A Hierarchical Unification of LIRICS and VerbNet Semantic Roles

Claire Bonial[*] **William Corvey**[*] **Martha Palmer**[*]
VolhaV. Petukhova[**] **Harry Bunt**[***]
[*]**Department of Linguistics, University of Colorado at Boulder**
[**]**Human Speech and Language Technologies, Vicomtech-IK4 - Visual Interaction Communication Technologies**
[***]**Department of Computer Science, Tilburg University**
{Claire.Bonial,William.Corvey,Martha.Palmer}@colorado.edu
vpetukhova@vicomtech.org
Harry.Bunt@uvt.nl

Abstract—**This research compares several of the thematic roles of VerbNet (VN) to those of the Linguistic InfRastructure for Interoperable ResourCes and Systems (LIRICS). The purpose of this comparison is to develop a standard set of thematic roles that would be suited to a variety of natural language processing (NLP) applications. We draw from both resources to construct a unified set of semantic roles that will replace existing VN semantic roles. Through the process of comparison, we find that a hierarchical organization of coarse-grained, intermediate and fine-grained roles facilitates mapping between semantic resources of differing granularity and allows for flexibility in how VN can be used for diverse NLP applications; thus, we propose a hierarchical taxonomy of the unified roleset. The comparison and subsequent development of the hierarchy reveals a level of granularity shared by both resources, which could be further developed into a standard set of thematic roles for the International Organization for Standardization (ISO).**

I. MOTIVATION

The ideal set of thematic roles should be able to concisely label the arguments of any relation; however, what this set of roles should be has long been a subject of dispute in the linguistic community. In our current endeavor to create a possible standard set of thematic roles for the International Organization for Standardization (ISO), we have undertaken a systematic comparison of two semantic resources: LIRICS[1] and VerbNet (VN) [1], [2]. We propose a unified set of semantic roles that will be incorporated into VN, which will allow VN to map more easily to an ISO standard set of semantic roles still under development. Additionally, we suggest a hierarchical arrangement of these roles based on semantic feature inheritance and feature constraints.

II. BACKGROUND

A. Overview of LIRICS

The EU-funded project LIRICS was set up as a spin-off of ISO TC 37/SC4, with the aim of exploring the possibility of establishing sets of annotation concepts, defined in accordance with ISO standard 12620 as so-called data categories, for syntactic, morphosyntactic, and semantic annotation and lexical markup. In the part of the project concerned with semantic annotation, several approaches and existing annotation schemes for semantic role labeling were analyzed and compared with respect to (1) description model; (2) granularity; (3) definition of semantic roles; and (4) consistency and reliability of annotation [3]. Based on this study, it was concluded that semantic roles should be defined:

○ as neither syntactic nor lexical structures but as semantic categories;
○ by virtue of distinctive semantic properties;
○ as not restricted to only a few specific verb (noun, adjective) classes;
○ as relational notions that link participants to an event, describing the way the participant is involved in an event (e.g. does he act intentionally; is he/it affected, changed, manipulated by other participants; does it come into existence through the event), rather than by internal properties.

A set of 29 semantic roles[2] was defined by listing for each a characteristic set of entailments. These entailments were converted into a set of orthogonal properties, e.g. [+/- intentionality], [+/- independent existence], etc. (see also [4] and [5]). For example, the *Theme* role is defined as a participant in an event or state who (i) is essential to the event taking place but does not have control over the way the event occurs; (ii) is not structurally changed by the event; in a state, is in a fixed position or condition throughout the state; (iii) in a state, is essential to the state being in effect, but it is not as central to the state as a participant in the *Pivot* role.

Different levels of granularity are distinguished, where a low-level semantic role inherits all the properties of a

---

[2]This set includes 11 roles which are central to any event, e.g. *Agent, Theme, Patient*; 10 adjunct roles, e.g. *Time, Location, Manner*; and 8 sub-roles for *Time* and *Location*, e.g. *Duration, Frequency, Path*. For definitions and examples, see http://semantic-annotation.uvt.nl/LIRICS_semroles.htm.

high-level role and has an additional feature, which reflects additional or more specific entailments. This results in a shallow hierarchy for several semantic roles, such as *Time*: *Initial_Time, Final_Time, Frequency* and *Duration*.

The LIRICS set of semantic roles was evaluated with respect to redundancy, completeness and usability for reliable, consistent annotation using a multilingual test suite including English, Dutch, Italian and Spanish (see [3] and [6]). Please refer to Table 1 for the complete set of LIRICS semantic roles.

