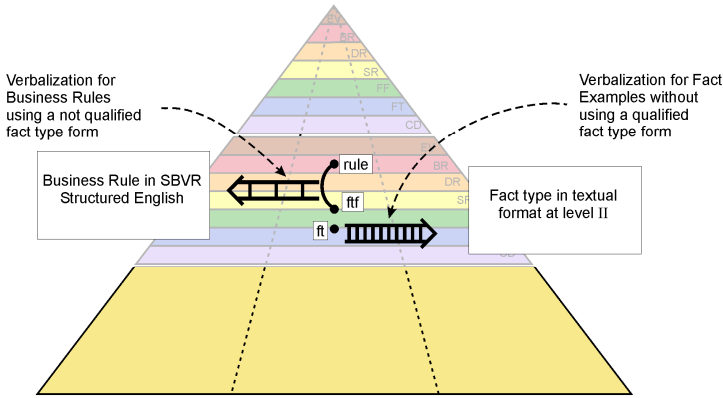
## 7 Verbalization for Fact Examples without Using a Fact Type Form

In section 3 we discussed the Verbalization for Business Rules. Now we use the same fact type which is represented by the fact type form rental incurs late return charge for the Verbalization for Fact Examples without using a fact type form. In this case we do not have the fact type form, as it is not yet available as we are in the development phase of deriving this fact type form. Therefore we have to verbalize this fact type. In table 1 the Verbalization for Fact Examples without using a fact type form is given with respect to the role of the fact type. We have given the fact type the abbreviation rilrc.

**Table 1.** The result of verbalizing the fact type's roles

Fact type	rilrc	has role	rental	.
Fact type	rilrc	has role	late return charge	.

These two different kinds of verbalization at the level of the domain-specific component of the conceptual schema are shown in Figure 3.



**Fig. 3.** Two different kinds of verbalization at level II

As shown above, by making systematic use of the knowledge triangle, these misunderstandings can be avoided.

The diagrammatic representation of the knowledge triangle of figure 1, 2 and 3 has been used in the last eight years by PNA as part of an approach that is now called CogNIAM. The knowledge triangle contains the core concepts of CogNIAM [12] in an easily memorizable way and in diagrammatical coherence. It can very quickly be seen that there are three levels, two of which share the same structure. These levels are:

- I. facts without a grammatical function, called ground facts in SBVR [13],
- II. facts with a domain-specific grammar function, called the domain-specific component of the conceptual schema in SBVR and
- III. facts with a generic or meta grammar function, called the generic component of the conceptual schema in SBVR.

The knowledge triangle has 7 knowledge classes in the domain-specific and generic component of the conceptual schema. One of these classes, events, is not part of SBVR. Why are events incorporated in the knowledge triangle? To facilitate respectful discussions with other communities such as UML.

The major common ground of ORM, SBVR, OWL, and CogNIAM is facts. The knowledge triangle uses nine parts (3 levels, at each level a left, middle and right side) to show the coherence between:

- A. A graphical or report representation of ground facts (middle part of the lowest level (level I)) and its textual representation (right and left part of lowest level).
- B. The textual or graphical representation of the domain-specific component of the conceptual schema (middle part of middle level (level II), having 7 knowledge classes) and the verbalization of rules in facts (right part of the middle level) as an *intermediate* step towards the next level, the generic level of the conceptual schema. The step from the middle part of the middle level to the left part of the middle level is extensively used in SBVR.

Please note that there is a distinction between three types of verbalization: one kind is used to express a rule in 'language that is readily understood by the business domain expert' [13, 10.1.1.2] and this is presented in annexes C, F and I of

SBVR. This is without doubt a useful form of verbalization. We call this Verbalization for Business Rules. However, CogNIAM has during many years used a second and third type of verbalization productively, to show the total coherence of all facts at all levels. One of the functions of Verbalization for Fact Examples without using a fact type form is to be a step towards the next level of grammar. This makes it possible to demonstrate that there are only three levels of facts and that the meta level describes itself, a major intellectual step in total understanding of the structure and modeling of conceptual schemas. The self description is clearly visible in figure 4, at the top of the knowledge triangle. The Verbalization for Fact Examples using a fact type form is used in the business communication of ground facts.

- C. The textual or graphical representation of the generic component of the conceptual schema (middle part of the upper level (level III)) and the verbalization of rules in facts as an intermediate step (right part of the upper level) towards the ‘next’ level. Because the generic component is self-describing, in this case the next level is the same (i.e. third) level. The step from the middle part of the upper level to the left part of the upper level is extensively used in SBVR.