B. Overview of VN

VN is a lexicon of approximately 5200 English verbs, the organization of which is based upon Levin's (1993) verb classification. Like Levin's classification, the verb classes of VN are organized according to the syntactic behavior of verbs; specifically, VN describes the sets of diathesis alternations that are compatible with each verb in the lexicon. For example, the verb *break* is compatible with the inchoative variation, in which the patient is the subject of the verb ('The window broke'), as well as the causative variation, in which an agent is the subject of the verb and the patient is the object ('The little boy broke the window'); however, the verb *appear* is compatible with the inchoative variation ('A rabbit appeared out of the magician's hat'), but is not compatible with the causative variation (*'The magician appeared a rabbit out of his hat') ([7]:3).

Although the basis of the classification is syntactic, the verbs of a given class do share semantic regularities as well because, as Levin hypothesized, the syntactic behavior of a verb is largely determined by its meaning [7]. Thus, there is a fundamental assumption in Levin's work that the syntactic frames compatible with a particular verb are a reflection of the underlying semantics; however, the relationship between syntax and semantics was not made explicit in her work. VN expands upon Levin's work by making this relationship explicit through the assignment of thematic roles to each syntactic argument in a given verb class as well as the use of semantic predicates denoting relations between participants and events. For example, the Put class of VN contains the example usage, 'I put the book on the table.' This is represented in VN as follows:

Syntactic representation:
NP V NP PP
Agent V Theme Destination

Semantic representation:
MOTION(DURING(E), THEME)
NOT(PREP: on (START(E), THEME, DESTINATION))
PREP: on (END(E), THEME, DESTINATION)
CAUSE(AGENT, E)

Thus, in each verb class, the thematic roles are used to link syntactic alternations to semantic predicates, which can serve as a foundation for further inferencing. For this reason, VN relies to an extent on syntactic features.

VN has been used for numerous NLP tasks, most notably, semantic role labeling [2], [8], [9]; additionally, VN has been used for the automatic generation of representations of unseen verbs (here, for example, as dialog systems: [10], [11]). Because VN is intended to be practical for a variety of NLP tasks and is organized into verb classes, the thematic roles in VN differ in several key aspects from those of LIRICS. In its current state VN makes use of (1) commonly used, coarse-grained roles like those of LIRICS (e.g. *Agent*), (2) roles that are specific to certain classes of events, which are intended to convey key semantic components of some verb classes (e.g. *Topic*, which is restricted to verbs of communication) (3) roles that are in part syntactically motivated (e.g. *Predicate*, which is used for classes with predicative complements) (4) roles that are distinguished by internal properties of the participant (e.g. [+animate]) [1], [2]. In contrast, LIRICS uses only roles that are not restricted to specific verb classes, not linked to particular syntactic structures, and not related to internal properties of participants. VN is distinct from LIRICS in these respects in part because, from its inception, VN has been an NLP application-oriented framework. For instance, early versions of VN were evaluated on Parameterized Action Representations (PARS) [1], [12], [13].

Table I provides a complete list of current VN thematic roles, as well as comparisons to those of LIRICS where possible. Note that these roles are those currently found in VN; however, there are plans to change these roles, as discussed in subsequent sections here, and in a previous related paper [14]. Notice that VN has roles that either do not have a mapping to a LIRICS role, or several roles map to a single LIRICS role. This reflects the greater specificity of VN roles, which can be unique to certain event types and therefore not included in the LIRICS roleset. Several of these cases are VN roles labeled with a '1' or '2', which are thought to be symmetrical participants (more than one participant that is involved in an event/state in the same way). These apply to classes of verbs where two participants cooperate equally in a shared action, such as the Chit_Chat and Contiguous_Location classes containing verbs like *chat* and *border* respectively.

Although the use of class-specific and/or syntactically motivated roles may be justified by some of the intended NLP applications of VN, such roles are not suitable for a set of standard thematic roles intended to capture purely semantic generalizations of any event or state without regard to syntactic realization. Nonetheless, from a practical perspective, VN has opted to keep roles that are specific to certain classes of events that may be helpful in distinguishing classes of verbs for particular NLP tasks. For example, research in automatic semantic role labeling has demonstrated the importance of the level of granularity of semantic roles: Yi et al. [15] and Loper et al. [16] both demonstrate that because VN labels are more generalizable across verbs than PropBank [17] labels, they are easier for semantic role labeling systems to learn; however, Merlo and Van Der Plas [18] found that the differing levels of granularity of PropBank and VN were both useful, and therefore suggest complementary use of both resources.

VN role	LIRICS role	VN (continued)	LIRICS (continued)
Actor	Agent	Patient	Patient
Actor 1	Agent	Patient1	Pivot
Actor 2	Partner	Patient 2	Patient
Agent	Agent	Predicate	-
Asset	Amount	Product	Result
Attribute	Attribute	Proposition	-
Beneficiary	Beneficiary	Recipient	Goal
Cause	Cause, Reason	Source	Source, Initial_Location
Destination	Final_Location	Stimulus	Theme
Experiencer	Pivot, Patient	Theme	Theme
Extent	Amount, Distance	Theme1	Theme
Instrument	Instrument	Theme2	Theme
Location	Location, Setting	Time	Time
Material	Source	Topic	Theme

TABLE I
VERBNET ROLES IN COMPARISON TO LIRICS.

Roles that are specific to certain VN classes can be especially helpful in distinguishing classes when the manual annotations of SemLink [16], [19], which includes VN role annotations, are used for training. These annotations reflect differences in the syntactic patterns of coarse-grained and fine-grained roles. For example, the fine-grained role *Topic*, specific to verbs of communication, is more likely to be realized in the form of a complement clause, in comparison to the more coarse-grained role *Theme*, which is more likely to be realized in the form of a noun phrase.

Thus, the challenge of integrating an ISO standard set of thematic roles within the organization of VN is one of finding a suitable mapping to the standard so that one can take advantage of the strengths of a set of roles applicable to all verbs, while maintaining the practicalities associated with roles that help to distinguish certain classes of verbs. To this end, a hierarchy of semantic roles is under development that will be implemented in the future as an improvement to VN.

In this hierarchy, coarse-grained roles applicable to all verbs will be superordinate roles. The addition of restrictions on these roles will form the basis of what can be thought of as subordinate roles. For example, in the hierarchy, *Experiencer* is subordinate to *Patient*, thus it inherits all features of *Patient*, but is characterized by the additional restriction [+awareness]; therefore, an *Experiencer* is a *Patient* who is aware of the process denoted by the verb that the participant is undergoing. The role *Patient* is underspecified for awareness. In some cases, the subordinate roles will be specific to certain event classes; for example, *Topic* is a *Theme* with the additional restriction [+information_content]; therefore, a *Topic* is a *Theme* in events of information transfer or communication verbs.

The motivation for this hierarchy is that users can select the level of granularity that is ideal for their task. The coarse-grained superordinate role can be used in place of a finer-grained role for tasks that require a roleset that has the broadest coverage across all verbs; conversely, fine-grained and class-specific roles can be used for tasks that benefit from information that helps to distinguish classes of verbs.

III. MERGING THEMATIC ROLESETS THROUGH HIERARCHICAL ORGANIZATION

We will now review key elements of the taxonomy and define the set of thematic roles included.

A. A Semantic Role Hierarchy

Please find an illustration of the hierarchy in Figure 1.[3] A hierarchy is a logical property of a set of elements, which can be easily visualized. The hierarchy discussed here is constructed with two goals in mind. First, it includes higher-level and intermediate nodes that allow all thematic roles to be subsumed under a single root node. Second, major distinctions between trees of thematic roles are theoretically-motivated. We propose a root node of *Participant*.[4] *Participant* dominates the proto-roles *Actor* and *Undergoer* [20], which in turn dominate both intermediate and currently instantiated semantic roles. Roles that are not in a parent-child relationship can co-occur. For instance, *Agent* and *Patient* co-occur frequently but *Patient* and *Experiencer* never co-occur in VN. As previously mentioned, parent-child relationships are governed by additional restrictions placed on the lower node, and in principle nodes that include restrictions may be thought of as either named semantic roles, such as *Recipient*, or as higher nodes plus feature sets (here, *Goal* [+concrete][+animate]).

B. Thematic Role Inventory and Definitions

A text description of the hierarchy relating feature constraints to semantic role definitions follows. Because each subordinate role inherits all features of superordinate roles, each subordinate role can be thought of as an instance of the superordinate role, with the additional characterization given by that role's constraining feature(s). Although the following augmented VN roleset contributes to the construction of an ISO standard, this standard is still under development and likely will not include some roles listed in italics below. The italicized roles, as defined here, are specific to certain

---

[3]Roles that will be maintained for VerbNet only are labeled with a 'VN' prefix; these roles will not be included in the planned ISO standard due to their specificity.