By using two steps, Verbalization for Fact Examples without using a fact type form and generalization, from a best practice in requirements engineering, it is possible to demonstrate in a fairly easy to follow and understandable way that there are only three levels of facts and the coherence of major concepts of the Fact Orientation Approaches. SBVR is the first official specification or standard in business computing where concept definitions are first class citizens. The concept definitions form the bridge between the formal and the informal world, hence are vital for business communication. One of the 7 knowledge classes at the domain-specific and the generic level, Concept Definitions form the basis for each of the conceptual schemas, the domain-specific component and the generic component. The major concepts to be introduced are:

- Fact instances
- Concept definitions
- Fact types
- Fact type forms (with their subtypes: sentential forms and noun forms)
- Structural rules (Constraints)
- Structural rules (Derivation rules)
- Behavioral business rules and
- Event rules.

## 8 Summary and Recommendations

We started from the level of ground facts (level I in figure 4 above). In the old days one would say we started at the database level. At this level the facts have no grammatical function. By applying verbalization for fact examples (arrow 1) to the diagrammatic representation at the ground fact level, the results are the facts in a textual

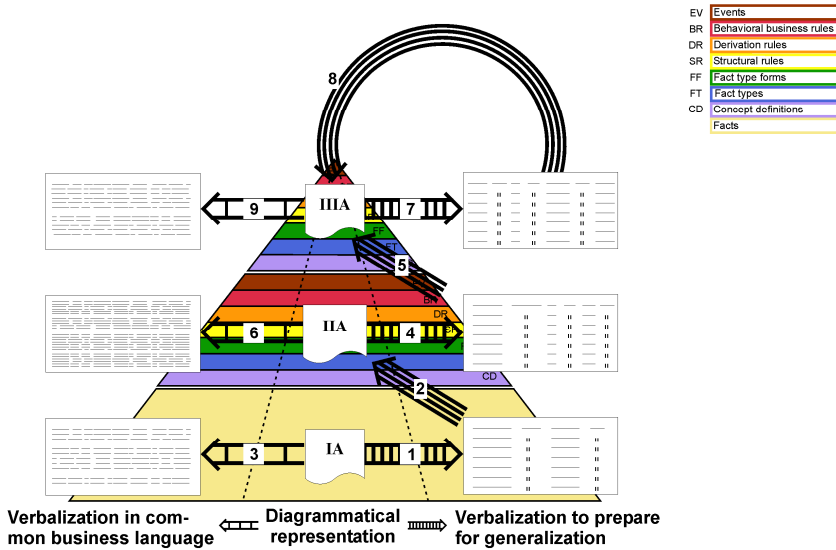


Fig. 4. Knowledge triangle with process aspects

format. By applying generalization (arrow 2) to the textual representation at the ground facts level the core of the domain-specific component of the conceptual schema was obtained in a diagrammatic format.

This diagrammatic format of the domain-specific component of the conceptual schema describes the meaning of terms at the ground fact level and it specifies the rules for fact populations, fact population transitions and for fact generation. Hence this determines the next lower level of facts and describes their semantics.

Next, by applying verbalization for fact examples (4) to the diagrammatic representation at the level of the domain-specific component of the conceptual schema, we obtain the textual format of the domain-specific component.

Continuing this process, by using generalization (5) at level II, the result is a diagrammatic representation of a core part of the generic component of the conceptual schema. This diagrammatic format of the generic component of the conceptual schema stipulates (6) the semantics and rules for the domain-specific component of the conceptual schema. Again, by applying Verbalization for Fact Examples (7) to the diagrammatic format of the generic component, we obtain a textual representation of a core part of the generic component of the conceptual schema.

As was illustrated previously, by applying generalization (8) at level III, the result was the identical representation of the generic conceptual schema, i.e. there is no higher conceptual level than level III.

The beauty of the generic component of the conceptual schema is that in effect it stipulates itself (9)!

It is our recommendation to distinguish the three kinds of verbalization by introducing clear terminology: we call the first kind of verbalization Verbalization for Business Rules (constraints, derivation rules or rules of guidance for humans) or Verbalization *with* Keywords. The second type we call Verbalization for Fact Examples



using a fact type form or Verbalization using a fact type form and *without* Keywords. The third we call Verbalization without using a fact type form and *without* Keywords. The major function of the process Verbalization with Keywords is to have the rules understood by business experts. The major function of Verbalization using a fact type form and without Keywords is to be able to communicate ground facts for business communication-purpose. The major function of Verbalization for Fact Examples without using a fact type form and without keywords is to have a deterministic procedure to derive the domain-specific component of the conceptual schema, to derive the generic component of the conceptual schema, often called the meta-schema and to demonstrate that the same procedure applied to the meta-schema will result in (a subset of) the meta-schema. Hence, all three types of verbalizations are useful, and each has its audience.

## References

1. Anderson Healy, K.: Special Report on SBVR. *Business Rules Journal* 9(3) (2008) ISSN: 1538-6325
2. Balsters, H., Carver, A., Halpin, T., Morgan, T.: Modeling Dynamic Rules in ORM. In: Meersman, R., Tari, Z., Herrero, P. (eds.) *OTM 2006 Workshops*. LNCS, vol. 4278, pp. 1201–1210. Springer, Heidelberg (2006)
3. Bollen, P.: Using Fact-Oriented Instructional Design. In: Meersman, R., Tari, Z., Herrero, P. (eds.) *OTM 2006 Workshops*. LNCS, vol. 4278, pp. 1231–1241. Springer, Heidelberg (2006)
4. Carver, A., Halpin, T.: Atomicity and Normalization. In: *Thirteenth International Workshop on Exploring Modeling Methods in Systems Analysis and Design (EMMSAD 2008)*, Montpellier, France (2008)
5. Chapin, D.: SBVR: What is now Possible and Why? *Business Rules Journal* 9(3) (2008), <http://www.BRCommunity.com/a2008/b407.html>
6. Damien, T., Vereecken, J., Christiaens, S., de Leenheer, P., Meersman, R.: T-Lex: A Role-Based Ontology Engineering Tool. In: Meersman, R., Tari, Z., Herrero, P. (eds.) *OTM 2006 Workshops*. LNCS, vol. 4278, pp. 1191–1200. Springer, Heidelberg (2006)
7. Hall, J.: Business Semantics of Business Rules. *Business Rules Journal* 5(3) (2004), <http://www.BRCommunity.com/a2004/b182.html>
8. Halpin, T., Curland, M.: Automated Verbalization for ORM2. In: Meersman, R., Tari, Z., Herrero, P. (eds.) *OTM 2006 Workshops*. LNCS, vol. 4278, pp. 1181–1190. Springer, Heidelberg (2006)
9. Halpin, T., Morgan, T.: *Information Modeling and Relational Databases*. Morgan Kaufmann, San Francisco (2008)
10. Hansen, J., dela Cruz, N.: Evolution of a Dynamic Multidimensional Denormalization Meta Model Using Object Role Modeling. In: Meersman, R., Tari, Z., Herrero, P. (eds.) *OTM 2006 Workshops*. LNCS, vol. 4278, pp. 1160–1169. Springer, Heidelberg (2006)
11. Nijssen, S.: On Experience with Large-scale Teaching and Use of Fact-Based Conceptual Schemas in Industry and University. In: *Proceedings of the IFIP WG 2.6 Conference on Data Semantics*, North-Holland Publishing Company, Hasselt (1986)
12. Nijssen, S., Bijlsma, R.: A Conceptual Structure of Knowledge as a Basis for Instructional Designs. In: *Proceedings of the Sixth International Conference on Advanced Learning Technologies (ICALT 2006)*, pp. 7–9. IEEE Computer Society, Los Alamitos (2006)

13. OMG (Object Management Group), Semantics of Business Vocabulary and Business Rules (SBVR), v1.0. Online as document 08-01-02 (2008),  
<http://www.omg.org/spec/SBVR/1.0/PDF>. SBVR 1.0 and supporting files  
<http://www.omg.org/spec/SBVR/1.0/>
14. Piprani, B.: Using ORM-Based Models as a Foundation for a Data Quality Firewall in an Advanced Generation Data Warehouse. In: Meersman, R., Tari, Z., Herrero, P. (eds.) OTM 2006 Workshops. LNCS, vol. 4278, pp. 1148–1159. Springer, Heidelberg (2006)
15. Ross, R.: Business Rule Concepts: Getting to the Point of Knowledge, 2nd edn. Business Rule Solutions LLC, Houston (2005)
16. Ross, R.: The Emergence of SBVR and the True Meaning of 'Semantics': Why You Should Care (a Lot!) ~ Part 1. Business Rules Journal 9(3) (2008),  
<http://www.BRCommunity.com/a2008/b401.html>
17. Sowa, J.: Fads and Fallacies about Logic. IEEE Intelligent Systems 22(2), 84–87 (2007),  
<http://www.jfsowa.com/pubs/fflogic.htm>
18. Vanthienen, J.: SBVR: The ABCs of Accurate Business Communication. Business Rules Journal 9(3) (2008), <http://www.BRCommunity.com/a2008/403.html>