[4]Following James Allen, personal communication, Nov. 2009.

event types or defined according to internal properties of participants, potentially making them overly specific for an ISO standard. In several cases, VN will be adopting the LIRICS definition of a role; these are attributed to LIRICS below. The revised roleset is as follows:

- **Participant**: entity involved in a state or event.
- **Actor**: Participant that is the instigator of an event.
- **Cause**: Actor in an event (that may be animate or inanimate) that initiates the event, but that does not act with any intentionality or consciousness; it exists independently of the event. -LIRICS
- **Agent**: Actor in an event who initiates and carries out the event intentionally or consciously, and who exists independently of the event. -LIRICS
- **Co-Agent**: *Agent who is acting in coordination or reciprocally with another agent while participating in the same event (specific to events with symmetrical participants).*[5]
- **Stimulus**: *Cause in an event that elicits an emotional or psychological response (specific to events of perception).*
- **Undergoer**: Participant in a state or event that is not an instigator of the event or state
- **Instrument**: Undergoer in an event that is manipulated by an agent, and with which an intentional act is performed; it exists independently of the event. -LIRICS
- **Theme**: Undergoer that is central to an event or state that does not have control over the way the event occurs, is not structurally changed by the event, and/or is characterized as being in a certain position or condition throughout the state. -revised from LIRICS
- **Co-Theme**: *Theme that participates in an event or state with another Theme; both participate equally (thereby distinguishing this role from Pivot; specific to events with symmetrical participants).*
- **Pivot**: Theme that participates in an event with another theme unequally. Pivot is much more central to the event (thereby distinguishing it from Co-Theme).
- **Topic**: *Theme characterized by information content transferred to another participant (specific to events of communication).*
- **Patient**: Undergoer in an event that experiences a change of state, location or condition, that is causally involved or directly affected by other participants, and exists independently of the event. -LIRICS
- **Co-Patient**: *Patient that participates in an event with another patient, both participate equally in the event (specific to events with symmetrical participants).*
- **Experiencer**: *Patient that is aware of the event undergone (specific to events of perception).*
- **Attribute**: Undergoer that is a property of an entity or entities, as opposed to the entity itself. -revised from

LIRICS

- **Beneficiary**: Undergoer in a state or an event that is (potentially) advantaged or disadvantaged by the event or state. -LIRICS
- **Place**: Participant that represents the state in which an entity exists. -revised from LIRICS
- **Location**: *Place that is concrete.*
- **Source**: Place that is the starting point of action; exists independently of the event. -revised from LIRICS
- **Material**: *Patient that exists at the starting point of action (inheritance from Source), which is transformed through the event into a new entity; concrete or abstract.*
- **Initial_Location**: Source that indicates the concrete, physical location where an event begins or a state becomes true. -revised from LIRICS
- **Goal**: Place that is the end point of action and exists independently of the event. -revised from LIRICS
- **Destination**: *Goal that is a concrete, physical location.*
- **Recipient**: *Destination that is animate.*
- **Result**: *Goal that comes into existence through the event.*
- **Product**: *Result that is a concrete object.*
- **Time**: Participant that indicates an instant or an interval of time during which a state exists or an event took place. -LIRICS
- **Initial_Time**: Time that indicates when an event begins or a state becomes true. -LIRICS
- **Final_Time**: Time that indicates when an event ends or a state becomes false. -LIRICS
- **Frequency**: Number of occurrences of an event within a given time span. -LIRICS
- **Duration**: Length or extent of time. -LIRICS
- **Value**: *Place along a formal scale.*
- **Extent**: *Value indicating the amount of measurable change to a participant over the course of the event.*
- **Asset**: *Value that is a concrete object.*

Again, several of these roles are unique to certain classes of events. However, by defining these roles according to their position in the hierarchy, one can opt to use a fine-grained role's superordinate role for tasks that benefit from a thematic roleset that is generalizable to all verbs. This not only provides flexibility in how VN can be used for diverse applications, it also allows VN to map easily to other semantic resources that may have differences in granularity, including the planned ISO standard. Notably, the roles *Predicate* and *Proposition* have been removed from the proposed thematic roleset because they are syntactically motivated and cannot be situated in the hierarchy using semantic constraints.

IV. DISCUSSION

This set of roles, organized into a hierarchy, is able to achieve an ideal balance between the strength of LIRICS generalizable roles that are applicable across all verbs and the advantages of finer-grained roles that can help to distinguish VN classes. This can be demonstrated by examining how simple examples would be annotated using the LIRICS framework and the revised VN roles. Consider the following examples,

---

[5]For all roles containing symmetrical participants, we assume that the ordering is non-arbitrary and may be pragmatically determined and/or relate to figure-ground considerations; however, the usefulness of this distinction may be application-specific; for further description of how and why 'Co'-roles are used, see the Discussion Section.

which list both coarse-grained and fine-grained levels of VN annotation where there is a distinction:

1) He talked about politics.
   LIRICS: He$_{AGENT}$ talked$_{RELATION}$ about_politics$_{THEME}$
   VN 1 (coarse-grained): He$_{AGENT}$ talked$_{RELATION}$ about_politics$_{THEME}$
   VN 2 (fine-grained): He$_{AGENT}$ talked$_{RELATION}$ about_politics$_{TOPIC}$

2) He sent the letter to Mary.
   LIRICS: He$_{AGENT}$ sent$_{RELATION}$ the_letter$_{THEME}$ to_Mary$_{GOAL}$
   VN 1 (coarse-grained): He$_{AGENT}$ sent$_{RELATION}$ the_letter$_{THEME}$ to_Mary$_{GOAL}$
   VN 2 (fine-grained): He$_{AGENT}$ sent$_{RELATION}$ the_letter$_{THEME}$ to_Mary$_{RECIPIENT}$

3) The contractor builds houses.
   LIRICS: The_contractor$_{AGENT}$ builds$_{RELATION}$ houses$_{RESULT}$
   VN 1 (coarse-grained): The_contractor$_{AGENT}$ builds$_{RELATION}$ houses$_{RESULT}$
   VN 2 (fine-grained): The_contractor$_{AGENT}$ builds$_{RELATION}$ houses$_{PRODUCT}$

4) I own twelve oxen.
   LIRICS: I$_{PIVOT}$ own$_{RELATION}$ twelve_oxen$_{THEME}$
   VN: I$_{PIVOT}$ own$_{RELATION}$ twelve_oxen$_{THEME}$

5) John collaborated with Paul on the task.
   LIRICS: John$_{AGENT}$ collaborated$_{RELATION}$ with_Paul$_{PARTNER}$ on_the_task$_{THEME}$
   VN: John$_{AGENT}$ collaborated$_{RELATION}$ with_Paul$_{CO-AGENT}$ on_the_task$_{THEME}$

6) The tourists admired the paintings.
   LIRICS: The_tourists$_{PIVOT}$ admired$_{RELATION}$ the_paintings$_{THEME}$
   VN 1 (coarse-grained): The_tourists$_{PATIENT}$ admired$_{RELATION}$ the_paintings$_{CAUSE}$
   VN 2 (fine-grained): The_tourists$_{EXPERIENCER}$ admired$_{RELATION}$ the_paintings$_{STIMULUS}$

As these examples demonstrate, the coarse-grained roles of VN largely overlap with the roles of LIRICS. It is perhaps this level of granularity, shared by both resources, that will prove to be the ideal level of granularity for the ISO standard. The final two examples show points where the two annotation schemas are quite different; the motivations for maintaining these differences in VN are discussed below.

First consider example (5). Verbs such as *chat, cooperate,* and *speak* correspond to events that usually involve two volitional participants, as in: 'Susan chatted/cooperated with Rachel.' Currently, VN uses the labels *Actor 1* and *Actor 2* to refer to each of these participants. In typical usage, *Actor 1* is the subject of the verb and *Actor 2* occurs in the oblique (e.g. 'with Rachel'). In theory, these labels capture the notion of two volitional actors involved in a single event, where one seems to be a true agent with pragmatic focus (*Actor 1*), while the other participant (*Actor 2*) fulfills the same agentive qualities (volitional instigator) without pragmatic focus.

While LIRICS does not have an exact mapping to *Actor 1* and *Actor 2*, it does have the complementary roles of *Agent* and *Partner*. In the LIRICS framework, an *Agent* is defined as a 'participant in an event who initiates and carries out the event intentionally or consciously, and who exists independently of the event,' while a *Partner* is defined as a 'participant in an event who is intentionally or consciously involved in carrying out the event, but who is not the principal agent of the event, and who exists independently of the event.' Upon examining this distinction between *Agent* and *Partner*, the LIRICS terms were preferable for the following reasons: (1) the labels *Agent* and *Partner* more clearly indicate that there are differing levels of agency between the two roles; (2) using the term *Actor 1* fails to illustrate that the argument is essentially an agent.

Although the use of the LIRICS roles seemed an ideal solution, a practical problem arose in implementing these roles. An adoption of *Agent* and *Partner* produces a potentially confusing incongruency among VN roles: parallel to *Actor 1* and *Actor 2*, VN has the roles *Theme 1, Theme 2, Patient 1* and *Patient 2*. *Theme 1* and *Theme 2*, for example, are used for verbs such as *border, coincide*, and *have*, which denote events that may involve two themes: 'Italy-*Theme 1* borders France-*Theme 2*.' The relationship between the two themes is analogous to the relationship between *Agent* and *Partner*: there is a pragmatically focused theme (*Theme 1*) and a secondary theme (*Theme 2*). In order to accommodate the parallel nature of these roles such that these relationships would be easily understandable to users, it is preferable to maintain the concept behind LIRICS *Agent* and *Partner*, but adjust the labels to *Agent* and *Co-Agent, Theme* and *Co-Theme* and finally *Patient* and *Co-Patient*. Perhaps more importantly, the adoption of *Agent* and *Partner* could also produce ambiguity among roles because there is the possibility that a sentence could involve both a *Partner* to an *Agent* and a *Partner* to a *Theme* or *Patient*, leading to two ambiguous *Partner* arguments. The 'Co-' terminology allows each argument to be easily distinguished by VN users.

Each of the 'Co-' roles should be thought of as semantically identical to the *Agent, Patient* and *Theme* roles. However, for certain applications, it may be useful to distinguish between these two participants in terms of a figure-ground relationship [21]; therefore, VN maintains distinct labels for each participant in an event. Which participant is the bare role and which is the 'Co-' role will be assigned based on syntactic considerations: the 'Co-' role will be assigned to the participant expressed as an oblique or prepositional argument. If this distinction is not useful for the application at hand, then all 'Co-' arguments can be deterministically converted to the bare role label *Agent, Patient* or *Theme* if desired.

Secondly, consider example (6). In the LIRICS framework, *Pivot* is a 'participant in a state that is characterized as being in a certain position or condition throughout the state, and that has a major or central role or effect in that state'. This

Old VN role	New VN role	LIRICS role	Old VN (cont.)	New VN (cont.)	LIRICS (cont.)
Actor	Agent	Agent	Patient	Patient	Patient
Actor 1	Agent	Agent	Patient1	Patient	Pivot
Actor 2	Co-Agent	Partner	Patient 2	Co-Patient	Patient
Agent	Agent	Agent	Predicate	-	-
Asset	Asset	Amount	Product	Result/Product	Result
Attribute	Attribute	Attribute	Proposition	-	-
Beneficiary	Beneficiary	Beneficiary	Recipient	Recipient	Goal
Cause	Cause	Cause	Source	Source/Initial_Loc	Source/Initial_Loc.
Destination	Destination	Final_Location	Stimulus	Stimulus	Theme
Experiencer	Experiencer	Pivot	Theme	Theme	Theme
Extent	Extent	Amount/Distance	Theme1	Theme	Theme
Instrument	Instrument	Instrument	Theme2	Co-Theme	Theme
Location	Location	Location	Time	Time	Time
Material	Material	Source	Topic	Topic	Theme

TABLE II
REVISED VERBNET ROLES IN COMPARISON TO LIRICS.

definition can apply, as seen in example (6), to arguments that would be considered *Experiencers* in VN. Although *Pivot* will be incorporated into VN, in the VN framework the definition of *Pivot* is somewhat narrower: *Pivot* strictly applies to a *Theme* accompanied by a secondary *Theme* that is much less central to the event. This definition arose out of careful consideration of how the *Theme* and *Co-Theme* roles would be used in replacing existing *Theme 1, Theme 2* roles. In certain VN classes, it was found that the current usage of *Theme 1* was inconsistent: for verbs in the Own and Require classes, using *Theme 1* to refer to the 'possessor' or 'requirer' seemed to obscure an important distinction between this type of participant and other *Theme 1* arguments, wherein the *Theme 1* is primarily being located (e.g. 'Italy' in 'Italy borders France'). For verbs in the Own and Require classes, *Theme 1* was not located; rather these verbs involve a state of ownership or need. Therefore, for the Own and Require classes, we did not adopt *Theme* and *Co-Theme* to replace *Theme 1* and *Theme 2*. Instead, we chose to adopt the label *Pivot*, with a somewhat revised definition, for participants in a state of ownership or need, and to use *Theme* to refer only to the owned or needed participants (as exemplified in (4)).

One may wonder why *Experiencer* was not used as the argument label for the participant in a state of ownership or need in a fashion similar to the usage of this role seen in example (6). In the VN framework, *Experiencer* differs from *Pivot* in that *Experiencers* undergo a change of state that the participant is aware of in reaction to a particular *Stimulus/Cause*, whereas *Pivots* continue in an unchanged state and are not reacting to another participant. In example (4), 'I own twelve oxen', the oxen clearly do not cause the state of ownership, instead they simply exist in a related state of being owned. However, in example (6), the paintings do evoke a change from an unspecified state to a state of admiration. Therefore, although the VN usage of *Pivot* is not incompatible with the LIRICS definition of this role, there seems to be a difference in how the two frameworks view certain stative verbs: in VN, the instigation of the state is highlighted through the use of *Experiencer* and *Stimulus*, whereas in LIRICS the continuing, static nature of a state is highlighted through the use of *Pivot* and *Theme*.

A summary of the revised VN roles in comparison to LIRICS roles is found in Table II.

V. CONCLUSION AND FUTURE WORK

This project extends previous work in reconciling LIRICS and VN semantic roles ([14]) by proposing a hierarchical organization of a unified set of roles. As shown, this hierarchy exploits variation in semantic role granularity to utilize the strengths of both projects. The hierarchy will be incorporated into the current version of VN as a reference with the most specific roles listed in each individual class, such that users may select the level of granularity most suited to an individual application. We expect that the hierarchy, in conjunction with VN semantic predicates, will enhance the inferences that may be drawn from VN annotations. By organizing semantic roles hierarchically, relationships between more fine-grained roles, such as *Recipient* and *Destination*, become clearer through their relationship via a higher node, here *Goal*.

Additionally, we believe that the organization of the VN roleset into a hierarchy will facilitate mapping between VN and other semantic resources, such as FrameNet [22], and PropBank. For example, in this research we have shown in the annotation examples that there is a level of the hierarchy that maps almost seamlessly to LIRICS. Other resources can be mapped easily to VN in a similar fashion, given that one level of the hierarchy should be compatible with the granularity of roles used in that resource. After the establishment of an ISO standard set of thematic roles, this flexibility will ensure that VN will be compatible with that standard.

Future work will extend the current hierarchy to incorporate intuitions from other prominent resources by using Semlink, a resource maintained at the University of Colorado that provides mappings between the semantic roles of VerbNet, FrameNet, and PropBank, and, as mentioned previously, provides a corpus in which the verbs have been annotated with VerbNet classes, FrameNet frames, and PropBank rolesets. Our hope is that the set of thematic roles identified here, in-

formed by comparisons to other resources, will ultimately aid in the construction of an ISO set of semantic roles, ensure the compatibility of VN with this standard, and meet the unique needs of various natural language processing applications.

REFERENCES

[1] K. K. Schuler, "Verbnet: A class-based verb lexicon," Ph.D. dissertation, Department of Computer and Information Science, University of Pennsylvania, 2005.

[2] K. Kipper, A. Korhonen, N. Ryant, and M. Palmer, "A large-scale classification of english verbs." *Language Resources and Evaluation Journal*, vol. 42, pp. 21–40, 2008.

[3] V. Petukhova and H. Bunt, "Lirics semantic role annotation: Design and evaluation of a set of data categories." in *Proceedings of the sixth international conference on language resources and evaluation (LREC 2008)*, Tilburg, 2008.

[4] D. Dowty, "Thematic proto-roles and argument selection." *Language*, vol. 67, pp. 547–619, 1991.

[5] J. Sowa, *Knowledge representation: logical, philosophical and computational foundations*. Pacific Grove: Brooks/Cole, 2000.

[6] H. Bunt, V. Petukhova, and A. Schiffrin, "Lirics deliverable d4.4. multilingual test suites for semantically annotated data. http://lirics.loria.fr."

[7] B. Levin, *English Verb Classes and Alternations: A Preliminary Investigation*. University of Chicago Press, 1993.

[8] R. Swier and S. Stevenson, "Unsupervised semantic role labelling," in *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, Barcelona, Spain, 2004, pp. 95–102.

[9] L. Shi and R. Mihalcea, "Putting pieces together: Combining framenet, verbnet and wordnet for robust semantic parsing," *Computational Linguistics and Intelligent Text Processing*, pp. 100–111, 2005.

[10] M. Swift, "Towards automatic acquisition from verbnet for spoken dialog processing," in *Proceedings of Interdisciplinary Workshop on the Identification and Representation of Verb Features and Verb Classes*, Saarbrucken, Germany, 2005.

[11] B. Crabbé, M. Dzikovska, W. de Beaumont, and M. Swift, "Increasing the coverage of a domain independent dialogue lexicon with verbnet," in *Proceedings of the Third Workshop on Scalable Natural Language. Understanding (ScaNaLU '06)*, 2006, pp. 25–32.

[12] N. Badler, M. Palmer, and R. Bindiganavale, "Animation control for realtime virtual humans," *Communications of the ACM*, vol. 42, no. 7, pp. 65–73, 1999.

[13] K. Kipper and M. Palmer, "Representation of actions as an interlingua," in *Proceedings of the Third Workshop on Applied Interlinguas, held in conjunction with ANLP-NAACL 2000*, 2000.

[14] C. Bonial, S. Brown, W. Corvey, M. Palmer, V. Petukhova, and H. Bunt, "Exploratory comparison of thematic roles in verbnet and lirics," in *Proceedings of the Sixth Joint ISO - ACL SIGSEM Workshop on Interoperable Semantic Annotation (ISA-6)*, 2011, pp. 39–44.

[15] S. Yi, E. Loper, and M. Palmer, "Can semantic roles generalize across genres?" in *Proceedings of the HLT/NAACL-2007*, Rochester, 2007, pp. 548–555.

[16] E. Loper, S. Yi, and M. Palmer, "Combining lexical resources: Mapping between propbank and verbnet," in *Proceedings of the Seventh International Workshop on Computational Semantics (IWCS-7)*, Tilburg, 2007.

[17] M. Palmer, D. Gildea, and P. Kingsbury, "The proposition bank: An annotated corpus of semantic roles," *Computational Linguistics*, vol. 31, no. 1, pp. 71–105, 2005.

[18] P. Merlo and L. Van Der Plas, "Abstraction and generalization in semantic role labels: Propbank, verbnet or both?" in *Proceedings of the 47th Annual Meeting of the ACL and 4th IJCNLP of the AFNLP*, Suntec, 2009, pp. 288–296.

[19] M. Palmer, "Semlink: Linking propbank, verbnet and framenet," in *Proceedings of the Generative Lexicon Conference, GenLex-09*, Pisa, Italy, 2009.

[20] R. J. Van Valin and R. LaPolla, *Syntax: structure, meaning, and function*. Cambridge University Press, 1997.

[21] L. Talmy, *Toward a Cognitive Semantics*. Cambridge: MIT Press, 2000.

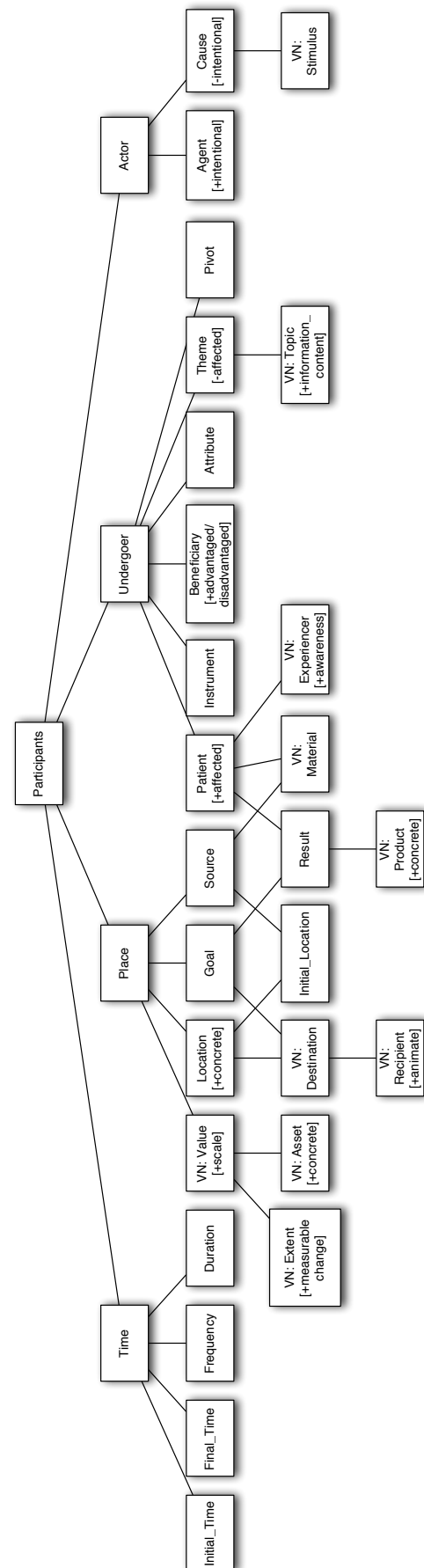[22] C. Fillmore and C. Baker, *Oxford Handbook of Linguistic Analysis*. Oxford University Press, 2010, ch. A Frame Approach to Semantic Analysis.

Fig. 1. Hierarchy of Unified Semantic Roles